Title
Sensorimotor Integration in Speech Production

Author
Kothare, Hardik

Publication Date
2020

Peer reviewed|Thesis/dissertation

Sensorimotor Integration in Speech Production

by

Hardik Kothare

DISSERTATION

Submitted in partial satisfaction of the requirements for degree of

DOCTOR OF PHILOSOPHY

in

Bioengineering

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO
AND
UNIVERSITY OF CALIFORNIA, BERKELEY

Approved:

DocuSigned by:
Srikantan Nagarajan
4C21BC1B9EB145E...

Srikantan Nagarajan

Chair

DocuSigned by:
John F Houde
[illegible]

John F Houde

DocuSigned by:
Keith Johnson
758CD264D5E4489...

Keith Johnson

Committee Members

To India, my motherland.

Acknowledgements

In his book 'So Long and Thanks for All the Fish', the author Douglas Adams says, "They were not the same eyes with which he had last looked out at this particular scene, and the brain which interpreted the images the eyes resolved was not the same brain. There had been no surgery involved, just the continual wrenching of experience". If I were to describe the process of getting a PhD, it would be hard to find a more appropriate phrase than 'the continual wrenching of experience'.

The principal reason why I could make it through this tempestuous process was the unwavering support of my advisors, Dr. Srikantan Nagarajan and Dr. John Houde. Sri's enthusiasm for research and science is infectious. His cheerful optimism encouraged me to look at the silver lining in every situation. With his unparalleled patience, John taught me that even the biggest problem has a solution if we get back to the basics. He is not only a brilliant scientist but also one of the kindest people one could ever work with. The BIL and SNL labs at UCSF are fortunate to have such scientific father figures worthy of emulation as Sri and John.

I would like to express my sincere gratitude to Dr. Keith Johnson who provided scientific guidance and a much-needed phonetic perspective to my projects as part of my qualifying exam and dissertation committees. I would also like to thank the rest of the members of my qualifying exam committee, Dr. Christoph Schreiner, Dr. Janine Lupo and the ever-witty Dr. Steven Cheung, who helped mould and refine my thinking as an engineer and a scientist. I have a deep sense of obligation towards Dr. Siddharth Kharkar for introducing me to Sri back in 2014 and setting me on this wonderful career trajectory that I have thoroughly enjoyed so far. I thank Dr.

Duygu Tosun, Dr. Susan Noworolski and Dr. Jennifer Henderson-Sabes for giving me the chance to hone my teaching skills.

Apart from all these academic mentors, I have been lucky to get the chance to work alongside a constellation of research superstars. This dissertation could come to fruition only because of their mentorship, guidance, help and friendship. Foremost among these teammates have been Danielle Mizuiri and Susanne Honma whose hard work and perseverance alone has been responsible for the success of every single lab member. Words are insufficient to express my thanks to these two pillars of the lab. The countless hours dedicated by Kamalini Ranasinghe in tutoring me in statistics and neuroscience have been some of the finest learning experiences I have had outside of a classroom. Carly Demopoulos is a steady champion of graduate students and junior lab members; I am grateful to her for a lot of things and especially for showing me the ropes when it comes to peer reviewing. Candid conversations with the generous and genial Karuna Subramaniam always helped bring some warmth straight from *Camp Nou* to the frozen environs of *HSE-608* a.k.a. '*Alaska*'; she has been a trustworthy comrade both in and out of the lab. I am deeply indebted to Vikram Ramanarayanan for his steadfast belief and support at every step of my career over the past so many years. The compassionate Zarinah Agnew taught me to keep in mind that being a caring individual is key to being a good researcher; I will always cherish this idea. I pay tribute to Leighton Hinkley for being the wise Gandalf to our merry fellowship of graduate students who always depend on his help and support for all things MEG. Carrie Niziolek and her tremendously helpful repositories of analysis pipelines have served as the intellectual North Star throughout my research career at UCSF. My heartfelt thanks to Anne Findlay and Corby Dale for always being on top of data management issues and always being so

accommodating when it came to troubleshooting bugs and errors related to the various scripts I used for MEG analysis. The jovial Phiroz Tarapore has had an impactful role in always making me feel welcome in the lab since the early days.

Megan Thompson has been an exemplar of PhD studentship. Her sense of humour kept us all going on long days in the lab. A big high five to my partner-in-research Inez Raharjo for also reaching this accomplishment alongside me. Our research projects mirrored each other right through all the ups and downs and Inez has been an invaluable research partner and friend through thick and thin. Cheers to our camaraderie! Yijing Gao and Jessica Gaines will surely ensure that the sense of fellowship among graduate students in the lab stays alive. I have learnt so much from Sarah Bakst and Noriko Tanigawa during some of the most delightful (scientific and non-scientific) conversations at academic conferences around the globe. Chang Cai managed to raise the bar for productivity and diligence so high that I am unsure whether I will ever be able to reach those levels. Ben Parrell's inputs on my projects have played a vital role in bringing this dissertation to its final form. Valentina Borghesani's multitasking and collaborative spirit has been truly inspirational. In addition to expertise in brain stimulation, Nina Sardesh also had the gift of planning lab lunches where we all had many a stimulating conversation about science, life and beyond. I got to learn so much from both at-work and after-work conversations with Katsuaki Kojima, Lucia Bulubas, Ana Souza, Elke De Witte, Nayara Mota, Coleman Garrett, Xihe Xie, Ashley Tay, Cathra Halabi, Kyunghee Kim, Kwang Kim and Jessie Chen. They have had a tremendous impact on shaping my outlook towards research. Abhishek Bhutada and Sarah Schneider deserve a special shout-out for their help in the smooth coordination of the Laryngeal Dystonia project. One of the most effective ways of learning was getting the chance to mentor a

fantastic bunch of high school students and undergraduate students over the past so many summers. In particular, David Klein, Anantajit Subrahmanya, Oscar Rausis, Jay Shroff, Isha Sai, Yash Shroff and Srivatsan Tennathur have made significant contributions to multiple projects big and small that spanned weeks, months and years.

I am grateful to the Bioengineering Association of Students (BEAST) and to the executive committee of the UC Berkeley-UCSF Graduate Program in Bioengineering for giving me the opportunity to serve in multiple leadership roles. Special thanks to SarahJane Taylor, Lisa Cabahug and Victoria Ross for their colossal help in navigating through administrative paperwork. I would also like to thank the Discovery Fellows Program and all its donors, especially Judith and George Marcus, for their belief in UCSF PhD students as advocates for basic and biomedical sciences.

There are countless others to thank: my co-authors, my friends, every teacher who has taught me everything I know.

I am where I am only because of my family's encouragement and backing. Thank you, Aai and Baba, for the gift of birth, for inculcating in me a sense of scientific curiosity and for always standing by me. I am eternally grateful for everything. Thank you, Amruta, for being such a loving big sister and also for being my first teacher even before I was old enough to go to school. This PhD journey has been possible **only because of** the love, support, belief and companionship of my husband, Saurabh. He saw me through every panic attack and bout of impostor syndrome and made sure that we celebrated every tiny accomplishment. He lent a sympathetic ear when I

expressed frustration over (now) trivial matters like bugs in my code and noisy data. He has been a dependable proofreader when my eyes got tired of reading the same text over and over again. He has made me a better scientist. I only have him to thank for keeping me sane through this journey. Thank you, Saurabh, for everything. I dedicate this work to you.

यथा चित्तं तथा वाचो यथा वाचस्तथा क्रिया…

As is the mind, so is the speech; as is the speech, so is the action…

-A Sanskrit proverb

Abstract

Sensorimotor Integration in Speech Production

Hardik Kothare

When we speak, there's a complex sequence of events occurring within us. A symphony, if you will, with the brain as the conductor; the respiratory muscles pushing air out of our bellow-like lungs; this exhaled air setting the vocal cords into vibration like the strings of a musical instrument; the tongue, lips, nose, teeth and jaw acting as an ensemble and obstructing this air flow to produce speech sounds. The smooth execution of this sequence of events requires the human brain to monitor sensory feedback during speech, correct for any errors and learn from any past errors. This phenomenon is called sensorimotor integration and is essential for efficient speech motor control. Various theoretical and computational models of speech production explain how sensorimotor integration occurs but many aspects of this process still remain underexplored. This dissertation starts by investigating how sensorimotor learning of vowels depends on the complex relationship between articulatory dimensions and acoustic space. Specifically, formant adaptation or response to altered formant frequency feedback depends on the direction of the shift in two-dimensional F1-F2 vowel space. Using magnetoencephalographic imaging, I then investigate how sensorimotor integration is affected during speaking in a voice disorder called Laryngeal Dystonia (or Spasmodic Dysphonia). Significant differences in neural activity at various nodes of the speech motor control network were observed in patients with Laryngeal Dystonia at various time points around the act of phonation. Lastly, the dissertation describes cortical dynamics of the speech motor control network in a neurodegenerative disorder affecting speech and language called the non-fluent

variant of Primary Progressive Aphasia (nfvPPA). These patients have significant motor speech impairments which were investigated using a pitch perturbation experimental paradigm. Neural and behavioural results showed that sensorimotor integration is severely impacted in patients with nfvPPA. Taken together, the work in this dissertation aims to help inform current computational models of speech production and underlines the important role of sensorimotor integration in human speech.

TABLE OF CONTENTS

List of Figures

List of Tables

Chapter 1: Introduction

That human beings are social animals is an indisputable fact. The human ability to communicate with each other enables us to build and strengthen bonds of cohesion with the people surrounding us. The year 2020, with worldwide lockdowns and bouts of social isolation due to a global pandemic, has further highlighted the role that communication plays in our personal, social and professional lives.

For many of us, the primary medium of communication is speech. We transmit our thoughts and ideas to others as sound or vibrations in the air. The production of speech sounds is an extraordinarily complex task that requires the orchestration of nearly 100 muscles. When we speak, we exhale air out of our lungs, thus setting our vocal folds, which are held in a tense position in our voice box or larynx, into vibration. This vibration creates a buzzing sound that is transformed into distinct consonant and vowel sounds by our articulators like the nose, lips, tongue and jaw which together change the shape of our oral and nasal cavities.

Despite the sheer complexity of the task of speech production, human beings are incredibly good at producing intelligible speech. One reason why this is possible is because speech production requires continuous monitoring of sensory feedback. It is not just our intended audience who are listening to our speech; we need to listen to our own speech to ensure that it is steady and fluent. Along with auditory feedback, humans also continuously monitor somatosensory feedback from the articulators involved in speech production. The brain relies on both these kinds of sensory feedback to correct for any speech errors while speaking and also to learn from prior speech

errors. This phenomenon of changing the speech motor output based on sensory feedback is called sensorimotor integration and it is key to error-free speech production.

We have come a long way in understanding how the brain produces and controls speech. Thanks to modern tools like neuroimaging and computational models, we have answers to several questions posed by speech scientists and neuroscientists. However, most of these answers are incomplete. We still have a long way to go before we have unlocked every secret and solved every mystery. This dissertation aims to provide yet another piece of the puzzle in the form of three chapters exploring sensorimotor integration in speech.

In Chapter 2, I explore the phenomenon of sensorimotor adaptation or the ability of the motor control system to adapt to a consistently altered sensory map and modify the motor output accordingly. Formants are frequency components of a speech signal that define vowel sounds. Generally, the first two formants are sufficient to distinguish vowels. When formant frequencies in auditory feedback are consistently altered while a person is speaking, the speaker registers a feedback error and responds by changing the formants of the produced speech (or compensating for the altered feedback) such that the feedback error is reduced. This is called formant adaptation in speech. Chapter 2 demonstrates how compensation in formant adaptation is not strictly oppositional to the applied feedback shift and that adaptation depends on the direction and magnitude of formant alteration in a two-dimensional space defined by the first two formants (F1 and F2).

One approach to examining how sensorimotor integration works is by investigating how it is impacted in patients with impaired speech. Chapters 3 and 4 consider two disorders impacting speech production in patients. In both these chapters I used magnetoencephalography (MEG), a neuroimaging technique that helps detect neural activity by sensing magnetic fields produced by electrical currents in the brain, to observe differences between neural activity in patients and that in healthy controls.

In Chapter 3, I explore neural activity around various phonatory events in patients with Laryngeal Dystonia (LD), a voice disorder that causes spasms in the vocal folds during voiced speech (and thus also called Spasmodic Dysphonia). There have been several studies with conflicting findings about which parts of the Central Nervous System (CNS) are impacted in LD. Most of these studies investigate ***where*** abnormalities exist in the brains of patients with LD. However, as mentioned before, speech production is a rapid process with various parts of the brain, respiratory muscles and articulatory muscles involved at various timepoints during the course of speech production. Surprisingly, there has been no study to investigate ***when*** neural abnormalities arise in patients with LD. MEG provides the temporal resolution required to ask this question. LD patients show abnormal neural activity in various parts of the speech motor control network during motor planning, initiation of phonation and during sustained phonation.

In Chapter 4, I explore sensorimotor integration in patients with the non-fluent variant of Primary Progressive Aphasia (nfvPPA), a neurodegenerative disorder affecting motor speech. Structural neuroimaging studies have shown that patients with nfvPPA exhibit atrophy in the cerebral cortex in regions that are crucial to speech production. Furthermore, these patients also

show aberrant connectivity between the various regions of the brain that form the speech production network. However, not much is known about the functional recruitment of the cortical regions that are part of the speech production network during speech in patients with nfvPPA. Understanding how these patients control speech and integrate sensory feedback with their motor output may help shed light on why motor speech is impacted in nfvPPA. Chapter 4 talks about how patients with nfvPPA are unable to rapidly adjust their motoric output in response to sudden changes to auditory feedback, suggesting deficits in sensorimotor integration. Indeed, differences in neural activity also provide supporting evidence to the hypothesis of abnormal sensorimotor integration.

Together the findings from the chapters in this dissertation pave the way for further exploratory studies by speech scientists, neuroscientists, phoneticians, neurologists, laryngologists and other researchers engaged in the task of decoding how the human brain produces speech.

Chapter 2: Sensorimotor adaptation of speech depends on the direction of auditory feedback alteration

2.1 Abstract

A hallmark feature of speech motor control is its ability to learn to anticipate and compensate for persistent feedback alterations, a process referred to as sensorimotor adaptation. Because this process involves adjusting articulation to counter the perceived effects of altering acoustic feedback, there are a number of factors that affect it, including the complex relationship between acoustics and articulation and non-uniformities of speech perception. As a consequence, sensorimotor adaptation is hypothesised to vary as a function of the direction of the applied auditory feedback alteration in vowel formant space. This hypothesis was tested in two experiments where auditory feedback was altered in real time, shifting the frequency values of the first and second formants (F1 and F2) of participants' speech. Shifts were designed on a subject-by-subject basis and sensorimotor adaptation was quantified with respect to the direction of applied shift, normalised for individual speakers. Adaptation was indeed found to depend on the direction of the applied shift in vowel formant space, independent of shift magnitude. These findings have implications for models of sensorimotor adaptation of speech.

2.2 Introduction

The intent of human speech communication is to transmit information or express emotions and it relies on successful audition by the intended recipient. However, speech communication also involves continuous monitoring of auditory feedback by the speaker (Lee, 1950; Fairbanks, 1954; Houde et al., 2002; Houde and Nagarajan, 2011; Guenther, 2016). The purpose of this monitoring may be to distinguish between self-generated and externally-generated speech

(Korzyukov et al., 2017; Subramaniam et al., 2018) and also to detect and correct speech errors (Levelt, 1983; Levelt, 1993).

Speech feedback monitoring has been examined in numerous experiments showing that real-time changes in auditory feedback cause speakers to modify speech production in a compensatory manner. Speakers have been shown to compensate for changes in fundamental frequency (F0) (Burnett et al., 1998; Jones and Munhall, 2000), vowel formant frequency (Houde and Jordan, 1998; Purcell and Munhall, 2006) and fricative centroid frequency (Shiller et al., 2009). Persistent alterations to sensory feedback cause long-term changes in motor behaviour where sensory feedback alterations are anticipated. This learnt compensatory process is called sensorimotor adaptation in speech and has been studied widely over the last couple of decades (Caudrelier and Rochet-Capellan, 2019).

Since sensorimotor adaptation involves adjusting articulation to counter the perceived effects of altering acoustic feedback, there are a number of factors that affect it. First, there is the complex relationship between acoustics and articulation (Stevens, 1968; Stevens, 1998; Johnson, 2012; Gick et al., 2013). In particular, we can consider the muscles, articulators and motor programs involved in changes to vowel height and vowel backness that create acoustic changes in F1 and F2 respectively. Articulatory changes in vowel height and backness are achieved by relatively independent control of different tongue muscles (Takano and Honda, 2007; Gick et al., 2013). For instance, the anterior genioglossus lowers and retracts the tongue tip and blade to produce low back vowels, the middle genioglossus can contract to lower the tongue body and pull it forward to produce low front vowels, whereas the posterior genioglossus can contract to pull the

tongue root forward to produce high front vowels (Gick et al., 2013). Therefore, the acoustic changes needed to counter different acoustic feedback alterations are likely accomplished by distinct articulatory motor programs that differ in their degrees of adaptability. In this way, adaptation response may depend on the changes in perceived vowel backness and vowel height created by the altered auditory feedback. Additionally, depending on the direction of the response to the altered auditory feedback, non-linearities in the acoustic-articulatory mapping (Stevens, 1989) would also create different changes in the magnitude of somatosensory and auditory feedback. This, in turn, would affect the magnitude of adaptation response since it is thought to be determined, in part, by the weighting of conflicting somatosensory and auditory feedback about the feedback error driving the adaptation (Lametti et al., 2012). Second, factors like categorical perception create non-uniformities in speech perception (Pisoni and Remez, 2005). The acoustic change needed to counter feedback alteration may differ depending on how acoustic changes are perceived in the neighbourhood of the speech sound being produced. Indeed it has been shown that shifts nearer to vowel category boundaries cause enhanced feedback compensation responses (Niziolek and Guenther, 2013) and that individual differences in perceptual categories influence the magnitude of sensorimotor adaptation (Daliri and Dittman, 2019). The current study sought to test the following hypothesis: As measured by its magnitude and orientation, the extent to which the adaptation response opposes the applied feedback shift varies as a function of the direction of the applied feedback shift in vowel formant space.

2.3 Methods

Eighteen (10 female) participants were recruited for the study (average age = 28.83, standard deviation =11.82 years). Data on participants' linguistic background was not collected prior to

the experiment but a survey on linguistic background was sent out to the participants after the experiments. 10 out of the 18 participants responded to this survey. Four respondents identified as trilingual or multilingual, three of them identified as bilingual and the remaining three identified as monolingual English speakers. Six respondents indicated that they were most fluent in English and the remaining four stated that they were as fluent in English as another language. Six respondents said that English was the first language they were exposed to and all respondents said that English was their language of formal education. None of the participants reported any hearing loss or a history of speech and language deficits. The procedures of testing were explained to the participants and informed consent was obtained. The study was approved by the Committee on Human Research of the University of California, San Francisco.

Participants were seated in an audio booth (Eckoustic C-14A LP Mod. Rev., Eckel Industries of Canada, Morrisburg, Ontario, Canada) and were wearing a headset microphone (MicroMic C520, AKG Acoustics, Vienna, Austria). Baseline vowel formant frequency values for 10 vowels (/ɛ/, /ɪ/, /i/, /e/, /æ/, /ɑ/, /o/, /ɔ/, /ʊ/ and /u/) were collected for all individuals using these prompt words respectively: ‘bet’, ‘fit’, ‘meet’, the first part of the diphthong in ‘late’, ‘bad’, ‘car’, ‘hope’, ‘bought’, ‘book’ and ‘pool’, three samples for each vowel. Formants were tracked using Linear Predictive Coding (LPC) and optimal LPC order for tracking was determined on a subject-by-subject basis (Schuerman et al., 2017). Vowel formant frequency values for every subject were calculated by averaging values for every vowel across the three samples. These average values were used as the basis for determining the altered feedback in Sec. 2.3.A and Sec. 2.3.B.

Two experiments were conducted on separate days and are described in detail below. Fourteen of the eighteen participants were able to participate in both experiments and the order of Experiment 1 and Experiment 2 was randomised for these subjects. The median difference between the date of Experiment 1 and Experiment 2 for these participants was 6 days. Two subjects could participate only in Experiment 1 and two could participate only in Experiment 2 due to scheduling conflicts.

During both experiments, participants wore circumaural headphones (DT 770 PRO 250 OHM, beyerdynamic, Heilbronn, Germany). Speech signals from the microphone were routed to a computer (Optiplex 9020 SFF, Dell, Round Rock, TX) with a Delta 44 sound card (M-Audio, Cumberland, RI) via a preamplifier (HR-MP2, Radio Design Labs, Prescott, AZ). These signals were analysed and re-synthesised by a real-time feedback alteration tool called Feedback Utility for Speech Processing (FUSP) designed by author JFH. This tool employs sinusoidal synthesis methods (Quatieri and McAulay, 1986) as described in previous studies (Katseff et al., 2012).

2.3.1 Experiment 1

This experiment consisted of six cases. Each case started with an unaltered block of 10 trials, followed by an altered block of 50 trials, and concluded with an unaltered washout block of 30 trials. During every trial, subjects were prompted to read the nonsense word ‘bep’ (/bɛp/). During the altered block, the first two formants were shifted from those of the subject’s production of /ɛ/ to those of a different vowel sound, so that subjects heard formant frequency values corresponding to a perceivably-different vowel sound, depending on the case (see Fig. 2.1A and Fig. 2.1B). The order of these six cases was randomised across participants. The shifts applied

were based on individual baseline vowel formant frequency values collected prior to the experiment as described above. In particular, the shift in F1 was calculated as the difference between F1 of /ɛ/ and F1 of either /ɪ/ (as in fit), /i/ (as in meet), /e/ (as in the first part of the diphthong in late), /æ/ (as in bad), /ɑ/ (as in car) or /u/ (as in pool), depending on the case; the shift in F2 was calculated similarly. Thus, the magnitude and direction of shifts varied across cases and subjects (Fig. 2.3, Supp Fig. A.1).

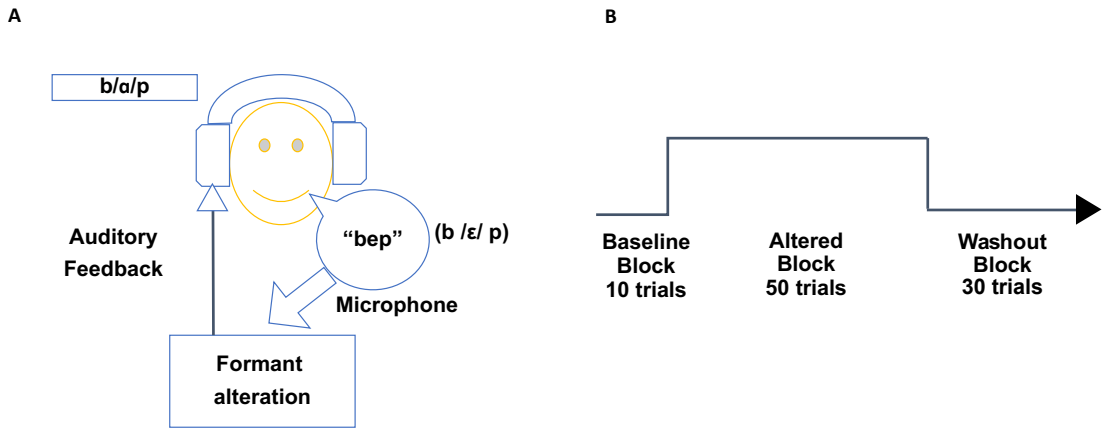


Fig. 2.1 Schematics of the experimental setup and the experimental design
A. *Schematic of the experimental setup. Participants spoke the syllable 'bep' into a microphone when prompted; this input signal from the microphone was relayed to a digital signal processing unit which altered the first two formants (F1 and F2) during the altered block. This processed signal (altered or unaltered, depending on the block) was then played through headphones.* ***B.*** *Schematic of the experimental design. Each case consisted of 90 trials which were divided into three blocks. The initial 10 trials were part of the baseline block where the participants' feedback was unaltered, i.e. they heard what they said. This block was followed by the altered block of 50 trials where the formants from the speech signal were altered and fed back through the headphones. The last block was the washout block of 30 trials where formant alteration was turned off and feedback returned to normal.*

2.3.2 Experiment 2

The sequence of blocks in this experiment was exactly the same as Experiment 1. The auditory feedback in the altered block was shifted towards the six target vowels from Experiment 1 but

the magnitude of shift was just 50 Hz in all six cases. Again, the direction of applied shift was determined based on the baseline vowel formant frequency values for individual subjects collected prior to both experiments.

2.3.3 Extraction of formant values

Each trial was visually inspected and played using custom-built speech analysis software called Wave Viewer (https://github.com/SpeechNeuroscienceLab/Wave-Viewer) written in MATLAB (The MathWorks, Inc., Natick, MA). This software analyses and displays the raw waveform as well as the spectrogram, formant tracks, time course of pitch and amplitude of the speech signal. Trials were marked bad if formants were tracked incorrectly on the spectrogram or if subjects phonated multiple times within a single trial. All trials that were not marked as bad were considered good trials. For each good trial, formant tracks for F1 and F2 were extracted for the entire duration of vowel phonation, excluding the transitions to the flanking consonants, and mean F1 and F2 values were calculated. All good trials from the last 20 trials of the altered block were taken into consideration to determine adaptation response, as done in other sensorimotor adaptation studies (Jones and Munhall, 2000; Kitago et al., 2013).

2.3.4 Vocal tract length normalisation

According to the source-filter theory of speech production (Fant, 1960), the acoustic properties of speech are a function of the shape and length of the supralaryngeal vocal tract. To account for differences in participants' vocal tract lengths based on gender and age and to ensure that behavioural differences are not influenced by speaker differences, we performed the ΔF method

of vocal tract length normalisation (Johnson, 2018). The ΔF value represents average spacing between vowel formants and is related to the talker's vocal tract length by the following formula:

$$\text{Vocal Tract Length} = \frac{c}{2\Delta F}$$

where c = 35000cm/s or the speed of sound in warm, moist air.

Each speaker's ΔF value was calculated using the following formula:

$\Delta F = \frac{1}{mn}\sum_{j}^{m}\sum_{i}^{n}[\frac{F_{ij}}{i-0.5}]$, where i = formant number and j = token number

For each participant, two formants (F1 and F2) and seven tokens (/ɛ/, /ɪ/, /i/, /e/, /æ/, /ɑ/ or /u/) were used to calculate ΔF. F1 and F2 values for every production in every trial and even those for the applied shifts were then vocal-tract-length-normalised by dividing them by ΔF.

Thus, vocal tract length normalisation converted F1-F2 space (Hz) into a normalised F1-F2 space for the purposes of the analysis. As an example, the shifts for one participant are shown in Fig. 2.2 A. Fig. 2.2 B shows the Experiment 2 shifts for the same participant as in Fig. 2.2 A.

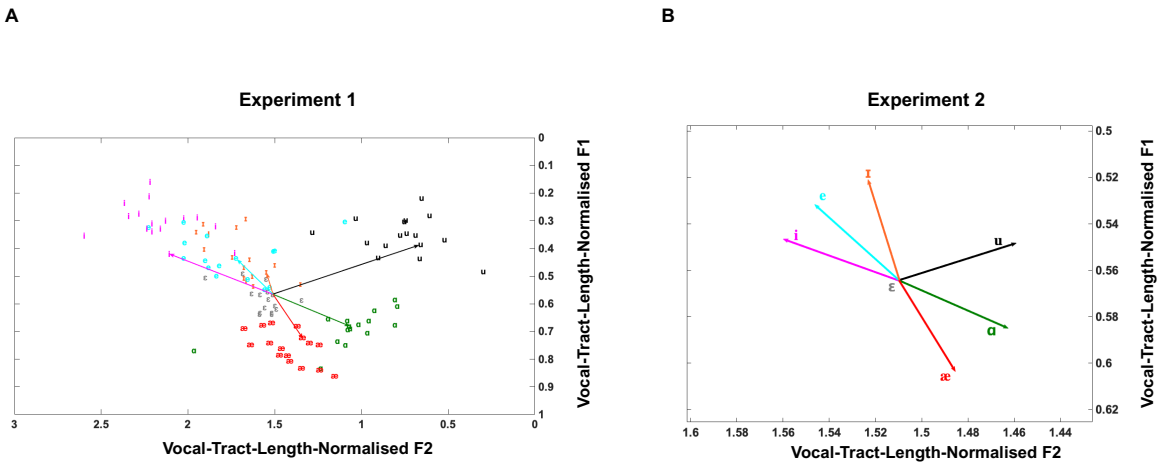


Fig. 2.2 Scatter plot of vowel locations and visual representations of formant shifts ***A.*** *A scatter plot of the vowel locations of all 16 participants who took part in Experiment 1 represented in vocal-tract-length-normalised F1-F2 space. The x-axis denotes vocal-tract-length-normalised F2 values and the y-axis represents vocal-tract-length-normalised F1 values. The vectors represent six shifts for one particular participant.* ***B.*** *A visual representation of the formant shifts for all six cases in Experiment 2 for the same example participant as in A. The x-axis denotes vocal-tract-length-normalised F2 values and the y-axis represents vocal-tract-length-normalised F1 values.*

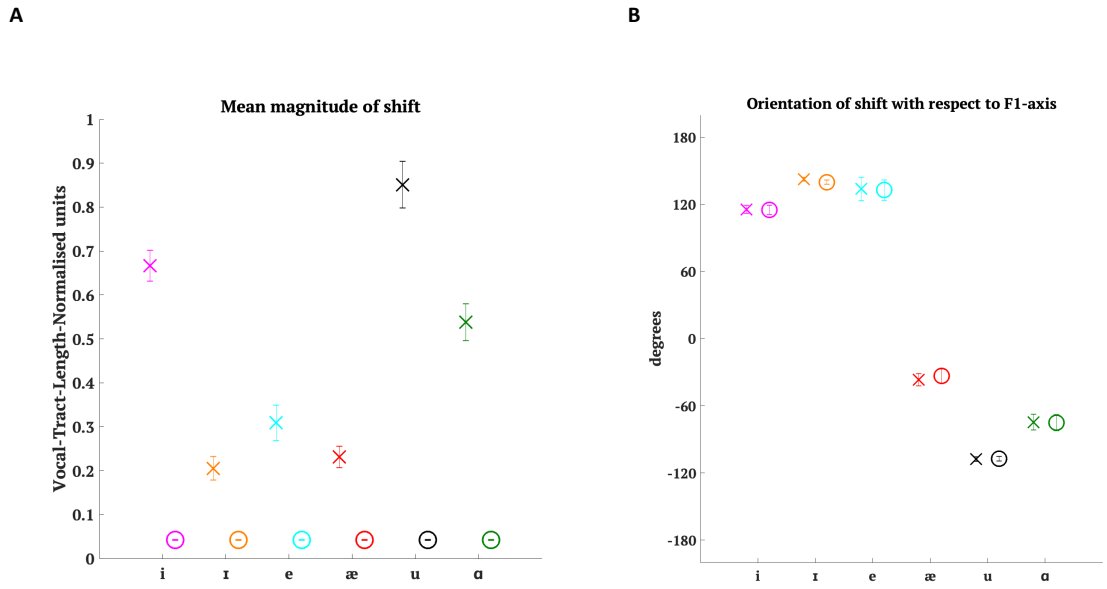


Fig. 2.3 Mean magnitudes of applied shift vectors and circular means of the angles between the applied shift vectors and the F1-axis
A. *Mean magnitude of the applied shift vector (in vocal-tract-length-normalised units) for all 6 cases in Experiment 1 (crosses) and Experiment 2 (circles) averaged across participants. Error bars depict ± 1 standard error of mean.* ***B.*** *Circular mean angle between the applied shift vector and the F1-axis (in degrees) for all 6 cases in Experiment 1 (crosses) and Experiment 2 (circles) averaged across participants. Positive values indicate that the angle is measured clockwise from a line parallel to the F1-axis to the shift vector (see Fig. 2.4A). Negative values indicate that the angle is measured anticlockwise from a line parallel to the F1-axis to the shift vector. Error bars depict ± 1 standard error of mean.*

2.3.5 Vector resolution of response vector

To look at the results of Experiment 1 and Experiment 2, both the applied shift and response to this shift were expressed as vectors in the two-dimensional vocal-tract-length-normalised F1-F2 space. For every case in both experiments, subjects' baseline vowel formant frequency values for /ɛ/ were determined by averaging the frequency values for the first 10 trials (baseline block). For every trial in every case, the Total Response (TR) was calculated as the length of the response vector in vocal-tract-length-normalised F1-F2 space from the point representing the baseline

vowel formant frequency values to the formant frequency values of the vowel production in that particular trial. The response vector was resolved into a component along the axis of applied shift and a component perpendicular to the axis of applied shift (Fig. 2.4A), henceforth called Compensatory Response (CR) and Orthogonal Response (OR) respectively. CR was thus the scalar projection of the response on the shift axis. CR values were assigned a negative sign if the response was in the direction opposite to the applied shift and a positive sign if it was in the same direction (example in Fig. 2.4B, see also Fig. 2.4A). OR values were assigned a negative sign if the response vector was located in the first or second quadrant of the cartesian coordinate system with the shift-axis as the reference and a positive sign if the response vector was located in the third or fourth quadrant (example in Fig. 2.4C, see also Fig. 2.4A). TR took the same sign as that assigned to CR. The angle between the response vector and the F1 - axis (ANG) was also measured for every trial.

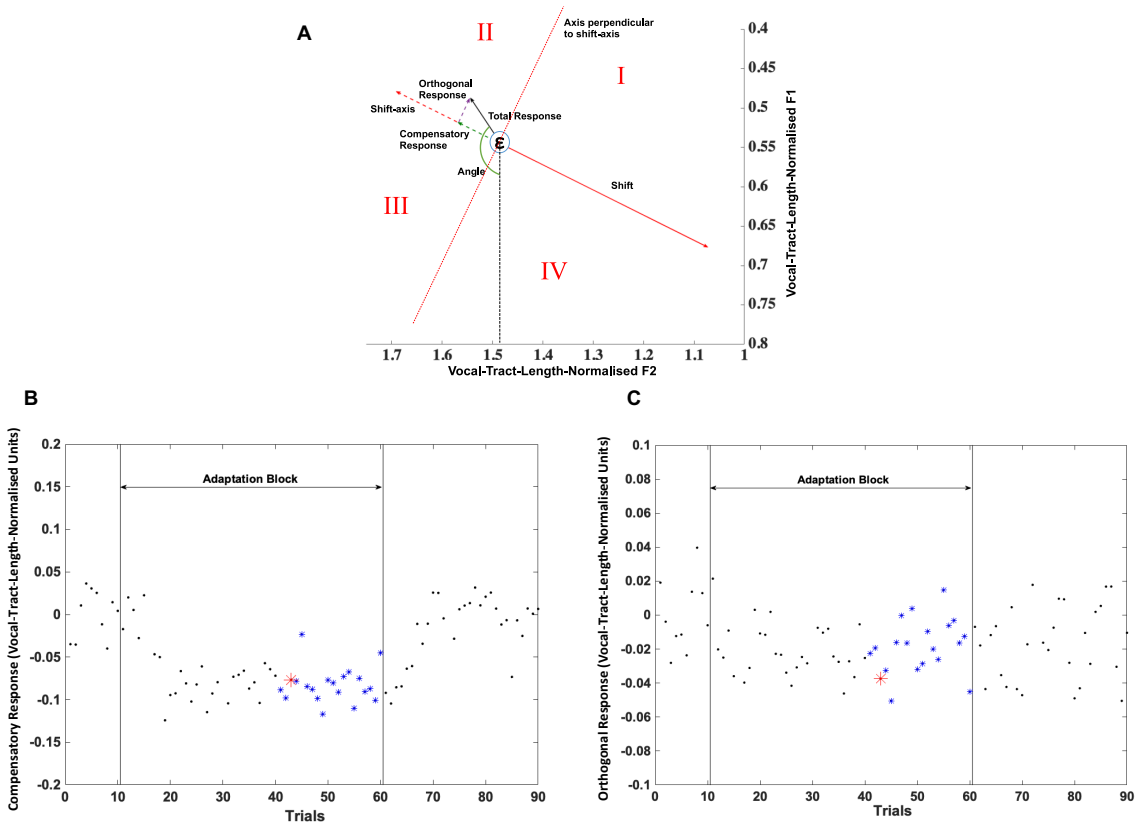


Fig. 2.4 Vector resolution

A. *In this example, the compensatory response is negative because it opposes the applied shift, the orthogonal response is negative as well because it lies in the second quadrant with respect to the direction of the shift. Quadrant numbers are mentioned in Roman numerals. Here, the total response takes the sign of the compensatory response and is therefore negative. The angle between the response vector and a line parallel to the F1-axis was measured in degrees.* ***B.*** *A scatter plot showing all 90 trials from one case for one participant on the x-axis with compensatory response (in vocal-tract-length-normalised units) represented by the y-axis. The last 20 trials of the adaptation block, which were considered in the statistical analysis, are represented by asterisks (blue; colour online). The large asterisk (red; colour online) is the example trial shown in A.* ***C.*** *A scatter plot showing all 90 trials from one case for one participant on the x-axis with orthogonal response (in vocal-tract-length-normalised units) represented by the y-axis. The last 20 trials of the adaptation block are represented by asterisks (blue; colour online). The large asterisk (red; colour online) is the example trial shown in A.*

We suggest that each of these four measures defines a particular aspect of sensorimotor adaptation. TR is a direct measure of the articulatory realisation and thus represents the size of the response. CR is a measure of the actual compensation in articulation to 'nullify' the effect of

the applied shift. OR represents the component of the response that doesn't contribute to compensation but contributes to the total size of the response. Collectively, CR and OR measure the efficiency of the response. ANG indicates the invariant direction of response, not with respect to the applied shift but in two-dimensional vocal-tract-length-normalised space and represents the acoustic and articulatory consequences of the adaptive response.

2.3.6 Statistical Analyses

In this subsection, we describe the statistical analyses we ran. We first talk about the tests we ran to test the dependence of TR, CR and OR on the independent variables in both experiments and across experiments. We then talk about how we dealt with circular quantities in our statistical analyses. We proceed to talk about the dependence of ANG on the independent variables in both experiments and across experiments. Lastly, we talk about how we controlled for false positives due to multiple significance testing.

2.3.6.1 TR, CR and OR in Experiment 1 and Experiment 2

In both experiments, the primary independent variable of interest was the direction of applied auditory feedback (angle of shift in vowel formant space). To test the hypothesis that the adaptation response depended on the direction of applied shift, we evaluated separate linear mixed effect models (LMM) for three of the dependent variables TR, CR and OR (implemented in SAS 9.4 (SAS Institute Inc., Cary, NC)). For each model and for data from both experiments, magnitude of shift and gender were included as covariates in the LMM. The linear mixed effects model provides a principled framework for examination of the effect of an independent variable of interest on a dependent variable of interest even in the presence of covariates that are

correlated with the independent variable of interest. For experiment 2, the applied shifts were uniform in magnitude in F1-F2 space across participants and cases (50 Hz). However, after vocal-tract-length-normalisation, the shift values were uniform only within-participant and not across participants. Therefore, magnitude of applied shift was also included as a covariate for experiment 2 (even though the magnitude difference across participants was minimal).

Across all the above LMMs, subjects were treated as a random effect. For each of these LMMs, we compared a fixed effects model without varying slopes to a model with varying slopes and we were unable to reject the null hypothesis of equal slopes using the Type III F-statistic for an interaction term (Littell et al., 2006). Therefore, we chose to report the models with random intercepts.

To evaluate the secondary hypotheses that adaptation responses depended on the relative backness or height of the applied shift, we ran separate and reduced LMMs for vowel backness and height. For relative vowel height, we grouped shifts towards vowels into two categories based on the relative height with respect to the produced vowel /ɛ/ (relatively higher: /i/, /ɪ/, /e/ and /u/, relatively lower: /æ/ and /ɑ/). For relative vowel backness, we grouped shifts into two groups according to relative vowel backness with respect to the produced vowel /ɛ/ (relatively to the front: /i/, /ɪ/ and /e/, relatively to the back: /æ/, /u/ and /ɑ/). For these reduced LMMs, across both experiments, the dependent variables were TR, CR and OR.

2.3.6.2 Comparison of TR, CR and OR across the two experiments

To examine differences in response measures across the two experiments, we used data from the 14 participants who took part in both experiments. To account for differences in the magnitude of applied shift across the two experiments, normalised measures were computed and used. Normalised Compensation Ratio (NCR), Normalised Orthogonal Ratio (NOR) and Normalised Total Ratio (NTR) were calculated for every subject and case by first averaging CR, TR and OR values across the last 20 trials of the adaptation block and dividing these averaged values by the magnitude of the applied shift. First, separate LMMs were computed with experiment number as an independent categorical variable, and NTR, NCR and NOR as the dependent variables. The angle of applied shift as an independent variable of interest and an interaction term (experiment number by angle of shift) were also included in these models. Second, separate and reduced LMMs were examined for the effect of relative vowel height and backness. For these reduced models, as before, shifts were grouped by either relative vowel height or backness. These reduced LMMs included experiment number and vowel height or backness as the independent variables of interest, and also included interaction terms. For these LMMs, subjects were treated as a random effect in all models and the models included random slopes and intercepts (Littell et al., 2006).

2.3.6.3 Dealing with circular quantities

The primary independent variable of interest (angle of shift) and one dependent variable (ANG) were circular quantities or values that were measured along a circle. Due to the periodic nature of such quantities, they may require statistical analyses designed for circular data (Mardia and Jupp, 2009; Cremers and Klugkist, 2018).

2.3.6.3.a Circular independent variable of interest (angle of shift)

To account for the independent variable, angle of shift, being a circular quantity, we used target vowel direction (/ɪ/, /i/, /e/, /æ/, /ɑ/ or /u/), a categorical variable, as a surrogate for angle of shift to represent the direction of applied shift. We reran all the LMMs involving angle of shift as the primary variable of interest for TR, CR and OR in both experiments as well as the normalised measures comparing values from Experiment and Experiment 2. We did not find evidence suggesting that replacing the circular quantity by a surrogate variable changes the findings of our study. Therefore, we chose to report the results from the LMMs with angle of shift as an independent variable in the subsequent sections because angle of shift captures the between-subject variability in the direction of applied shift (because angle of shift is a continuous variable).

2.3.6.3.b Circular dependent variable (ANG)

To account for the dependent variable, ANG, being a circular quantity, we first verified that our data followed the von Mises distribution (von Mises, 1918) using Watson's U^2 goodness-of-fit test (Watson, 1961; Lockhart and Stephens, 1985) (Experiment 1: U^2-statistic = 2.8248, $p < 0.01$, R-squared = 0.2255; Experiment 2: U^2-statistic = 1.9357, $p < 0.01$, R-squared = 0.2243; Experiment 1 vs Experiment 2: U^2-statistic = 4.3611, $p < 0.01$, R-squared = 0.2208). For each experiment, we then ran three separate parametric Watson-Williams (WW) tests as one-way ANOVA tests for circular data (Watson and Williams, 1956; Berens, 2009; Cremers and Klugkist, 2018) with ANG as the dependent variable and target vowel direction, target vowel relative height and target vowel relative backness as the categorical independent variables in

those three tests respectively. For the comparison of response measures across both experiments, we ran three separate parametric Harrison-Kanji (HK) tests as a two-way ANOVA for circular data (Harrison and Kanji, 1988; Berens, 2009) with ANG as the dependent variable. The two independent categorical variables for the three tests were, respectively: experiment number and target vowel direction, experiment number and target vowel height, experiment number and target vowel backness. An interaction term between the independent variables was also included for each HK test.

2.3.6.4 Multiple correction

Finally, to control for false positives due to multiple significance testing, p-values for significance were adjusted using the Benjamini-Hochberg False Discovery Rate ($\alpha = 0.05$) Procedure (Benjamini and Hochberg, 1995) and only findings that survived this adjusted significance threshold were reported.

2.4. Results

2.4.1 Experiment 1

Results showed (Fig. 2.5A, Supp Fig. A.2) that the compensatory response (CR) was in the direction opposite to the applied shift in all cases except the shift towards /u/. Although most CR values indicated that participants tried to oppose (or compensate for) the shift whether the angle of shift was positive or negative, the value of CR varied along with the angle of shift [$F(1,1828) = 39.97$, $p < 0.0001$]. If we look at CR values for shifts towards /i/ and /ɪ/, although the angles of shift were close to each other and on the same side with respect to the F1-axis (Fig. 2.3B), their CR values were very different from each other. CR also varied along with the magnitude of shift

[F(1,1828) = 119.51, p < 0.0001]. The largest shifts, on average, were towards /u/ (Fig. 2.3A) and produced a following response. The shifts towards /ɪ/ and /æ/ were smaller than the rest but still produced differing CR values.

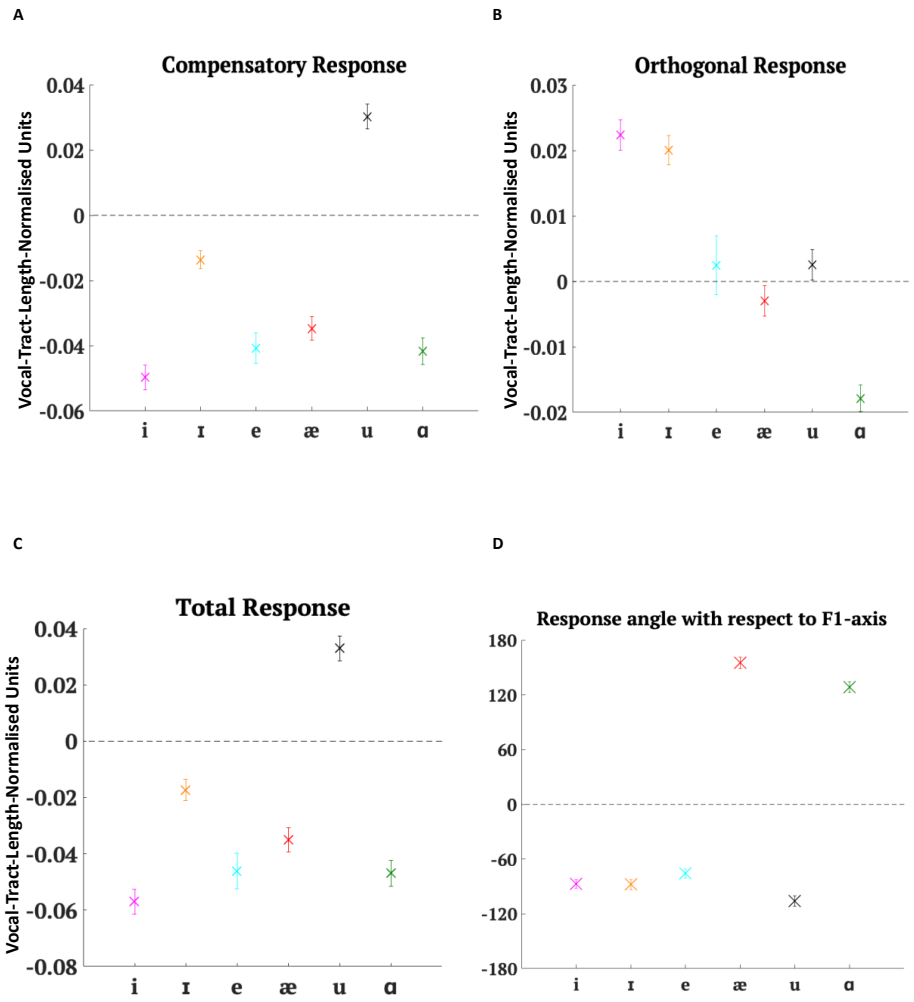


Fig. 2.5 Experiment 1 results

Mean results for all 6 cases in Experiment 1 averaged across participants. Error bars depict ± 1 standard error of mean. ***A.*** *Mean compensatory response (in vocal-tract-length-normalised units). Negative values indicate a response opposing the applied shift whereas positive values indicate a response following the applied shift.* ***B.*** *Mean orthogonal response magnitude (in vocal-tract-length-normalised units). Negative values indicate a response in the first and second quadrant with the direction of the shift vector as the reference. Positive values indicate a response in the third or fourth quadrant.* ***C.*** *Mean total response (in vocal-tract-length-normalised units). Negative values indicate a response opposing the applied shift whereas*

positive values indicate a response following the applied shift. ***D.*** *Circular mean angle between the response vector and the F1-axis (in degrees). Negative values indicate that the angle is measured clockwise from the F1-axis to the response vector. Positive values indicate that the angle is measured anticlockwise from the F1-axis to the response vector.*

If we examine the orthogonal response (OR) (Fig. 2.5B), we observe that shifts towards /e/, /æ/ and /u/ produced a small OR but shifts towards /i/, /ɪ/ and /ɑ/ engendered a comparatively larger OR. On average, with the notable exception of /u/, the pattern seems to match that of angle of shift, with positive angles of shift producing a positive OR and negative angles of shift producing a negative OR. Indeed, OR was dependent on angle of shift [$F(1,1828) = 87.97$, $p < 0.0001$]. Looking at the figure, there seems to be a natural grouping based on whether the shifts were towards vowels that are relatively higher (a decrease in F1 values) than /ɛ/ or whether the shifts were towards vowels that are relatively lower (an increase in F1 values). OR, however, is not dependent on magnitude of shift.

In Fig. 2.5C, we see that total response (TR) shows a pattern that is very similar to CR. It was dependent on angle of shift [$F(1,1828) = 44.27$, $p < 0.0001$] and magnitude of shift [$F(1,1828) = 93.48$, $p < 0.0001$]. As expected, responses to shifts towards /u/ stood out as an exception because of their following nature.

When we look at ANG values in Fig. 2.5D, we again observe clear groupings in responses to shifts towards vowels that are relatively higher than /ɛ/ versus shifts towards vowels that are relatively lower than /ɛ/. Because shifts towards /i/, /ɪ/ and /e/ caused responses that were compensatory in nature and their angles of shift with respect to F1- axis were positive, it makes sense to see that their ANG values were negative. Similarly, since shifts towards /ɑ/ and /æ/ also

produced compensatory responses and their angles of shift were negative, their ANG values were positive. It is then obvious that /u/ with its following response would have an average ANG value that is negative. Thus, ANG, the angle made by the response vector with the F1-axis or the orientation of the response in formant frequency space regardless of the applied shift, depended on the target vowel direction [$F(5,1845) = 154.06$, $p < 0.0001$].

OR was dependent on whether the shift was towards higher vowels or lower vowels [$F(1,1829) = 91.28$, $p < 0.0001$] and so was ANG [$F(1,1845) = 737.24$, $p < 0.0001$]. Even CR [$F(1,1829) = 34.47$, $p < 0.0001$] and TR [$F(1,1829) = 22.27$, $p < 0.0001$] showed dependency on relative vowel height. CR and TR values for shifts towards /æ/ and /ɑ/, on average, did not differ much from those for shifts towards /i/, /ɪ/ and /e/. However, the differences found could be explained by the exceptional following behaviour in the case of shifts towards /u/.

Also, CR [$F(1,1829) = 37$, $p < 0.0001$], TR [$F(1,1829) = 38.65$, $p < 0.0001$], OR [$F(1,1829) = 91.49$, $p < 0.0001$] and ANG [$F(1,1845) = 474.83$, $p < 0.0001$] were all dependent on whether shifts were towards vowels to the front of /ɛ/ versus to the back of /ɛ/. Responses to shifts towards /u/ may be responsible for the effect in CR and TR, whereas in OR it could be the relatively large negative value for shifts towards /ɑ/.

To summarise, in Experiment 1, all response measures, Compensatory Response (CR), Orthogonal Response (OR), Total Response (TR) and the angle between the response vector and F1-axis (ANG), depend on the angle of applied shift. Additionally, CR and TR also depend on

the magnitude of shift. All four measures were also dependent on the relative vowel height and vowel backness of the target vowel of the altered auditory feedback.

2.4.2 Experiment 2

Results from Experiment 2 showed that the smaller shifts produced two clear groups of CR values (Fig. 2.6A, Supp Fig. A.3). For shifts in the directions of /i/, /ɪ/ and /e/, the responses were, on average, compensatory in nature whereas responses to shifts towards /æ/, /u/ and /ɑ/ were, on average, following in nature. There was a clear subdivision based on shifts towards vowels that are relatively to the front of /ɛ/ (increase in F2) versus shifts towards vowels that are relatively to the back of /ɛ/ (decrease in F2). For angles of shift that were positive, participants seemed to compensate whereas they seemed to follow when angles of shift were negative. Thus, CR depended on angle of shift [$F(1,1823) = 75.55$, $p < 0.0001$]. CR also depended on magnitude of shift [$F(1,1823) = 7.05$, $p = 0.008$].

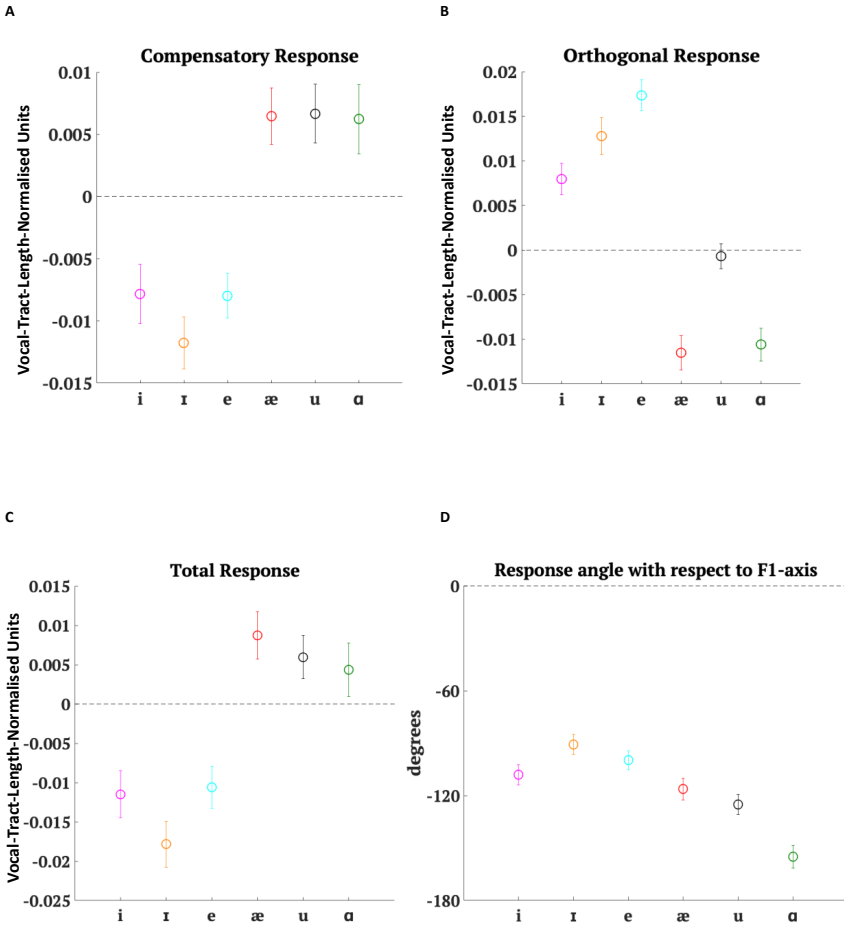


Fig. 2.6 Experiment 2 results

Mean results for all 6 cases in Experiment 2 averaged across participants. Error bars depict ± 1 standard error of mean. ***A.*** *Mean compensatory response (in vocal-tract-length-normalised units). Note: The y-axis scale differs from the y-axis scale in Fig. 2.5A.* ***B.*** *Mean orthogonal response magnitude (in vocal-tract-length-normalised units). Note: The y-axis scale differs from the y-axis scale in Fig. 2.5B.* ***C.*** *Mean total response (in vocal-tract-length-normalised units). Note: The y-axis scale differs from the y-axis scale in Fig. 2.5C.* ***D.*** *Circular mean angle between the response vector and the F1-axis (in degrees). Note: The y-axis scale differs from the y-axis scale in Fig. 2.5D.*

In Fig. 2.6B, we can observe that although the average OR value for shifts towards /u/ was small just like in Experiment 1, the average value in Experiment 2 was negative. While OR values for shifts towards /u/ were an exception in Experiment 1, OR values in Experiment 2 were positive for angles of shift that were positive with respect to F1-axis (i.e. shifts with an increase in F2

values) and negative for angles of shift that were negative (shifts with a decrease in F2 values). There was a clear divide along shifts towards vowels that are relatively to the front of /ɛ/ and shifts towards vowels that are relatively to the back of /ɛ/. Thus, OR was dependent on angle of shift [$F(1,1823) = 101.75$, $p < 0.0001$]. OR was not dependent on magnitude of shift.

TR values (Fig. 2.6C) showed a pattern similar to CR. TR was dependent on angle of shift [$F(1,1823) = 76.98$, $p < 0.0001$] with positive angles of shift causing a net compensatory response and negative angles of shift causing a net following response. TR was dependent on magnitude of shift [$F(1,1823) = 4.13$, $p = 0.0423$].

In Fig. 2.6D, ANG values showed an interesting pattern in Experiment 2. All shifts produced ANG values that were negative on average. This makes sense when you note that positive angles of shift caused a compensatory response and thus should have negative ANG values. Similarly, since negative angles of shift caused a following response, one would expect to see negative ANG values for these shifts. ANG values varied according to the target vowel direction [$F(5,1840) = 12.3$, $p < 0.0001$] .

In Experiment 2, although the main divide was not along the lines of vowel height, TR [$F(1,1824) = 32.14$, $p < 0.0001$], CR [$F(1,1824) = 31.44$, $p < 0.0001$], OR [$F(1,1824) = 171.79$, $p < 0.0001$] and ANG [$F(1,1840) = 33.53$, $p < 0.0001$] all depended on whether the shifts were towards higher vowels or towards lower vowels as compared to /ɛ/. The larger effect size in the case of OR as compared to CR and TR was perhaps a result of the average OR value of shifts towards /u/ being closer to the average OR values of the other relatively higher vowels.

The response measures that showed a clear dichotomy between shifts towards vowels that are relatively to the front of /ɛ/ and shifts towards vowels that are relatively to the back of /ɛ/ (Fig. 2.6A, 2.6B, 2.6C) were CR [$F(1,1824) = 71$, $p < 0.0001$] , OR [$F(1,1824) = 196.28$, $p < 0.0001$] and TR[$F(1,1824) = 67.01$,$p < 0.0001$]. Even ANG [$F(1,1840) = 37.72$, $p < 0.0001$] depended on relative vowel backness.

To summarise, in Experiment 2 as well, all response measures, Compensatory Response (CR), Orthogonal Response (OR), Total Response (TR) and the angle between the response vector and F1-axis (ANG), depended on angle of shift in Experiment 2 just as in Experiment 1. Additionally, CR and TR also depended on variations in the magnitude of applied shift. All four measures also depended on whether the applied shift involved an increase or decrease in F1 value and whether they involved an increase or decrease in F2 value.

2.4.3 Accounting for perceptual differences in magnitude of applied shift

The applied shifts in our experiments were in Hertz values. In Experiment 2, the shift was 50Hz for all cases. Vocal tract length normalisation of the vowel formant space maintained shifts of equal magnitudes across directions of shift for a given subject but the magnitudes varied slightly across subjects, as mentioned before in section 2.3.6.1.

However, shifts designed in Hertz values, although uniform in F1-F2 vowel space, may not be uniform on a psychoacoustic or perceptual scale. Each subject's individual perceptual scale for formant frequency values may be too idiosyncratic to be captured on a fixed scale like the mel

scale (Greenwood, 1997). Nevertheless, to account for the possibility of perceptual differences for different shifts in both experiments, we converted all shift magnitudes to mels (Supp Fig. A.4) (O'Shaughnessy, 1987) and reran the statistical models including magnitude of shift in mels as a covariate instead of magnitude of shift in vocal-tract-length-normalised units. We did not observe any differences in the fixed effect of angle of shift in either experiment. It is important to note that this conversion to mels would cause shift magnitudes to vary both within and across subjects.

2.4.4 Comparison of response measures across experiments

In Fig. 2.7A, it can be seen that for all shifts, the magnitude of Normalised Compensatory Response (NCR), i.e. the compensatory response divided by the magnitude of applied shift, was larger in Experiment 2 than in Experiment 1. In proportion to the applied shift magnitude, participants had larger CR values in Experiment 2 as compared to Experiment 1. Therefore, NCR was dependent on experiment number [$F(1,3225) = 30.25$, $p < 0.0001$], angle of shift [$F(1,3225) = 119.9$, $p < 0.0001$] and there was a significant interaction between experiment number and angle of shift [$F(1,3225) = 28.89$, $p < 0.0001$]. Since NCR values were normalised CR values and since CR was dependent on angle of shift in both experiments, it comes as no surprise that NCR would depend on angle of shift. Moreover, the significant interaction term indicates that the pattern of covariation of NCR with angle of shift depended on experiment number. NCR depended on target vowel backness [$F(1,3225) = 106.01$, $p < 0.0001$]. This dichotomy between shifts towards vowels to the front of /ɛ/ and vowels to the back of /ɛ/ is clearer in Experiment 2 than in Experiment 1 seen in the interaction term between experiment number and target vowel backness [$F(1,3225) = 34.59$, $p < 0.0001$]. Similarly, NCR also depended on target vowel height

[F(1,3225) = 35.29, p < 0.0001] and its interaction with experiment number [F(1,3225) = 33.28, p < 0.0001].

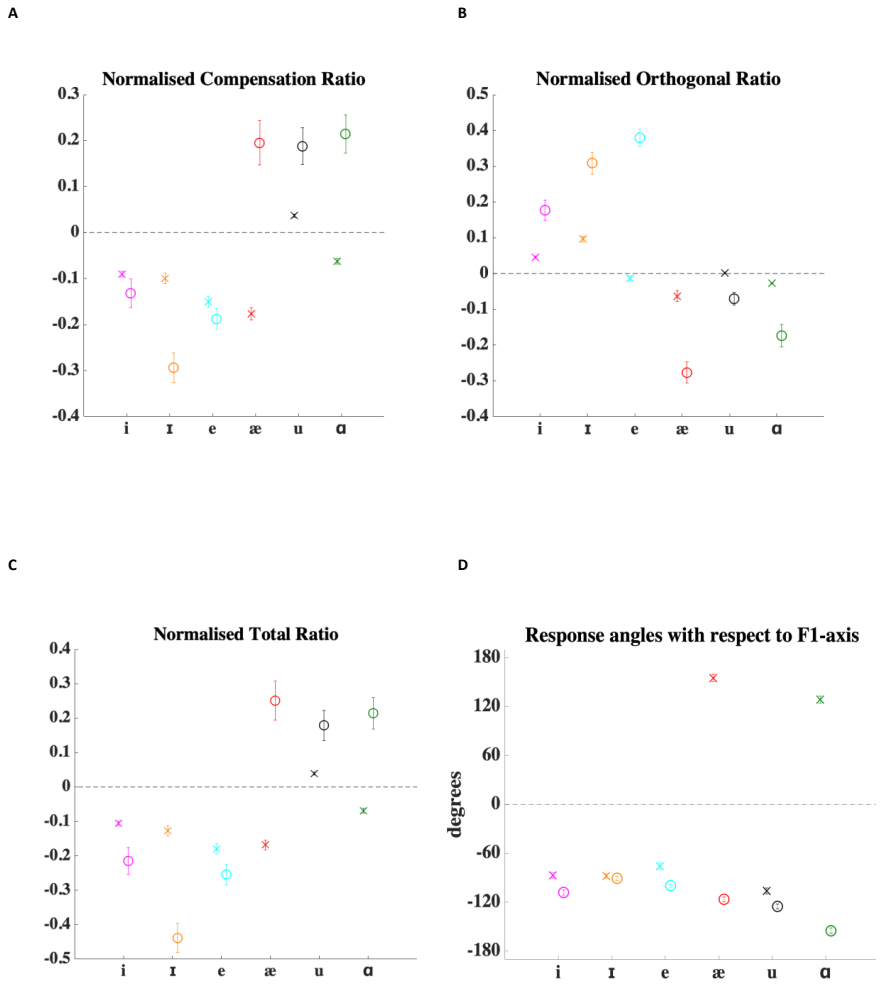


Fig. 2.7 Comparison of responses across experiments
Mean results for 14 of the 18 participants who took part in both Experiment 1 and Experiment 2. Error bars depict ± 1 standard error of mean. ***A.*** *Average compensatory responses for both experiments were normalised (i.e. divided by the applied shift magnitude) and were called Normalised Compensation Ratios. The Normalised Compensation Ratio values for Experiment 1 (crosses) and Experiment 2 (circles) are shown.* ***B.*** *Average orthogonal responses for both experiments were normalised (i.e. divided by the applied shift magnitude) and were called Normalised Orthogonal Ratios. The Normalised Orthogonal Ratio values for Experiment 1 (crosses) and Experiment 2 (circles) are shown.* ***C.*** *Average total responses for both experiments were normalised (i.e. divided by the applied shift magnitude) and were called Normalised Total Ratios. The Normalised Total Ratio values for Experiment 1 (crosses) and Experiment 2 (circles) are shown.* ***D.*** *Circular mean angle between the response vector and the F1-axis (in degrees) for Experiment 1 (crosses) and Experiment 2 (circles) are shown.*

Fig. 2.7B tells us that Normalised Orthogonal Response (NOR) values too were larger in Experiment 2 than in Experiment 1. Smaller shifts in Experiment 2 produced proportionally larger OR values. However, NOR was not found to be dependent on experiment number [$F(1,3225) = 1.9$, $p = 0.1678$]. Looking at the pattern of NOR across cases, it seems that the mean NOR values for each experiment were not that different from each other. NOR was dependent on angle of shift [$F(1,3225) = 102.39$, $p < 0.0001$] and there was a significant interaction between experiment number and angle of shift [$F(1,3225) = 42.21$, $p < 0.0001$]. This reiterates that the angle of shift determined OR values and how OR covaried with angle of shift depended on experiment number. There was also a significant target vowel backness effect for NOR [$F(1,3225) = 188.07$, $p < 0.0001$] and this effect had an interaction with experiment number [$F(1,3225) = 100.41$, $p < 0.0001$]. The shifts towards front vs shifts towards back divide was clearly much more evident in Experiment 2. NOR depended on the target vowel height [$F(1,3225) = 145.79$, $p < 0.0001$] and its interaction with experiment number [$F(1,3225) = 69.34$, $p < 0.0001$].

Normalised Total Response (NTR) (Fig. 2.7C) showed a pattern that was similar to NCR. TR was proportionally larger for the smaller shifts in Experiment 2 as compared to Experiment 1. Although NCR was dependent on experiment number and NOR was not, NTR was indeed dependent on experiment number [$F(1,3225) = 14.51$, $p = 0.0001$]. NTR also varied according to angle of shift [$F(1,3225) = 125.27$, $p < 0.0001$] and there was a significant interaction between experiment number and angle of shift [$F(1,3225) = 36.35$, $p < 0.0001$]. Similar to the effects seen in NCR and NOR, NTR was also dependent on the target vowel backness [$F(1,3225) = 106.58$, p

< 0.0001] and its interaction with experiment number [$F(1,3225) = 34.65$, $p < 0.0001$]. Shifts towards vowels to the front of /ɛ/ produced a consistent oppositional response in Experiment 2 whereas shifts toward vowels to the back of /ɛ/ consistently produced following responses. NTR also depended on the target vowel height [$F(1,3225) = 42.18$, $p < 0.0001$] and its interaction with experiment number [$F(1,3225) = 32.45$, $p < 0.0001$].

The angle between the response vector and F1-axis (ANG) values could not be normalised by magnitude of shift values but on comparison, the pattern in Experiment 1 was different from that in Experiment 2, as observed before. However, statistically from the HK test, we did not find that ANG was dependent on experiment number [$\chi2 = 2.7231$, $p = 0.2563$]. When ANG values from both experiments were taken into consideration, they were still dependent on target vowel direction [$\chi2 = 243.1855$, $p < 0.0001$]. The interaction term between experiment number and target vowel direction was also significant [$\chi2 = 149.5644$, $p < 0.0001$], indicating that the way ANG values covaried with the angle of shift depended on the experiment number.

Like the other measures of response, ANG was also dependent on target vowel backness [$\chi2 = 183.5941$, $p < 0.0001$] and target vowel height [$\chi2 = 197.9663$, $p < 0.0001$] and their interactions with experiment number respectively [$\chi2 = 65.8612$, $p < 0.0001$ and $\chi2 = 137.1501$, $p < 0.0001$].

Further details about the responses in both experiments can be seen upon examination of the two-dimensional plot of produced formant changes (solid lines with standard error ellipses) in response to applied shift vectors (dashed lines) (Fig. 2.8A for Experiment 1 and Fig. 2.8B for Experiment 2). The representation of both the applied shifts and the corresponding responses

shown here is based on vector averages across both participants and trials (Note however, actual applied shift vectors were participant-specific and adaptation responses were calculated for each subject relative to these participant-specific shift vectors). It can indeed be seen from Fig. 2.8A that in Experiment 1, responses to shifts towards /i/, /ɪ/, /e/, /æ/ and /ɑ/ were compensatory in nature whereas responses to shifts towards /u/ were following in nature. In Fig. 2.8B, we see that in Experiment 2, responses to shifts towards /i/, /ɪ/ and /e/ remain compensatory in nature, whereas responses to shifts towards /æ/, /u/ and /ɑ/ were following in nature. The figure highlights the qualitative similarities and differences between the two experiments. In both experiments, the responses to shifts towards /u/ are following and shifts towards /i/ and /ɪ/ show a compensatory response. The only major qualitative differences between responses in Experiment 1 and 2 are for the shifts towards vowels that are lower and to the back of /ɛ/, namely /æ/ and /ɑ/, which change from compensatory in Experiment 1 to following in Experiment 2.

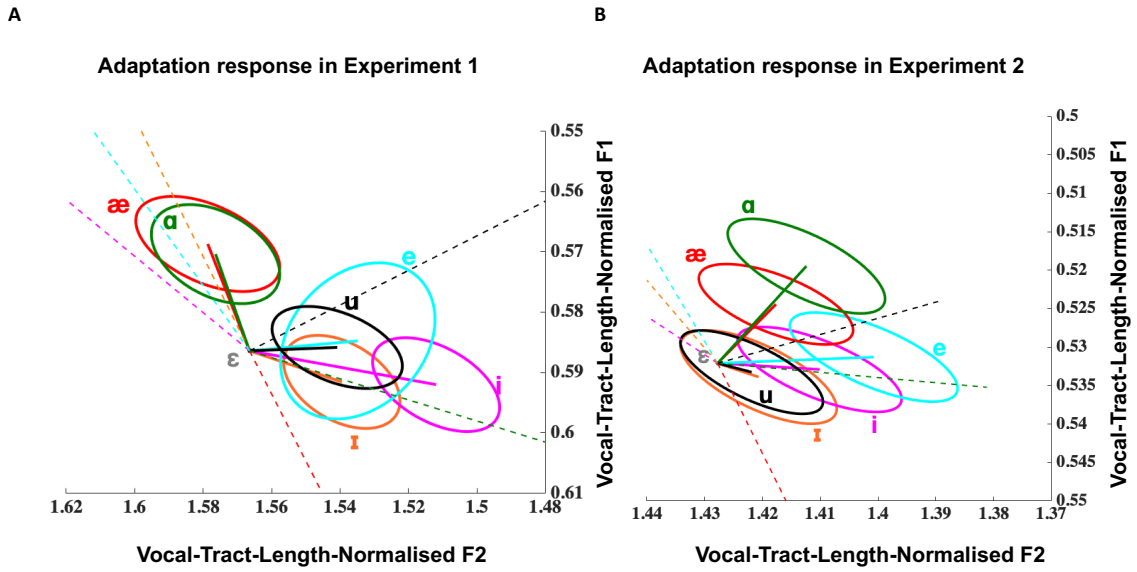


Fig. 2.8 Two-dimensional comparison of responses across experiments
A & B. *Mean baseline vocal-tract-length-normalised ɛ formant frequency values for 14 of the 18 participants who took part in both Experiment 1 and Experiment 2 and standard error ellipses representing adaptation responses to the shifts applied (shifts represented by dashed lines). Note: The representation of both the applied shifts and the corresponding responses shown here is based on vector averages across both participants and trials. However, actual applied shift vectors were participant-specific and adaptation responses were calculated for each subject relative to these participant-specific shift vectors. For each experiment, the average shift and response vectors were computed relative to the corresponding baseline vocal-tract-length-normalised ɛ formant frequency values for that experiment.*

2.5 Discussion

In this study, we conducted two experiments to investigate the possibility that sensorimotor formant adaptation varies as a function of the direction of formant shift. Both experiments looked at responses to altered formant feedback in different directions from the same vowel. In Experiment 1, the shifts were from /ɛ/ to six other vowels. In Experiment 2, the shifts were in the same six directions as in Experiment 1 but of equal magnitudes in Hertz values. For both experiments, we quantified shift magnitudes and responses in F1-F2 space that was normalised for vocal tract length (Johnson, 2018). We comprehensively characterised the vector describing

the participants' formant adaptation responses and found that formant adaptation indeed depends on the direction of the applied shift. Across both experiments, response characteristics also depended on relative height and backness of the vowel target to which feedback was altered. These results and their implications for models of speech motor control are discussed below.

2.5.1 Formant adaptation depends on direction and magnitude

Results from Experiment 1 showed that responses were dependent on the direction of the applied shift independent of differences in applied shift magnitude across the six different vowel targets. For five of these six shifts, the response was compensatory in nature with /u/ being the only exception. All response measures depended on the direction of applied shift in vocal-tract-length-normalised F1-F2 space. To further examine the specificity of the effect of shift direction on adaptation responses, in Experiment 2, for each subject, the applied shifts were in the same six directions as in Experiment 1 but of equal magnitudes in formant frequency space. Results again showed that adaptation responses depended on the direction of shift. For both experiments, the main effect of angle of shift on Total Response (TR), Compensatory Response (CR) and Orthogonal Response (OR) held true even after controlling for differences in magnitudes of applied shift in perceptual units.

There were differences in adaptation response observed across the two experiments. Responses for smaller shifts of Experiment 2 were not merely a scaled version of responses to larger shifts in Experiment 1. In particular, responses to shifts towards /æ/ and /ɑ/, which were compensatory in Experiment 1, were following in Experiment 2.

Together, these findings suggest that the simple model of compensatory responses opposing the applied feedback shift is not accurate. In fact, the pattern of compensatory responses is quite a complex function of direction and magnitude of the applied shift.

2.5.2 Possible explanations for direction dependence of adaptation

Sensorimotor adaptation in speech is a response to persistent feedback alteration. It can be described as learning to change articulation to compensate for a perceived error between predicted and actual sensory feedback, i.e. a sensory feedback prediction error. Sensorimotor adaptation in the production of vowels has previously been shown to depend on the size of the formant alteration (Katseff et al., 2012) and the vowel produced when experiencing the alteration in single-word production in laboratory settings (Mitsuya et al., 2015) or natural connected speech (Lametti et al., 2018). Mitsuya et al. showed that adaptation to F1 shifts depended on whether the F1 shift was an increase or a decrease, providing preliminary evidence that adaptation may vary as a function of shift direction. Results from the current study suggest that adaptation also depends on the direction of alteration in vowel formant space. Models of speech production currently do not account for these results. To begin to explore how these models could be modified, we consider various contributory factors that could potentially explain the direction dependence of adaptation observed in this study.

One possible explanation for why adaptation differs as a function of direction in vowel formant space involves considering the combination of two factors, acoustic to articulatory non-linearities inherent in speech production (Stevens, 1989) and the hypothesis that sensorimotor adaptation is a balance between compensating for sensory prediction errors in audition and somaesthesis

(Katseff et al., 2012; Lametti et al., 2012). The amount of articulatory change needed to counter formant feedback perturbation varies as a function of the direction of that perturbation in vowel formant space. This variation in articulatory response arises from the non-uniformity of the relationship between articulation and acoustic consequences. In particular, the quantal theory of speech production (Stevens, 1989) states that changes of articulatory configuration during speaking cause changes to the resulting acoustic output in a non-monotonic manner, i.e. in some parts of the vocal tract, a small difference in articulatory positioning corresponds to a large acoustic difference whereas in other parts even a large difference in articulatory positioning is not sufficient to cause a large change in the acoustic consequences. Because of this, different degrees of articulatory change are needed to counter acoustic perturbations in different directions. These differing articulatory changes would cause speakers to experience varying levels of somatosensory feedback change. In particular, compensatory responses to shifts towards higher vowels may differ from compensatory responses to shifts towards lower vowels because of changes in expected somatosensory feedback through the lateral contacts of the tongue at the palate.

The articulatory change that corrects for an auditory feedback prediction error will in turn generate a somatosensory feedback prediction error in the opposite direction. The resulting trade-off between correcting for auditory feedback prediction errors and competing somatosensory feedback prediction errors may partly account for the observed direction dependence of adaptation.

A second factor contributing to direction dependence could arise from articulatory constraints on producibility of compensatory responses, i.e. the ability to reconfigure the human vocal tract to a state that achieves compensation. There are physical limits to the dimensions of the vocal tract and compensatory responses may not be possible for all directions of applied shifts. Moreover, the required compensatory responses may involve moving the vocal tract into regions where the speaker has limited phonetic experience based on their language.

A related interesting finding from the current study is that responses to formant shifts also depended on the height and backness of the vowel target to which feedback was altered, relative to the intended vowel production. Changes to vowel height and backness are achieved by moving different parts of the tongue. Changes towards back vowels may require changes in the tongue body whereas changes towards front vowels may require changes to the tongue blade (Stevens, 1998). Therefore, it is natural that the compensatory adjustments would group according to whether the compensatory change is one of fronting or backing. Similarly, because different muscles are involved in tongue raising or tongue lowering (Gick et al., 2013), compensatory adjustments would also group according to whether the compensatory change is one of lowering or raising.

The observed effects of vowel height and backness are consistent with the two factors describing tongue position as described by Harshman et al. (Harshman et al., 1977) that can be explained by tongue biomechanics (Perrier et al., 2007). Furthermore, as stated before, compensatory responses to shifts towards higher vowels may differ from compensatory responses to shifts towards lower vowels because of corresponding changes in expected somatosensory feedback

through the lateral contacts of the tongue at the palate. Greater palatal contact would be more consistent with auditory feedback conveying the percept of higher vowels and so the absence of this palatal contact may have an effect on the adaptation response to shifts towards higher vowels.

A possible third factor contributing to the observed direction dependence of adaptation is categorical-like perception of vowels (Kuhl, 1991). We know that behavioural and cortical responses to formant feedback alterations depend on whether the applied shifts cross a vowel category boundary (Niziolek and Guenther, 2013). In the current study, the shifts for Experiment 1 were carefully chosen for each participant such that the resulting feedback crossed vowel category boundaries and was a perceivably-different vowel sound. Shifts in different directions would cross different numbers of category boundaries, thus affecting the perceptual salience of the applied shifts. This would result in directional dependence of adaptation responses. However, in Experiment 2, although we did not measure categorical boundaries in our study, it is likely that all the shifts stayed within the categorical boundary of /ɛ/ for three reasons. First, Niziolek and Guenther (Niziolek and Guenther, 2013) shifted formant feedback from /ɛ/ such that it did or did not cross the categorical boundary, which they measured perceptually, to either /æ/ or /ɪ/. The average shift applied in their study for crossing the categorical boundaries was 122.2 mels. Second, we observe from baseline formant measurements of the production of different vowels in our study that formant distances in F1-F2 space from /ɛ/ to the two closest vowels, /æ/ and /ɪ/, were ~165 mels. Even if we assume that the category boundary is at the midpoint between these vowels, shifts of ~83 mels would be required to cross categorical boundaries. Third, the natural variability of the productions of the vowel /ɛ/ across participants was found to be ~30 mels (two

times the standard deviation of the mean baseline /ɛ/ production = 29.91 mels) which is likely to be well within the vowel category of /ɛ/. Nevertheless, in spite of all shifts being within the same category, we observed direction dependence of the adaptation response. Therefore, taken together, the results of Experiments 1 and 2 suggest that direction of applied shift is by itself a determining factor for the adaptation response. There is some evidence that sensorimotor adaptation is sensitive to the ease with which the altered vowel feedback can be assimilated into a known phonological category (Mitsuya et al., 2013). Adaptation is also sensitive to the lexical status of the altered vowel feedback sound (Bourguignon et al., 2014). While we did not focus on lexical and phonological categories in our experiments, it would be an important research path to pursue in the future.

The manipulations in Experiment 1 were focused on distinctiveness of the shifted auditory feedback to a perceivably-different vowel sound and how one responds to these feedback alterations. Although the prompt word for both experiments was the same (‘bep’), the formant frequency values of the shifts were determined prior to the experiments in a separate pre-test session where participants produced words containing different vowels. These words had different consonantal environments. Rhotacisation, diphthongisation and other coarticulatory contexts encountered in the production of these words would affect vowel formant frequencies and therefore the formant shifts used in our experiments. However, since the consonantal environment was fixed in the actual experiments we do not think these factors affect the adaptation responses which were always measured relative to the applied shift. We also note that the durational characteristics of a particular vowel may covary with its formant frequency values. The feedback may be perceived as unnatural because of differences in vowel length between the

natural production of a particular vowel and its altered-feedback version. It is true that our feedback shifts did not impact the duration of the perceived vowel. Five of the six feedback shifts in Experiment 1 would be considered shifts from a short vowel to a long vowel (with the only exception being /ɪ/, where the shift would be from a short vowel to another short vowel (Rositzke, 1939). Nevertheless, this may not be an important factor affecting our findings for the following reasons. First, vowel duration has small or relatively modest effects on vowel identity (Hillenbrand et al., 2000). Furthermore, in American English (the language of our study participants) the distinction between short and long vowels is not so sharp as in other dialects of English like British English (Wells, 1962). Second, vowel durations are sensitive to the consonant context (House and Fairbanks, 1953). In particular, a well-known phenomenon called pre-fortis clipping applies to our experiments because the coda consonant in our prompt was a fortis obstruent (/p/). It has been shown that this consonant context can reduce long vowel durations by almost 40-50% (House, 1961; Kluender et al., 1988; Wells, 1990). Therefore, having the vowel formants shifted to long vowel identities while retaining the short vowel duration of the produced vowel /ɛ/ will not sound unnatural due to perceptual expectations of pre-fortis clipping of long vowel durations before the coda consonant /p/.

2.5.3 The curious case of following responses

There were a number of following responses observed in our experiments. Here we discuss the possible reasons for such responses.

Participants, on average, tended to follow the shift from /ɛ/ towards /u/ in Experiment 1. There could be two reasons why this may have occurred: 1) /u/ was the only rounded vowel amongst

the six shifts and may provide rich somatosensory feedback during rounding. An altered auditory feedback sounding like /u/ without the presence of lip-rounding somaesthesis may drive the participants towards 'rounding' their production by producing something like /u/. We only tracked the first two formants in the current study but this line of investigation could be explored in future experiments by looking at the effect of feedback manipulations on participants' F3 values. 2) Producibility of compensatory responses depends on articulatory constraints, also mentioned in the previous section. Responses for the shift towards /u/ in Experiment 1 may be following in nature because an opposing response to the shift would require an articulatory configuration that would produce a formant pattern lying outside the vowel space for most people, i.e. an articulatory configuration that has never been achieved by the participant during speaking. In a study investigating compensation strategies for labial perturbation of the rounded vowel [u] (Savariaux et al., 1995), the authors suggest that complete acoustic compensation may be impossible due to speaker-dependent articulatory constraints. They further suggest that these constraints are due more to speaker-specific internal representation of articulatory-to-acoustic relationships rather than to any anatomical or neurophysiological limitations. In our study, these constraints on production may manifest themselves as a following response.

In Experiment 2, following responses were observed for shifts towards /u/, /æ/ and /ɑ/. Previous studies have suggested that large feedback shifts can be interpreted as targets rather than production errors (Burnett et al., 1998; Behroozmand et al., 2012) causing a following response. Here, we suggest that this following phenomenon can also be observed for smaller shifts.

2.5.4 Implications for models of speech production

Our results have implications for current models of sensorimotor behaviour in speech production (Houde and Nagarajan, 2011; Tourville and Guenther, 2011; Parrell et al., 2019). While in theory these models include influences beyond auditory feedback that control responses to auditory feedback perturbations, these models have not adequately elaborated in detail the effects of extra-auditory influences like somaesthesis and articulatory constraints on determining auditory feedback responses. Incorporating the effects of these extra-auditory influences in models of speech motor control would provide a quantitative framework to assess whether these factors alone or their combination can lead to the pattern of direction dependence of adaptation we found in this study.

2.5.5 Limitations

It is important to acknowledge the limitations of our study. Firstly, although the participants were all English-speakers, they were recruited from the San Francisco Bay Area where people come from diverse linguistic backgrounds (Hall-Lew, 2010) and may be exposed to varying degrees of acoustic and articulatory goals during everyday speech. Although we used shifts tailored to each subject's vowel space in this study, how these findings vary across different linguistic groups needs to be further explored. Secondly, the shifts that we used do not cover the whole gamut of an individual's vowel space. We had to limit the scope of our study, due to constraints like experimental time, to six shifts that we felt were a good representation of back vowels, front vowels, closed vowels and open vowels. While there's a general relation between vowel height-backness and F1-F2 (Stevens and House, 1955; Fant, 1960), the actual relationship is more complex (Stevens, 1998) and future studies should look at articulatory measures along with

acoustic measures. Thirdly, our focus was on the first two formants that were shifted and analysed in the current experimental design. In order to look at phenomena like lip-rounding in rounded vowels where F3 is implicated (Stevens and House, 1955), we would track and shift F3 in such experiments. Lastly, we did not test our participants for differences in vowel discriminability and vowel category boundaries. The shifts applied in Experiment 2 were equal in absolute frequency values but may not be equal on a psychoacoustic scale that takes perceptual differences into account. Future versions of this experiment could include perceptual testing and perceptually-equal shifts.

2.6 References

Behroozmand, R., Korzyukov, O., Sattler, L., and Larson, C. R. (2012). "Opposing and following vocal responses to pitch-shifted auditory feedback: Evidence for different mechanisms of voice pitch control," The Journal of the Acoustical Society of America 132, 2468-2477.

Benjamini, Y., and Hochberg, Y. (1995). "Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing," J R Stat Soc B 57, 289-300.

Berens, P. (2009). "CircStat: a MATLAB toolbox for circular statistics," J Stat Softw 31, 1-21.

Bourguignon, N. J., Baum, S. R., and Shiller, D. M. (2014). "Lexical-perceptual integration influences sensorimotor adaptation in speech," Frontiers in human neuroscience 8, 208.

Burnett, T. A., Freedland, M. B., Larson, C. R., and Hain, T. C. (1998). "Voice F0 responses to manipulations in pitch feedback," The Journal of the Acoustical Society of America 103, 3153-3161.

Caudrelier, T., and Rochet-Capellan, A. (2019). "Changes in speech production in response to formant perturbations: An overview of two decades of research," (Peter Lang).

Cremers, J., and Klugkist, I. (2018). "One Direction? A Tutorial for Circular Data Analysis Using R With Examples in Cognitive Psychology," Frontiers in psychology 9.

Daliri, A., and Dittman, J. (2019). "Successful auditory motor adaptation requires task-relevant auditory errors," Journal of neurophysiology 122, 552-562.

Fairbanks, G. (1954). "Systematic research in experimental phonetics:* 1. A theory of the speech mechanism as a servosystem," Journal of speech and Hearing Disorders 19, 133-139.

Fant, G. (1960). Acoustic theory of speech production (Mouton, s'Gravenhage,).

Gick, B., Wilson, I., and Derrick, D. (2013). Articulatory phonetics (Wiley-Blackwell, Malden, MA).

Greenwood, D. D. (1997). "The Mel Scale's disqualifying bias and a consistency of pitch-difference equisections in 1956 with equal cochlear distances and equal frequency ratios," Hearing research 103, 199-224.

Guenther, F. H. (2016). Neural control of speech (MIT Press).

Hall-Lew, L. (2010). "Ethnicity and Sociolinguistic Variation in San Francisco," Language and Linguistics Compass 4, 458-472.

Harrison, D., and Kanji, G. K. (1988). "The development of analysis of variance for circular data," Journal of Applied Statistics 15, 197-223.

Harshman, R., Ladefoged, P., and Goldstein, L. (1977). "Factor analysis of tongue shapes," The Journal of the Acoustical Society of America 62, 693-707.

Hillenbrand, J. M., Clark, M. J., and Houde, R. A. (2000). "Some effects of duration on vowel recognition," The Journal of the Acoustical Society of America 108, 3013-3022.

Houde, J. F., and Jordan, M. I. (1998). "Sensorimotor adaptation in speech production," Science 279, 1213-1216.

Houde, J. F., and Nagarajan, S. S. (2011). "Speech production as state feedback control," Frontiers in human neuroscience 5, 82.

Houde, J. F., Nagarajan, S. S., Sekihara, K., and Merzenich, M. M. (2002). "Modulation of the auditory cortex during speech: an MEG study," Journal of cognitive neuroscience 14, 1125-1138.

House, A. S. (1961). "On vowel duration in English," The Journal of the Acoustical Society of America 33, 1174-1178.

House, A. S., and Fairbanks, G. (1953). "The influence of consonant environment upon the secondary acoustical characteristics of vowels," The Journal of the Acoustical Society of America 25, 105-113.

Johnson, K. (2012). Acoustic and auditory phonetics (Wiley-Blackwell, Malden, MA).

Johnson, K. (2018). "Vocal tract length normalization," UC Berkeley PhonLab Annual Report 14.

Jones, J. A., and Munhall, K. G. (2000). "Perceptual calibration of F0 production: evidence from feedback perturbation," The Journal of the Acoustical Society of America 108, 1246-1251.

Katseff, S., Houde, J., and Johnson, K. (2012). "Partial compensation for altered auditory feedback: a tradeoff with somatosensory feedback?," Language and speech 55, 295-308.

Kitago, T., Ryan, S. L., Mazzoni, P., Krakauer, J. W., and Haith, A. M. (2013). "Unlearning versus savings in visuomotor adaptation: comparing effects of washout, passage of time, and removal of errors on motor memory," Frontiers in human neuroscience 7, 307.

Kluender, K. R., Diehl, R. L., and Wright, B. A. (1988). "Vowel-length differences before voiced and voiceless consonants: An auditory explanation," Journal of phonetics 16, 153-169.

Korzyukov, O., Bronder, A., Lee, Y., Patel, S., and Larson, C. R. (2017). "Bioelectrical brain effects of one's own voice identification in pitch of voice auditory feedback," Neuropsychologia 101, 106-114.

Kuhl, P. K. (1991). "Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not," Percept Psychophys 50, 93-107.

Lametti, D. R., Nasir, S. M., and Ostry, D. J. (2012). "Sensory preference in speech production revealed by simultaneous alteration of auditory and somatosensory feedback," The Journal of neuroscience : the official journal of the Society for Neuroscience 32, 9351-9358.

Lametti, D. R., Smith, H. J., Watkins, K. E., and Shiller, D. M. (2018). "Robust Sensorimotor Learning during Variable Sentence-Level Speech," Current biology : CB 28, 3106-3113 e3102.

Lee, B. S. (1950). "Some effects of side-tone delay," The Journal of the Acoustical Society of America 22, 639-640.

Levelt, W. J. (1983). "Monitoring and self-repair in speech," Cognition 14, 41-104.

Levelt, W. J. M. (1993). Speaking: From intention to articulation (MIT press).

Littell, R. C., Milliken, G. A., Stroup, W. W., Wolfinger, R. D., and Schabenberber, O. (2006). SAS for Mixed Models, Second Edition (SAS Publishing).

Lockhart, R. A., and Stephens, M. A. (1985). "Tests of fit for the von Mises distribution," Biometrika 72, 647-652.

Mardia, K. V., and Jupp, P. E. (2009). Directional statistics (John Wiley & Sons).

Mitsuya, T., MacDonald, E. N., Munhall, K. G., and Purcell, D. W. (2015). "Formant compensation for auditory feedback with English vowels," The Journal of the Acoustical Society of America 138, 413-424.

Mitsuya, T., Samson, F., Ménard, L., and Munhall, K. G. (2013). "Language dependent vowel representation in speech production," The Journal of the Acoustical Society of America 133, 2993-3003.

Niziolek, C. A., and Guenther, F. H. (2013). "Vowel category boundaries enhance cortical and behavioral responses to speech feedback alterations," The Journal of neuroscience : the official journal of the Society for Neuroscience 33, 12090-12098.

O'Shaughnessy, D. (1987). Speech communication : human and machine (Addison-Wesley, Reading, Mass.).

Parrell, B., Ramanarayanan, V., Nagarajan, S., and Houde, J. (2019). "The FACTS model of speech motor control: Fusing state estimation and task-based control," PLoS computational biology 15, e1007321.

Perrier, P., Perkell, J., Payan, Y., Zandipour, M., Guenther, F., and Khalighi, A. (2007). "Degrees of freedom of tongue movements in speech may be constrained by biomechanics," arXiv preprint arXiv:0709.1405.

Pisoni, D. B., and Remez, R. E. (2005). The handbook of speech perception (Wiley Online Library).

Purcell, D. W., and Munhall, K. G. (2006). "Adaptive control of vowel formant frequency: evidence from real-time formant manipulation," The Journal of the Acoustical Society of America 120, 966-977.

Quatieri, T., and McAulay, R. (1986). "Speech transformations based on a sinusoidal representation," IEEE Transactions on Acoustics, Speech, and Signal Processing 34, 1449-1464.

Rositzke, H. A. (1939). "Vowel-length in General American speech," Language 15, 99-109.

Savariaux, C., Perrier, P., and Orliaguet, J. P. (1995). "Compensation strategies for the perturbation of the rounded vowel [u] using a lip tube: A study of the control space in speech production," The Journal of the Acoustical Society of America 98, 2428-2442.

Schuerman, W. L., Nagarajan, S., McQueen, J. M., and Houde, J. (2017). "Sensorimotor adaptation affects perceptual compensation for coarticulation," The Journal of the Acoustical Society of America 141, 2693-2704.

Shiller, D. M., Sato, M., Gracco, V. L., and Baum, S. R. (2009). "Perceptual recalibration of speech sounds following speech motor learning," The Journal of the Acoustical Society of America 125, 1103-1113.

Stevens, K. N. (1968). The Quantal Nature of Speech: Evidence from Articulatory-acoustic Data.

Stevens, K. N. (1989). "On the quantal nature of speech," Journal of phonetics 17, 3-45.

Stevens, K. N. (1998). Acoustic phonetics (MIT Press, Cambridge, Mass.).

Stevens, K. N., and House, A. S. (1955). "Development of a quantitative description of vowel articulation," The Journal of the Acoustical Society of America 27, 484-493.

Subramaniam, K., Kothare, H., Mizuiri, D., Nagarajan, S. S., and Houde, J. F. (2018). "Reality Monitoring and Feedback Control of Speech Production Are Related Through Self-Agency," Frontiers in human neuroscience 12, 82.

Takano, S., and Honda, K. (2007). "An MRI analysis of the extrinsic tongue muscles during vowel production," Speech communication 49, 49-58.

Tourville, J. A., and Guenther, F. H. (2011). "The DIVA model: A neural theory of speech acquisition and production," Language and cognitive processes 26, 952-981.

von Mises, R. (1918). "Über die" Ganzzahligkeit" der Atomgewichte und verwandte Fragen," Physikal. Z. 19, 490-500.

Watson, G. S. (1961). "Goodness-of-fit tests on a circle," Biometrika 48, 109-114.

Watson, G. S., and Williams, E. J. (1956). "On the construction of significance tests on the circle and the sphere," Biometrika 43, 344-352.

Wells, J. C. (1962). "A study of the formants of the pure vowels of British English," (University of London London, UK).

Wells, J. C. (1990). "Syllabification and allophony," Studies in the pronunciation of English: A commemorative volume in honour of AC Gimson, 76-86.

Chapter 3: Temporal specificity of abnormal neural oscillations during phonatory events in Adductor Laryngeal Dystonia

3.1 Abstract

Laryngeal Dystonia (LD) is a debilitating disorder of voicing in which the laryngeal muscles are intermittently in spasm. This prevents the vocal folds from vibrating efficiently and results in involuntary interruptions during speech. The underlying causes of LD remain largely unknown. Prior imaging studies have found aberrant activity in the central nervous system during LD phonation. However, these studies could not resolve at what timepoints during phonation these abnormalities emerge. To investigate this question, we used magnetoencephalography (MEG) to monitor neural activity and associated behavioural responses time-locked to glottal movement onset, voice onset and onset of an externally-applied pitch feedback perturbation in adductor LD patients and controls. MEG scanning was performed in 17 patients and 12 controls. Four additional patients participated only in speech psychophysics studies without imaging. During scanning, subjects were prompted to start vocalising the vowel /ɑ/ and hold it for the duration of the prompt (2.4 s). Glottal onset was recorded using surface electromyography of pre-phonatory laryngeal muscular activity. On every trial, between 200ms and 500ms after voice onset, the pitch of their auditory feedback was briefly perturbed by +/- 100 cents for a period of 400ms and vocal responses to this change were recorded. We examined induced beta-band (12-30 Hz) and high-gamma-band (65-150 Hz) neural oscillations time-locked to glottal movement onset, voice onset and pitch perturbation onset in patients and controls and performed non-parametric statistical tests to observe group differences. Patients showed an elongated interval between laryngeal movement onset and phonatory onset. Patients exhibited abnormal neural activity around glottal movement onset, voice onset and even after pitch perturbation onset in both beta

band and high gamma band. Notably, patients' vocal response to pitch perturbation was not different from that in controls. The results suggest that patients with LD have significant abnormalities in neural control of speech around various phonatory events at almost all nodes of the speech motor control network. They exhibit impairment in preparatory or feedforward aspects of vocal control along with abnormal processing of sensory feedback.

3.2 Introduction

Laryngeal Dystonia (LD), also known as Spasmodic Dysphonia, is a voice disorder of neurological aetiology that affects the laryngeal muscles causing intermittent spasms (Ludlow, 2011). These spasms prevent the vocal folds from vibrating efficiently and result in involuntary interruptions only during voiced speech and not during other vocalisations like coughing and laughing (Bloch *et al.*, 1985), making it a focal or task-specific dystonia (Ludlow, 2011). LD affects approximately one in 100,000 people (Castelon Konkiewitz *et al.*, 2002; Ludlow *et al.*, 2008; Steeves *et al.*, 2012) and is a chronic condition with largely ambiguous or unknown causes. More than 80% of patients with LD are impacted by the adductor type of LD which affects the adductor muscles that bring the vocal folds closer (or adduct the vocal folds) causing them to shut too tightly, thus disrupting the initiation of voicing (Parnes *et al.*, 1978; Nash and Ludlow, 1996; Blitzer, 2010; Ludlow, 2011).

Currently, treatment options for LD are limited (Sulica, 2004) with most patients opting for temporary symptom relief in the form of speech therapy and/or Botulinum toxin (botox) injections (Ludlow *et al.*, 1988; Murry and Woodson, 1995; Boutsen *et al.*, 2002; Paniello *et al.*, 2008). Enhanced knowledge of abnormalities in the central nervous system (CNS) may

contribute towards the design of novel and perhaps more effective treatments for LD and also help us better understand how existing treatments work or can be improved.

A number of studies have found that distinct areas of the CNS exhibit abnormalities in LD. Some studies have looked for structural abnormalities via post-mortem analysis (Simonyan *et al.*, 2010) or MRI-based morphometry measurements of grey matter volume, cortical thickness, cortical surface area and local white matter integrity (Simonyan and Ludlow, 2012; Kostic *et al.*, 2016; Waugh *et al.*, 2016; Kirke *et al.*, 2017; Bianchi *et al.*, 2019). Others have looked for abnormal activity during speaking via Positron Emission Tomography (PET) (Ali *et al.*, 2006) or Functional Magnetic Resonance Imaging (fMRI) (Haslinger *et al.*, 2005; Simonyan and Ludlow, 2010; Kirke *et al.*, 2017; Kiyuna *et al.*, 2017). Taken together, these studies have found LD-associated abnormalities in: (1) areas of the CNS that exhibit abnormalities across all dystonias: primary motor cortex (M1; especially laryngeal motor cortex [LMC]), primary somatosensory cortex (S1), thalamus, cerebellum, and basal ganglia; (2) areas of the parietal and premotor cortices that a recent study has shown are associated with task-specific dystonias (Bianchi *et al.*, 2019); (3) areas more specifically associated with speech motor control and speech processing (Price, 2012): e.g., the inferior frontal gyrus (IFG), middle frontal gyrus (MFG), superior temporal gyrus (STG), middle temporal gyrus (MTG), and the frontal operculum (FOp). LMC in particular has been studied with a number of differing modalities. Studies of evoked response potentials (Khosravani *et al.*, 2019) and studies using transcranial magnetic stimulation (Samargia *et al.*, 2014, 2016) have generally shown hyper-excitability of LMC in LD.

Notwithstanding the commonalities, the studies investigating abnormal activity during speaking in adductor LD have shown conflicting findings (see Supp. Table B.1 for a meta-analytical synopsis of these studies) with both increases and decreases in activity and connectivity within the speech motor control network in LD. One possible source of these discrepancies is the lack of temporal resolution in fMRI and PET imaging (Gosseries *et al.*, 2008). Vocal motor control is a dynamic process that involves the orchestration of various parts of the brain, respiratory muscles and speech articulators (Simonyan and Horwitz, 2011). There are a number of events related to onset of phonation and according to models of speech motor control (Houde and Nagarajan, 2011; Houde and Chang, 2015; Guenther, 2016; Parrell and Houde, 2019), the different phonation events engage different control processes. Neural activity before initial glottal closure is thought to be preparatory or related to feedforward control of speech. Neural activity immediately after glottal closure also includes responses to somatosensory feedback. After the onset of phonation, neural activity also includes responses to the onset of auditory feedback. Thus, in addition to knowing where in the central nervous system aberrant activity occurs it is also important to know during which vocal events (e.g. initial glottal closure, beginning of phonation or voice onset, sustained phonation) abnormalities related to LD emerge.

A neuroimaging modality with the temporal resolution needed to examine the rapid sequence of activations associated with phonation onset is magnetoencephalography (MEG). MEG in combination with advanced source reconstruction algorithms makes it possible to preserve spatial resolution while examining cortical activity on the order of milliseconds. In this study, we examined neural activity using MEG during the onset and continuation of phonation in patients with Adductor Laryngeal Dystonia (henceforth referred to as patients with LD) and compared it

with neural activity in a control group. To the best of our knowledge, no prior study has looked at the time course of cortical activity prior to and during sustained phonation in LD (but see Khosravani *et al*., 2019 for an analysis of spectral power using electroencephalography (EEG) during early vocalisation and late vocalisation in patients with LD). Furthermore, to isolate whether LD involves deficits in auditory feedback processing, we also briefly perturbed the pitch of participants' auditory feedback during sustained phonation. Studies have shown that participants tend to compensate for this auditory feedback perturbation by making changes to their vocal motor output (Burnett *et al.*, 1998; Kort *et al.*, 2014). Thus, perturbing pitch feedback not only allows us to test vocal behavioural responses to an altered auditory feedback but also examine neural activity induced by the onset of perturbation.

In sum, the study design enabled us to make the following inferences:

1. Abnormal activity in the premotor and motor cortices prior to glottal movement onset would mean impairment in feedforward or preparatory control of voicing.
2. Abnormalities after glottal movement onset but prior to phonation (voice onset) would suggest deficits in somatosensory feedback processing whereas abnormalities after voice onset would suggest deficits in somatosensory and/or auditory feedback processing.
3. Abnormalities after the onset of pitch perturbation would suggest abnormal motor responses to auditory feedback.

3.3 Materials and methods

3.3.1 Participants

For this research study, 22 patients (15 female, Mean age = 57.38 years, Standard deviation = 9.69 years) were recruited from the University of California, San Francisco (UCSF) Voice and Swallowing Center and through postings on the National Spasmodic Dysphonia Association's website. Additionally, 13 controls (6 female, Mean age = 49.45 years, Standard deviation = 17.7 years) were recruited by word-of-mouth and from healthy research cohorts at the UCSF Memory and Aging Center. A two-sample heteroscedastic t-test was used to determine that the cohorts did not differ in age ($p = 0.156$, see Table 3.1). 17 patients and 13 controls took part in the imaging part of the study which consisted of magnetoencephalography (MEG) and structural MRI acquisition. Additionally, these 17 patients and 13 controls underwent a thorough clinical voice evaluation by author SS. Five additional patients participated only in the pitch perturbation task without imaging. MEG data from two patients and two controls had to be excluded because of large dental or movement artefacts. Vocal behavioural data for five of the 22 patients had to be excluded because of poor pitch tracking. All patients were examined and diagnosed by a team of laryngologists and speech-language pathologists. Two of the 22 patients were also diagnosed to have vocal tremor. Data from five additional controls belonging to the same age range was included to improve the signal-to-noise ratio only for high-gamma-band analysis locked to pitch perturbation onset. Eligibility criteria for patients were: 1. a diagnosis of adductor LD, 2. symptomaticity during research participation. Eligibility criteria for control participants were: 1. no structural abnormalities in their MRI, 2. no hearing loss and 3. absence of neurological disorders. The study was approved by the Committee on Human Research of the University of California, San Francisco. Informed consent was obtained from each participant prior to the study.

Table 3.1 Participant demographics and Voice Evaluation Measurements

Demographics	Controls (n = 13)	Patients with LD (n = 22)	p-value
Age in years	49.45 ± 17.70	57.38 ± 9.69	0.156
Females, number (%)	6 (46.15)	15 (68.18)	0.288
White race, number (%)	8 (61.54)	17 (77.27)	0.444

Values for Age are means ± standard deviation, age range is 27.4-71.34 for controls and 28.83-72.46 for patients. Statistical tests: unpaired two-tailed t-test for age; Fisher's exact test for sex and race. Race was self-reported by participants.

Measure	Controls (n = 13)	Patients with LD (n = 17)	p-value
Voice Handicap Index (VHI) (Range = 0 - 120)	2 (0 - 8)	63 (57 - 66)	3.834×10^{-6}
Consensus Auditory Perception Evaluation of Voice (CAPE-V), Overall Severity (Range = 0 - 100)	5 (0 - 6)	30 (12 - 40)	8.721×10^{-6}
Laryngeal Diadochokinesis (L-DDK) Rate (syllables per second)	5.39 ± 0.70	4.46 ± 0.60	0.001

Scores are median (lower quartile - upper quartile) for VHI and CAPE-V.
Scores are means ± standard deviation for L-DDK Rate.
Statistical tests: Wilcoxon rank - sum test for VHI and CAPE-V; unpaired two-tailed t-test for L-DDK Rate.

3.3.2 MRI Acquisition

T1-weighted structural MRI images were acquired using a 3-Tesla MRI scanner (Discovery MR750, GE Medical Systems) with an 8-channel head coil at the UCSF Margaret Hart Surbeck Laboratory for Advanced Imaging. An inversion recovery spoiled gradient echo sequence was used to acquire 128 axial slices (repetition time = 6.664ms, echo time = 2.472ms, inversion time = 900ms, slice thickness = 1mm, in-plane voxel dimensions = 0.5mm x 0.5mm). Individual structural MRIs were spatially normalised to a standard Montreal Neurological Institute (MNI) template using SPM8 (https://www.fil.ion.ucl.ac.uk/spm/software/spm8/).

3.3.3 MEG Imaging and Experimental Design

The MEG scanner used for imaging was a CTF system (Coquitlam, British Columbia, Canada) consisting of 275 axial gradiometers. Participants were scanned in supine position and signals were collected at a sampling rate of 1200Hz. Head position was recorded relative to the sensor array using three fiducial coils placed at the nasion and left and right preauricular points. These points were co-registered with individual spatially-normalised structural MRI images to generate head shapes.

The experimental task consisted of 120 trials. During each trial, participants were prompted to start vocalising the vowel /ɑ/ upon seeing a green dot on a screen in front of them (see Fig. 3.1). They were asked to hold the phonation for the duration of the visual prompt (~2.4s). Pre-phonatory laryngeal muscular activity (glottal movement onset) was recorded in participants using surface EMG. First, an abrasive gel was used to prepare the skin and to lower the impedance. Then, two of the surface EMG electrodes were pasted on either side of the larynx over the thyroid cartilage,

a third electrode was placed infrahyoid and midline superior to the two laryngeal electrodes and EMG signal was recorded using the double differential technique. Conductive paste was used to increase conductivity between the skin and the electrodes. The ground electrode was placed on the participant's forehead. Simultaneous electrocardiography (ECG) and electrooculography (EOG) signals were also collected to ensure that the electrophysiological signals picked up by the EMG electrodes were devoid of noise from eye blinks, eye movements and heartbeats. Participants' vocal output was picked up by an optical microphone (Phone-Or Ltd., Or-Yehuda, Israel), passed through a digital signal processing system (DSP) and fed back to them via insert earphones (ER-3A, Etymotic Research, Inc., Elk Grove Village, IL). On every trial, between 200ms and 500ms after voice onset, the DSP perturbed the pitch of their auditory feedback by either +100 cents or -100 cents ($1/12^{th}$ of an octave) for a period of 400ms and vocal responses to this change were recorded. The DSP was implemented on a PC as a real-time vocoder program called Feedback Utility for Speech Processing (FUSP) which has been used in previous studies in the lab (Kort *et al.*, 2014; Naunheim *et al.*, 2019; Ranasinghe *et al.*, 2019).

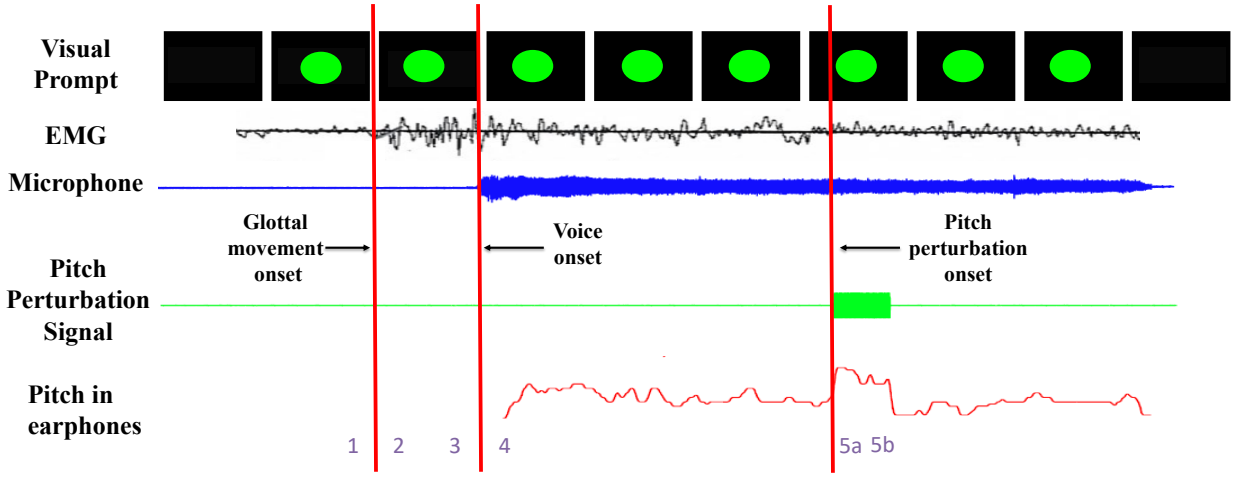


Fig. 3.1 Schematic diagram of the task

Participants were prompted to vocalise the vowel /a/ into a microphone for as long as they saw the green dot on the display. Participants could hear themselves through earphones throughout every trial. Glottal movement onset was measured using surface electromyography. Subjects' pitch was altered using a digital signal processing unit, between 200-500ms after voice onset, either up or down by 100 cents (1/12th of an octave) for 400ms and sent this shifted signal to the participants' earphones.

The numbers at the bottom of the figure represent various time windows as follows: 1 = before glottal movement onset, 2 = after glottal movement onset, 3 = before voice onset, 4 = after voice onset, 5a = early response to pitch perturbation onset, 5b = late response to pitch perturbation onset.

3.3.4 Data Analysis

3.3.4.1 Pitch Data Processing

Audio signals for participants' speech and the altered feedback were both collected at 11,025 Hz. The pitch time course for phonation in each trial was determined using an autocorrelation-based pitch tracking method (Parsons, 1987). All trials were visually inspected. Pitch tracks for all the trials were extracted and were aligned from 200ms before perturbation onset to 1000ms after perturbation onset. Trials with pitch tracking errors or with incomplete utterances were marked

bad and excluded. For the remaining good trials, pitch was converted from Hertz to cents using the following formula:

$$\mathrm{Pitch}_{\mathrm{cents}}(t) = 1200 \log_2\left(\frac{\mathrm{Pitch}_{\mathrm{Hz}}(t)}{\mathrm{Pitch}_{\mathrm{ref}}}\right)$$

where $\mathrm{Pitch}_{\mathrm{Hz}}(t)$ is the pitch value in Hertz at timepoint t and $\mathrm{Pitch}_{\mathrm{ref}}$ is the reference pitch calculated as the mean pitch during a window spanning from 50ms prior to perturbation onset to 50ms after perturbation onset. Participants responded to the feedback perturbation by deviating from their baseline pitch track. For each participant, absolute responses to both upward and downward perturbations were calculated and pooled together. To do this, the downward responses to upward perturbations were flipped and combined with upward responses to downward perturbations, thus making all compensatory responses positive. Mean responses for patients and controls were plotted. For time windows showing qualitative differences in responses, a linear mixed effects model was run in SAS 9.4 (SAS Institute Inc., Cary, NC) with group as the independent variable and pitch response in cents as the dependent variable. To account for multiple timepoints from each participant's mean vocal pitch response time-course' participant identity was included as a repeated measure with a random intercept.

Additionally, peak deviation from the baseline pitch track was calculated for every trial. This peak deviation had a positive sign if the response was opposing the shift (or compensatory) and negative if the response was following the shift. The mean across trials of this peak deviation was called the mean compensation to pitch perturbation for each participant.

3.3.4.2 Baseline pitch variability

To examine whether patients with LD had the laryngeal motor control capacity and the pitch range to compensate for pitch perturbation, we examined baseline pitch variability, in patients and controls, both within-trial and across trials during a pre-perturbation baseline time window defined from 200ms before perturbation onset to perturbation onset. Two-sample heteroscedastic t-tests were used to compare variability in patients with that in controls.

3.3.4.3 Correlations between responses to pitch perturbation and voice evaluation measures

To examine whether compensatory responses to pitch perturbation can predict disease severity, we selected three measures of voice impairment (see Table 3.1) from the participants' clinical voice evaluation: i) A subjective measure: Voice Handicap Index (VHI), a validated 30-item self-reported voice assessment (Jacobson *et al.*, 1997), ii) A perceptual measure: Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V) (Zraick *et al.*, 2011) and a iii) A laryngeal motor control measure: the Laryngeal Diadochokinesis rate (L-DDK rate) (Verdolini and Palmer, 1997; Bodenlos, 2013; Lombard and Solomon, 2019). Three Pearson's correlation coefficients were calculated between participants' mean compensation to pitch perturbation and each of the above measures.

3.3.4.4 MEG Data Processing

Third-order gradient noise correction and Direct Current (DC) offset correction were performed on the MEG datasets. A notch filter was implemented at 120Hz (width = 2Hz) to reduce power line noise. EMG signals, voice signals and MEG signals were examined simultaneously and three markers were added to mark the onset of glottal movement, onset of voicing and onset of pitch

perturbation. Subsequent neural analyses were locked with respect to these phonatory event markers. Trials with abnormal signals due to head movement, eye blinks or saccades were excluded. We focused our analyses on neural responses in the beta band (12-30Hz) and high gamma band (65-150Hz). Signals in beta band have been shown to play a role in motor planning, top-down motor-auditory interactions and the production of overt speech (Hinkley *et al.*, 2016; Sörös *et al.*, 2017; Abbasi and Gross, 2020). Also, our previous studies have shown that pitch perturbations affect high-gamma-band oscillations (Kort *et al.*, 2016; Ranasinghe *et al.*, 2019). Spatiotemporal source localisation for induced neural activity was mapped on to the spatially-normalised structural MRI for each individual subject using time-frequency-optimised adaptive spatial filtering (8mm lead field) in the Neurodynamic Utility Toolbox for MEG (NUTMEG: http://nutmeg.berkeley.edu) (Dalal *et al.*, 2008; Hinkley *et al.*, 2020). This source space reconstruction provided a voxel-by-voxel estimate of neural activity derived using a linear combination of a spatial weight matrix and a sensor data matrix. Active windows were defined as the time period after each of the three phonatory event markers. Noise-corrected pseudo-F statistics were computed by comparing the active window to a control window (pre-stimulus for glottal movement onset and voice onset analysis, pre-perturbation for perturbation onset analysis). Within-group and between-group statistical analyses were performed using statistical non-parametric mapping methods incorporated into the NUTMEG toolbox (Dalal *et al.*, 2011). For voice onset analysis, the phonatory onset interval or the duration between glottal movement onset and voice onset was included as a covariate in the statistical model. To correct for multiple comparisons across time and space, corrected p-value thresholds were calculated for $\alpha = 0.05$ and a False Discovery Rate (FDR) of 5%. Furthermore, cluster correction was performed to exclude

clusters with less than 18 congruent voxels as done in previous studies (Naunheim *et al.*, 2019). These steps minimised the possibility of observing spurious effects due to noise.

3.3.4.5 Phonatory onset interval analysis

Using the glottal movement onset marker and the voice onset marker added during the MEG analysis, phonatory onset interval was calculated for every trial in every participant. A two-sample heteroscedastic t-test was run to look at differences in the mean phonatory onset interval in patients as compared to controls, which has previously been shown to be elongated in LD patients (Ludlow and Connor, 1987).

3.3.4.6 Time windows of interest

To better interpret the differences in activity between patients and controls, the entire time course of the trial was divided into six time windows around the three phonatory event markers (see Fig. 3.1). Time window 1 extended from 125ms before and up to glottal movement onset. Time window 2 extended from glottal movement onset to 125ms after onset. Time window 3 consisted the 125ms preceding voice onset, while the 125ms immediately succeeding voice onset constituted time window 4. Time windows 5a and 5b were the early (0-225ms) and late (225-425ms) responses to pitch perturbation onset. These analyses shown in Fig. 3.2 to 3.7 make reference to these time windows.

3.4 Results

3.4.1 Behavioural results

3.4.1.1 Larger phonatory onset interval in patients with LD

Analysis of the time interval between glottal movement onset and voice onset markers (time windows 2 and 3) in the current study showed (Fig. 3.2) a significantly larger duration ($p = 1.14 \times 10^{-47}$; two-sample heteroscedastic t-test) in LD patients (Mean = 356.5ms, Standard Error of the mean = 4.94ms) than in controls (Mean = 251.55ms, Standard Error of the mean = 4.1ms). These results are consistent with a previous study showing voice onset delay in patients with LD (Ludlow and Connor, 1987).

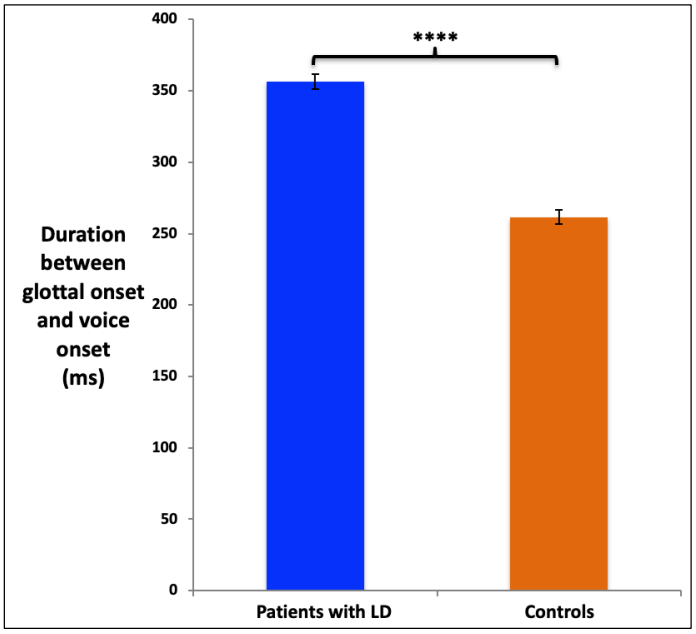


Fig. 3.2 Phonatory onset interval
Patients with LD have a larger phonatory onset interval (duration between glottal movement onset and voice onset) as compared to controls ($p = 1.140 \times 10^{-47}$; two-sample heteroscedastic t-test). Error bars indicate standard error of the mean.

3.4.1.2 Pitch perturbation vocal responses do not differ in patients with LD

No statistical differences were found between patients' and controls' vocal responses to pitch perturbation (Fig. 3.3A). Although there appeared to be a slight reduction in patients' vocal response from 250 to 400ms after pitch perturbation onset (part of time window 5b), this reduction was not significantly different from controls' vocal response in the same window [LMM, $F(1, 2001) = 0.01$, $p = 0.909$].

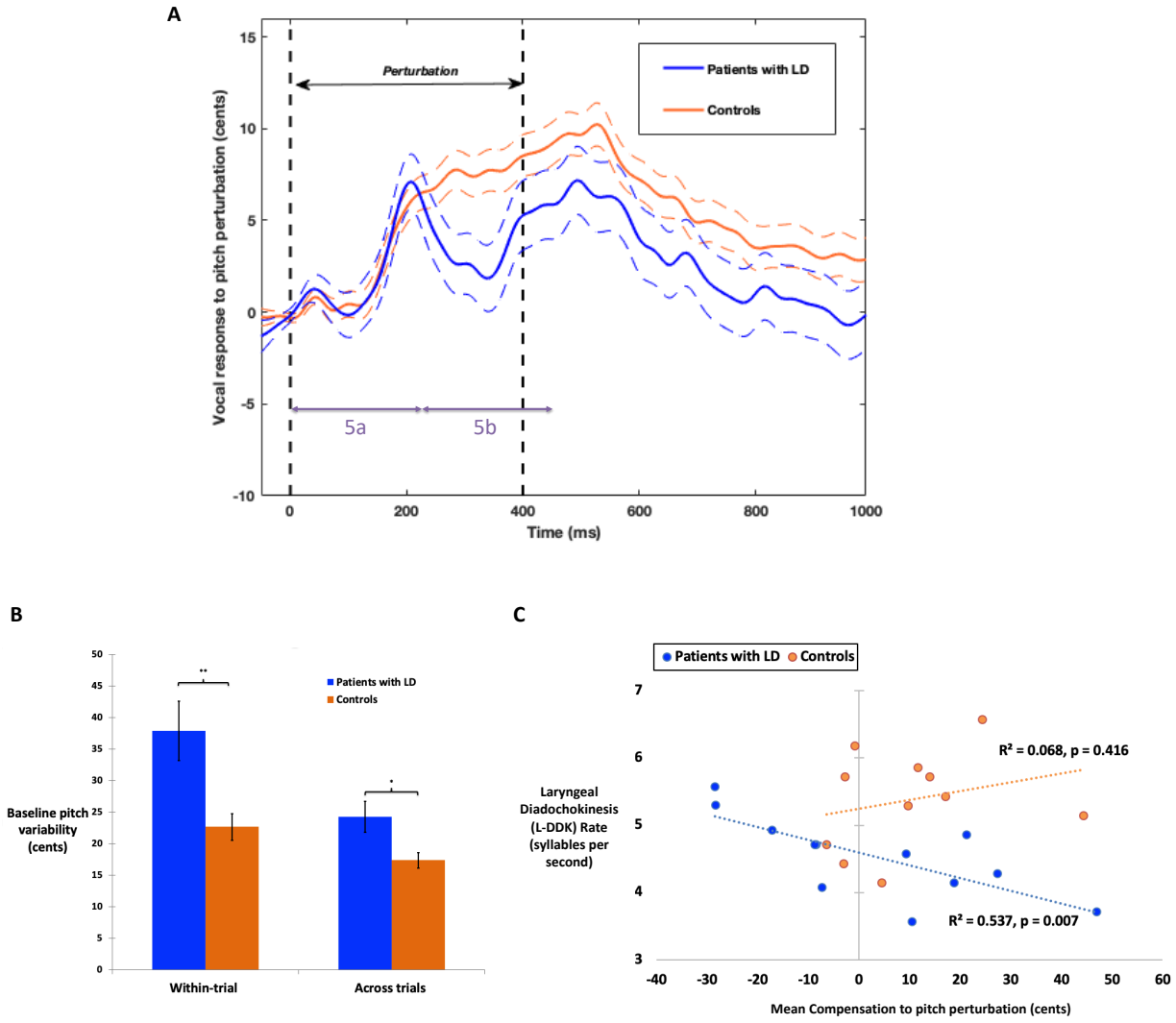


Fig. 3.3 Pitch control in patients with LD and controls

(A) *Vocal response to pitch perturbation: Patients with LD (n = 17) have a response to pitch perturbation that is not statistically different from that in controls (n = 12). The solid lines are the mean responses and the flanking dashed lines indicate standard error of the mean responses. Time windows 5a and 5b represent early and late response to pitch perturbation* ***(B)*** *Baseline vocal range: Baseline pitch variability (200ms prior to perturbation onset) in patients with LD (n = 15) differs from that in controls (n = 12) both within-trial and across trials. Error bars indicate standard error of the mean variability.* ***(C)*** *Correlation between mean compensation and laryngeal diadochokinesis (L-DDK) rate: In patients with LD, the L-DDK rate is negatively correlated with mean compensation to pitch perturbation. Patients (n = 12) who have a higher syllable rate tend to follow the direction of the pitch shift and the ones with a lower syllable rate tend to have a larger compensation value for pitch shifts. This correlation does not hold true in controls (n = 11).*

3.4.1.3 Larger baseline pitch variability in patients with LD

Patients showed larger pitch variability in cents both within-trial ($p = 0.008$; two-sample heteroscedastic t-test) and across trials ($p = 0.02$; two-sample heteroscedastic t-test) in a 200ms pre-perturbation time window, i.e. time window 4 (Fig. 3.3B) indicating that patients not only possessed the vocal range to compensate for pitch perturbation but also, in fact, had greater variations in laryngeal control during sustained phonation, perhaps owing to the presence of spasms.

3.4.1.4 Compensation to pitch perturbation in LD predicted severity of disease

Although no correlations were found, in both patients and controls, between mean compensation and VHI ($R^2 = 0.004$, $p = 0.839$ in patients and $R^2 = 0.150$, $p = 0.239$ in controls) or mean compensation and CAPE-V Overall Severity ($R^2 = 0.037$, $p = 0.549$ in patients and $R^2 = 0.016$, $p = 0.709$ in controls) , there was a negative correlation (Fig. 3.3C) between mean compensation and patients' L-DDK rate ($R^2 = 0.537$, $p = 0.007$). Patients with lower L-DDK rates, i.e. with greater disease severity, had greater values of mean compensation. Patients with higher L-DDK rates, i.e. with lesser disease severity, either had smaller values of mean compensation or tended to follow the direction of the pitch shift. There was no correlation between controls' mean compensation values and L-DDK rates ($R^2 = 0.068$, $p = 0.416$).

3.4.2 Neural activity

3.4.2.1 Bilateral reduction in inferior frontal and increase in parietal beta-band activity around glottal movement onset

Widespread differences in beta-band neural activity between controls and patients with LD were observed before and after glottal movement onset (Fig. 3.4A, Supp. Table B.2). Even prior to glottal movement (time window 1), the greatest reductions in beta-band activity in patients were seen in the left ventral motor cortex (vMC), the left ventral premotor cortex (vPMC) and the left inferior frontal gyrus (IFG). The reduced activity in the left vPMC and the left IFG persisted after glottal movement onset (time window 2), but differences were smaller after onset. Reduced activity was also observed in the anterior left STG, anterior left MTG, right middle frontal gyrus (MFG), right cerebellum and right IFG after glottal movement onset. In contrast, patients with LD also showed persistently increased activity, both prior to and after glottal movement onset, bilaterally in a large cluster in the inferior parietal lobule (IPL), around and posterior to the angular gyrus. Increased activity was also observed in the right superior frontal gyrus (SFG) or dorsal premotor cortex from -25ms before glottal movement onset to +125ms after glottal movement onset and the left dorsal premotor cortex at +125ms after glottal movement onset.

3.4.2.2 Reduced right cerebellar activation before and increased bilateral cortical activation after glottal movement onset in high-gamma band

Differences in high-gamma-band neural activity between controls and patients with LD showed an interesting change in polarity from before glottal movement onset to after glottal movement onset (Fig. 3.4B, Supp.Table B.3). From -125 to -75ms before glottal movement onset (time window 1), patients with LD showed reduced activity in the right cerebellum. However, after

glottal movement onset (time window 2), patients showed increased activity in the left primary somatosensory cortex (postcentral gyrus/S1) from +25 to +125ms. Increased activity in patients with LD was also observed in the right anterior temporal lobe and right IFG from +75 to +125ms after glottal movement onset. At +125ms after glottal movement onset, patients also showed increased high-gamma-band activity bilaterally in the STG and MTG.

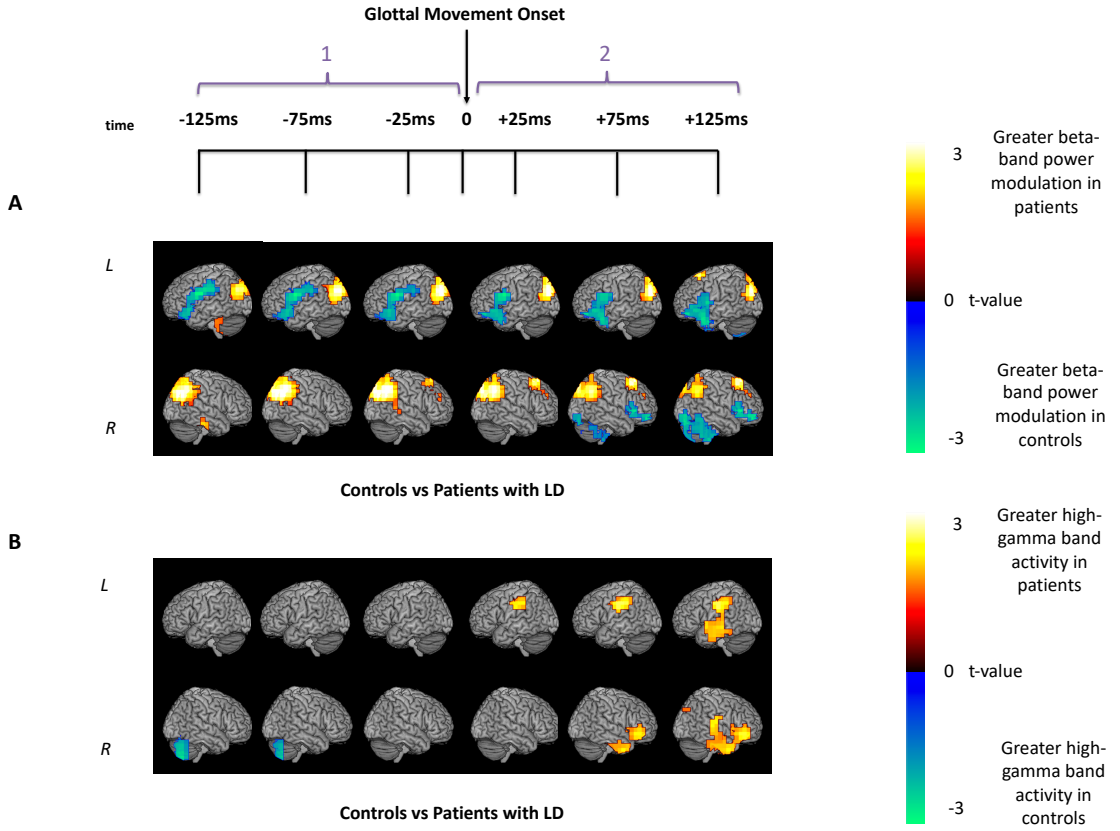


Fig. 3.4 Differences in neural activity around glottal movement onset between controls and patients with LD

A. *Non-phase-locked beta-band activity (12-30Hz) differences between patients and controls locked to glottal movement onset. As compared to controls (n = 11), patients with LD (n = 15) show significant differences in beta-band activity both before and after glottal movement onset. False Discovery Rate (FDR) correction for a rate of 5% and cluster correction at a threshold of 18 voxels and $p < 0.05$ were performed. For beta-band neural activity locked to glottal movement onset in each group alone, refer to Supp. Fig. B.1.* ***B.*** *Non-phase-locked high-gamma-band activity (65-150Hz) differences between patients and controls locked to glottal movement onset. As compared to controls (n = 11), patients with LD (n = 15) show significantly increased*

activity after glottal movement onset. False Discovery Rate (FDR) correction for a rate of 5% and cluster correction at a threshold of 18 voxels and $p < 0.05$ were performed. For high-gamma-band neural activity locked to glottal movement onset in each group alone, refer to Supp. Fig. B.4.

3.4.2.3 Bilateral increase in dorsal sensorimotor cortical activation and reduced cerebellar activation in beta band around voice onset

Prior to voice onset in beta band (time window 3), bilateral hyperactivity in patients' IPLs and hypoactivity in their left IFG and anterior left STG and MTG was observed (Fig. 3.5A, Supp. Table B.4), reflecting differences persisting after glottal movement onset (time window 2) observed in Fig. 3.4A. Importantly, both before and after voice onset, patients showed increasing bilateral hyperactivity along the primary somatosensory cortex (postcentral gyrus/S1) and the superior parietal lobule (SPL). Patients also showed hyperactivity in the superior left MFG from -125ms to +25ms. In contrast, patients showed reduced activity bilaterally in the cerebellum and the inferior occipital lobe (greater differences in the right hemisphere) after voice onset (time window 4). Patients also showed reduced activity in the right MFG and right IFG after voice onset. Note that these results were not impacted by differences in phonatory onset interval between the two groups, because this interval was included as a covariate in the statistical analysis to compare neural activity locked to voice onset.

3.4.2.4 Bilateral increase in activation in ventral sensorimotor cortex, prefrontal cortex and temporal lobe in high-gamma band around voice onset

Patients with LD showed widespread hyperactivity in high-gamma band both before and after voice onset (Fig. 3.5B, Supp. Table B.5). This hyperactivity was seen bilaterally in the ventral

sensorimotor cortex, temporal lobe, medial and ventral parts of the left prefrontal cortex and right ventral prefrontal cortex. Hyperactivity in the left temporal lobe increased from before voice onset (time window 3) to after voice onset (time window 4) whereas hyperactivity in the right temporal lobe decreased along with time. Hyperactivity was also observed in patients with LD in the right precuneus from -125 to -75ms before voice onset (time window 3). Patients with LD showed hypoactivity in the right inferior parietal lobule at +75ms after voice onset (time window 4).

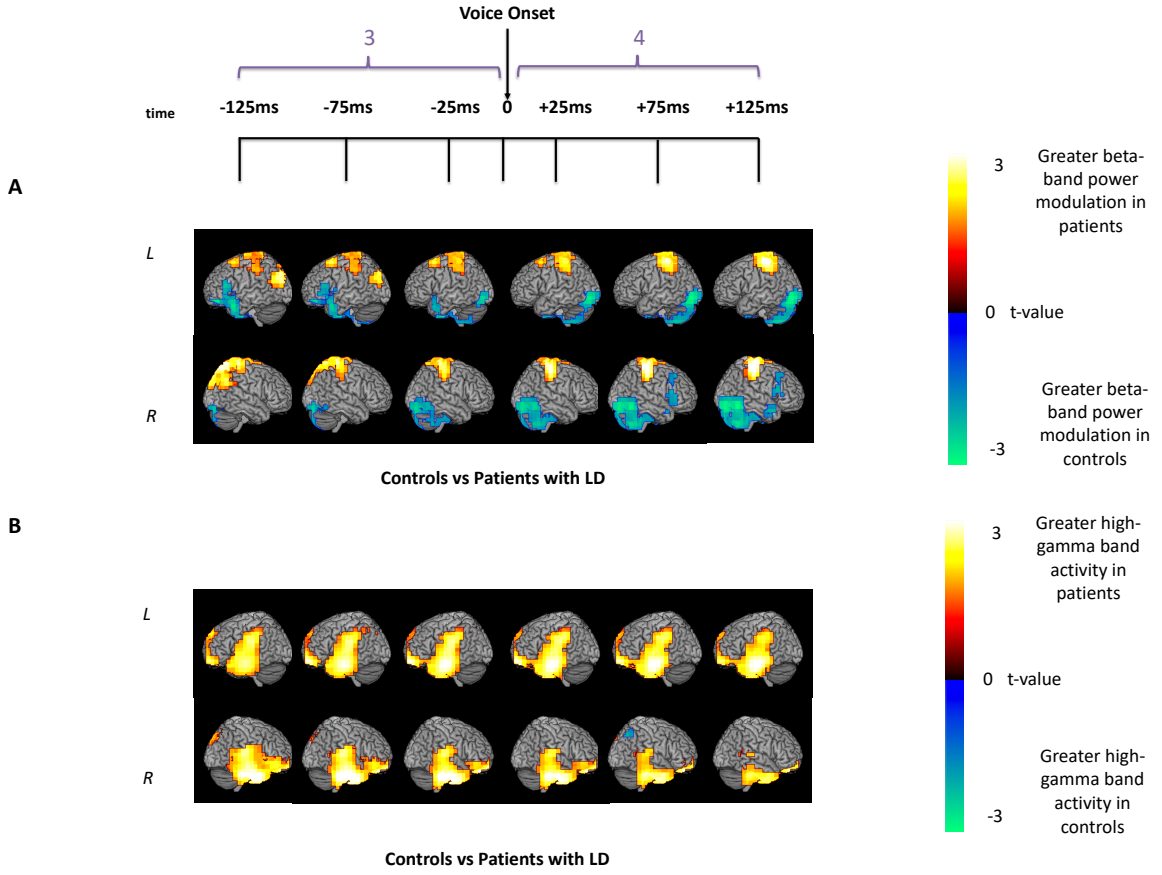


Fig. 3.5 Differences in neural activity around voice onset between controls and patients with LD

A. *Non-phase-locked beta-band activity (12-30Hz) differences between patients and controls locked to voice onset. Phonatory onset interval was added as a covariate in the statistical analysis. As compared to controls (n = 11), patients with SD (n = 15) show significant*

differences both before and after voice onset. False Discovery Rate (FDR) correction for a rate of 5% and cluster correction at a threshold of 18 voxels and $p < 0.05$ were performed. For beta-band neural activity locked to voice onset in each group alone, refer to Supp. Fig. B.2. ***B.*** *Non-phase-locked high-gamma-band activity (65-150Hz) differences between patients and controls locked to voice onset. Phonatory onset interval was added as a covariate in the statistical analysis. As compared to controls (n = 11), patients with SD (n = 15) show consistent differences in both hemispheres from before voice onset through voice onset. These differences increase after voice onset in the left hemisphere and decrease in the right hemisphere. False Discovery Rate (FDR) correction for a rate of 5% and cluster correction at a threshold of 18 voxels and $p < 0.05$ were performed. For high-gamma-band neural activity locked to voice onset in each group alone, refer to Supp. Fig. B.5.*

3.4.2.5 Bilateral increased frontoparietal beta-band activity after pitch perturbation onset

Examination of beta-band neural activity locked to pitch perturbation onset (Fig. 3.6A, Supp. Table B.6) showed that patients had greater activity bilaterally in the IFG and in the right vPMC and vMC with activity peaking from 125 to 225ms after perturbation onset. Frontal hyperactivity in the right hemisphere was more widespread than in the left hemisphere. Patients also showed greater activity in the left SPL from 175 to 375ms, right IPL from 375 to 425ms, right SFG from 225 to 425ms and anterior right STG from 25 to 425ms. Although this parietal hyperactivity was bilateral, the left hemisphere led the right hemisphere in terms of time of activity increase. Cerebellar activations were mixed – there was early left hemisphere reduction (time window 5a) and right hemispheric activity increased (time window 5b) paralleling the increased activity in the right frontal lobe.

Note that all the above neural differences were observed although there were no behavioural differences in the vocal pitch responses to pitch feedback perturbations.

3.4.2.6 Bilateral reduced cerebellar, prefrontal and temporal high-gamma-band activity after pitch perturbation onset

Patients with LD showed widespread reduced high-gamma-band activity after pitch perturbation onset (Fig. 3.6B, Supp.Table B.7). Hypoactivity in the left cerebellum preceded that in the right cerebellum and also peaked earlier. Bilateral hypoactivity in the MFG and IFG dissipated as time progressed whereas that in the ventral prefrontal cortex persisted from 25 to 425ms. There was significant hypoactivity in the left inferior temporal lobe from 175 to 375ms and from 125 to 425 in the right temporal and occipital lobes. Hyperactivity in patients with LD was also observed in the left inferior parietal lobule around and posterior to the angular gyrus from 175 to 275ms after pitch perturbation onset and the right dorsal superior frontal gyrus at 325ms.

Again, note that all the above neural differences were observed although there were no behavioural differences in the vocal pitch responses to pitch feedback perturbations.

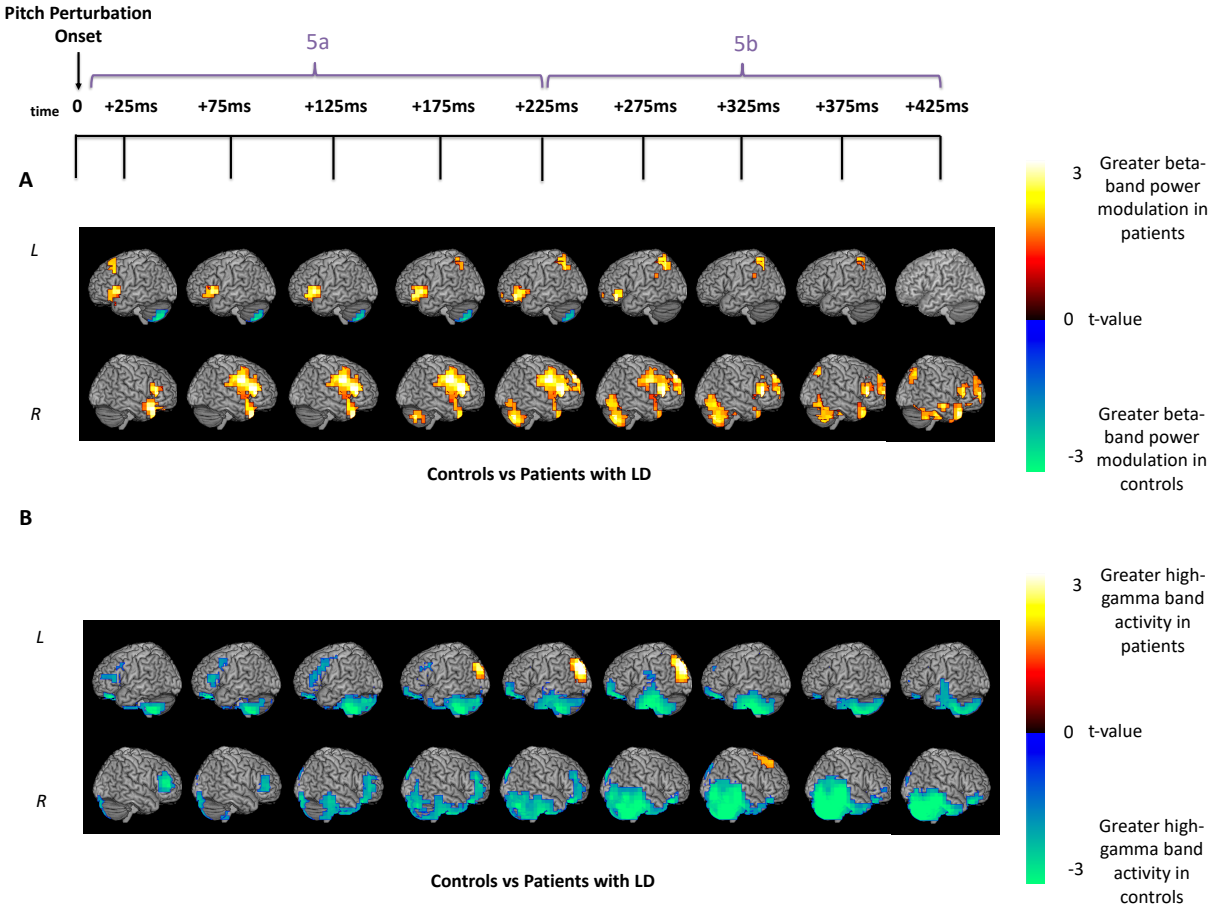


Fig. 3.6 Differences in neural activity around pitch perturbation onset between controls and patients with LD

A. *Non-phase-locked beta-band activity (12-30Hz) differences between patients and controls locked to pitch perturbation onset. Patients with LD (n = 15) show greater beta-band activity as compared to controls (n = 11) in a number of regions. False Discovery Rate (FDR) correction for a rate of 5% and cluster correction at a threshold of 18 voxels and $p < 0.05$ were performed. For beta-band neural activity locked to pitch perturbation onset in each group alone, refer to Supp. Fig. B.3.* ***B.*** *Non-phase-locked high-gamma-band activity (65-150Hz) differences between patients and controls locked to pitch perturbation onset. Patients with LD (n = 15) show mostly lesser high-gamma-band activity as compared to controls (n = 16) along with greater activity in some regions. False Discovery Rate (FDR) correction for a rate of 5% and cluster correction at a threshold of 18 voxels and $p < 0.05$ were performed. For high-gamma-band neural activity locked to pitch perturbation onset in each group alone, refer to Supp. Fig. B.6.*

3.5 Discussion

Laryngeal Dystonia is a focal or task-specific dystonia with symptoms arising only during voiced speech. While steady-state phonation is not symptom-free in LD, the fact that repeated productions

induce more symptoms (Simonyan and Ludlow, 2010) suggests that phonation onsets are particularly problematic for patients with LD. Investigating the timing of abnormal activations around phonatory events may help shed light on a number of unanswered questions about LD. The current study sought to examine the spatiotemporal neural dynamics of phonatory control during the initiation of sustained phonation in patients with LD. Differences in neural activity were found between patients and controls at various timepoints during phonation. In order to interpret these findings, we will refer to the State Feedback Control (SFC) model of speech production (Houde and Nagarajan, 2011; Houde and Chang, 2015).

The SFC model for phonation (see Fig. 3.7) begins with the idea that at any moment in time, the larynx has a laryngeal motor state and that depending on the state, it can produce either somatosensory feedback (like at glottal movement onset) or both somatosensory and auditory feedback (like at voice onset). When a speaker decides to phonate, the higher frontal cortex activates the phonation control network consisting of, but not limited to, the motor cortex, the somatosensory cortex and the auditory cortex. This activation of the phonation network causes the primary motor cortex (M1) to output laryngeal controls which initiate glottal movement and an efference copy of these controls drives the ventral premotor cortex (vPMC) to output a prediction of the somatosensory consequences of the glottal movement. In this study, we used surface EMG to detect when the larynx began to move in response to these controls and also began to generate somatosensory feedback. Any abnormalities in neural activity seen before this point would have to do with the output of the motor cortex and the higher frontal cortex (IFG) and the generation of feedback predictions by the premotor cortex. Once the somatosensory feedback reaches the primary somatosensory cortex (S1), it is compared with the predicted feedback. Any mismatch

with the prediction results in corrections to the estimated laryngeal state in the vPMC, which in turn further modifies the output of controls by M1. We used a microphone to monitor sound output from the larynx and thus detect voice onset. This onset of phonation generates both somatosensory and auditory feedback. Like somatosensory feedback, auditory feedback is also compared with a feedback prediction. Any mismatch with the auditory prediction results in corrections to the estimated state and output of laryngeal controls. Thus, any abnormalities seen after voice onset would indicate abnormal responses to either the somatosensory or auditory feedback of phonation. To isolate neural responses to only auditory feedback, we perturbed the pitch of the auditory feedback of the participants' sustained phonation. This perturbation creates a deliberate mismatch between the prediction and the actual sensory feedback and generates auditory feedback prediction errors in the auditory cortex. This error ideally causes a compensatory vocal response driven by M1 by shifting vocal pitch production in the direction opposite to the shift.

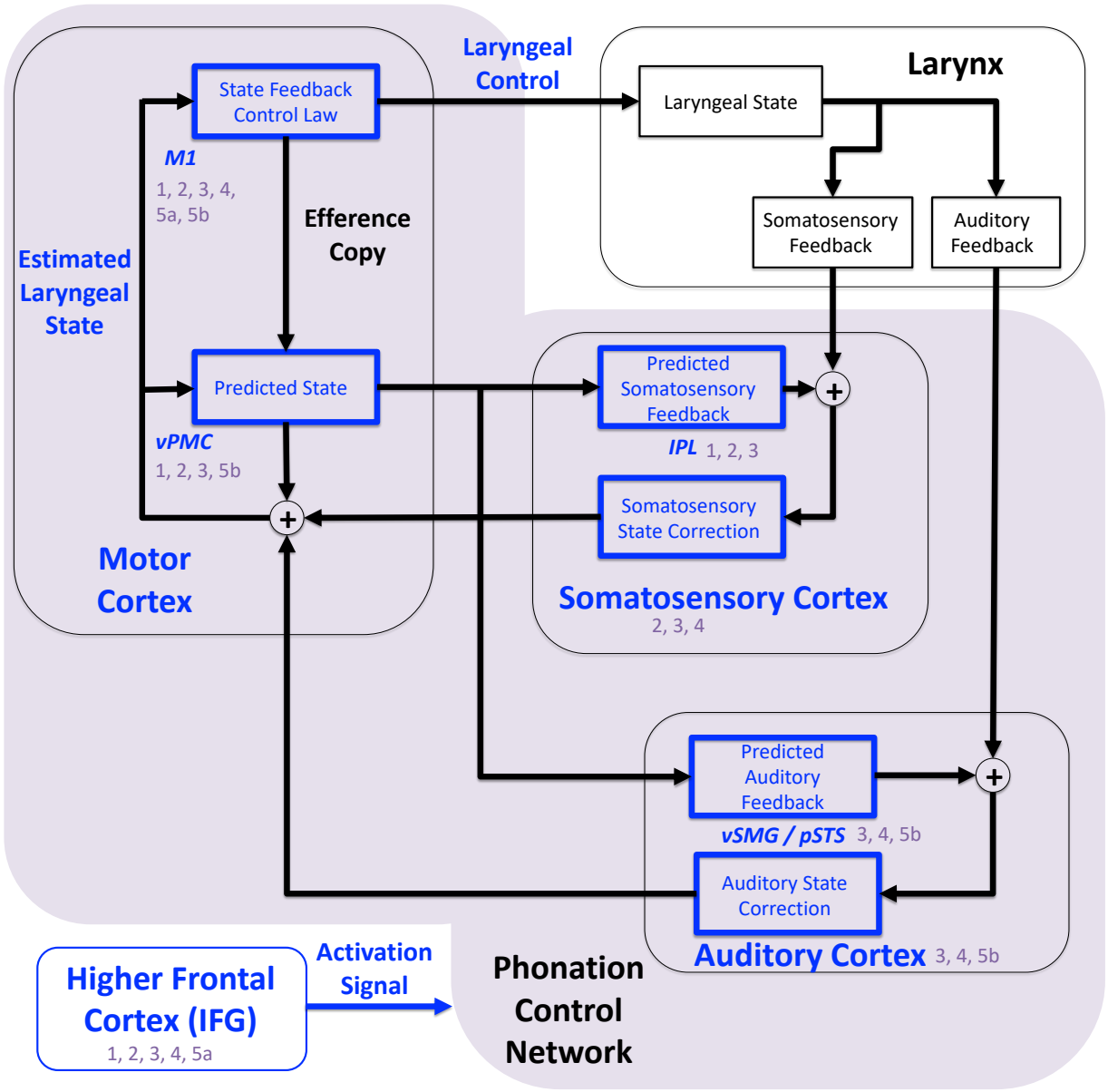


Fig. 3.7 Schematic diagram of the State Feedback Control (SFC) model and networks impacted in LD

According to the SFC model, laryngeal control is based on an estimate of the current laryngeal state maintained by a comparison between the predicted state and the incoming feedback. State corrections are generated when there is a mismatch between the predicted state and the actual state as conveyed by the feedback signals. Brain regions marked in blue appear to be abnormally impacted in LD. The numbers in purple indicate the time windows in which abnormalities are observed.

Abbreviations used: M1 = Primary Motor Cortex, vPMC = Ventral Premotor Cortex, IPL = Inferior Parietal Lobule, S1 = Primary Somatosensory Cortex, A1 = Primary Auditory Cortex, vSMG = Ventral Supramarginal Gyrus. pSTS = Posterior Superior Temporal Sulcus.

Comparison of beta-band neural activity between controls and patients with LD showed (Fig. 3.4) that even before glottal movement onset, patients showed reduced activity in the left vPMC (the location where predictions of the estimated laryngeal state are generated as proposed by the SFC model) and the left ventral motor cortex (VMC), the part of the motor cortex that is involved in laryngeal articulatory control (Simonyan and Horwitz, 2011; Bouchard *et al.*, 2013). This suggests difficulties in sending motor commands to the larynx from M1 and in generating predictions about the somatosensory consequences of these motor commands. Additionally, patients with LD also show hypoactivity in the right cerebellum before glottal movement onset in the high-gamma band. The cerebellum is involved in the co-ordination, planning and sequencing of movements (Manto *et al.*, 2012) and is implicated in various aspects of speech motor control (Ackermann *et al.*, 2007; Ackermann, 2008; Ghosh *et al.*, 2008; Parrell *et al.*, 2017; Lametti *et al.*, 2018; Houde *et al.*, 2019). The SFC model posits that along with the vPMC, the cerebellum predicts the next state of the vocal tract. Thus, abnormalities in cerebellar activation consistent with abnormalities in vPMC further suggest impaired speech motor planning and state prediction in patients with LD. Beta-band hypoactivity in the left VMC disappears after glottal movement onset indicating that the trouble may be in generating motor commands to initiate glottal movement but not to sustain phonation. However, hypoactivity in the left vPMC continues even after glottal onset suggesting a general impairment in generating feedforward somatosensory predictions of glottal movement.

Sustained bilateral beta-band hyperactivity in patients' inferior parietal lobule (IPL) both before and after glottal movement onset may perhaps be related to this inability to predict the somatosensory consequences of laryngeal control. Furthermore, patients show persistent hyperactivity in the high-gamma band after glottal movement onset in the left ventral

somatosensory cortex and the left supramarginal gyrus, indicative of an abnormal response to the onset of somatosensory feedback. This observation presents temporal specificity to the abnormal activation of the primary somatosensory cortex that has been observed in previous studies (Simonyan and Ludlow, 2010). Patients exhibit hyperactivity in the right anterior temporal lobe, right IFG, bilateral STG and MTG as they approach voice onset, perhaps indicative of hyperactivity associated with auditory predictions.

This high-gamma-band hyperactivity spreads bilaterally along the temporal lobes and continues even after voice onset. Beta-band hypoactivity in left vPMC and hyperactivity around the IPL also continues as patients approach voice onset. We speculate that impairment in generating feedforward predictions of the laryngeal state may also be responsible for the differences in phonatory onset interval observed in the current study as well as in previous studies (Ludlow and Connor, 1987). Additionally, the larger phonatory onset interval in patients with LD may also be due to more time needed to build up subglottal pressure sufficient to overcome their tightly adducted vocal folds. Patients show both beta-band and high-gamma-band hyperactivity in S1 (more dorsal in the beta band and more so in the left hemisphere in high-gamma band) in the time interval between glottal movement onset and voice onset and this hyperactivity continues after voice onset. This is again indicative of abnormal somatosensory feedback processing also seen in previous studies (Simonyan and Ludlow, 2010; Daliri *et al.*, 2020). Notably, this abnormal somatosensory feedback processing is not only observed after voice onset, when patients would begin to experience spasms, but also before voicing begins. This suggests that hyperactivity in somatosensory feedback processing may be a trait in patients with LD and not a state induced by symptom-eliciting tasks. The dorsal presentation of beta-band hyperactivity, around what is known

as the truncal sensorimotor cortex, in these patients before and after voice onset could perhaps indicate that the respiratory muscles work harder in patients with LD to push air out of a tightly closed glottis due to spasms in the adductor muscles in these patients. Also, beta-band hyperactivity after voice onset is limited to the somatosensory cortex and is markedly absent from the primary auditory cortex (A1) and other higher order regions involved in the processing of auditory feedback (posterior superior temporal sulcus (pSTS) and ventral supramarginal gyrus (vSMG)). Previous studies have found that auditory feedback control mechanisms do not contribute to cortical hyperactivity in the vocal motor control network of patients with LD (Daliri *et al.*, 2020). However, hyperactivity in the high-gamma band is observed in regions involved in auditory feedback processing both before and after voice onset. This hyperactivity may suggest that patients with LD do have trouble in generating auditory feedback predictions and the temporal granularity provided by MEG helps identify this abnormality.

To further investigate the role of auditory feedback in abnormalities in speech production in patients with LD, we perturbed the pitch of their auditory feedback during phonation. Behaviourally, patients show no differences from controls in generating a compensatory vocal response to this pitch shift indicating that their ability to generate corrective motor commands in response to auditory feedback prediction errors may be intact. However, when we take a look at neural responses to pitch perturbation, we see that there is increased cortical beta-band activity in patients, especially in the right vPMC and vMC regions known to be involved in feedback-based articulator control (Tourville *et al.*, 2008). This increased activity, along with hyperactivity in the left IFG, may be indicative of an increased effort in vocal control required to compensate for the auditory mismatch. In contrast, patients show widespread reduced activity in the high-gamma band

after pitch perturbation onset, notably bilaterally in the cerebellum and the frontal and temporal lobes. An interesting observation is that in both frequency bands, activation differences in the left cerebellum precede those in the right cerebellum. However, patients do show hyperactivity in the left posterior parieto-occipital junction midway into the perturbation when beta-band hyperactivity peaks in the right hemisphere. As participants compensate for pitch perturbation to reduce the auditory prediction error, there arises a somatosensory prediction error due to this motoric compensation. This hyperactivity in the high gamma band may then perhaps be the manifestation of abnormalities in somatosensory prediction because of a combination of spasms in the larynx and a mismatch in somatosensory prediction due to vocal compensation to altered auditory feedback.

Additionally, patients' mean compensatory response negatively correlates with their L-DDK rate while there is no relationship observed in controls' responses. Patients who don't perform well on the L-DDK task (i.e. those who produce fewer syllables per second) have a higher mean compensation to pitch perturbation. Patients whose laryngeal function is not as impaired tend to have a smaller mean response or a mean response in the direction of the applied shift. In a previous study by us (Subramaniam *et al.*, 2018), we found that participants who made smaller corrective responses to pitch perturbations are better at judging self-agency (or accurate identification of self-generated information) reflecting an increased reliance on internal predictions to guide their speech output rather than reliance on externally-altered feedback. Based on these findings, we can speculate that perhaps the less-impaired patients who have smaller responses to pitch perturbation rightly attribute the alteration in their auditory feedback to an external source and treat it as a vocal target instead of something that needs correction. Whereas the patients whose laryngeal control is

more impaired and have a higher mean compensatory response have less confidence in their vocal production, regard the altered auditory feedback as self-generated (perhaps due to their spasms) and motorically correct for this auditory mismatch.

We observed evidence in support of impaired laryngeal motor control in the form of significant hypoactivity in the left VMC and right cerebellum before glottal movement onset. Does impaired laryngeal motor control in LD arise in M1/cerebellum or is the hypoactivity a signature of motor impairments that are further downstream from the M1 (subcortical regions or musculature) reflected through cortico-thalamic or cortico-muscular loops? The motor cortex is thought to produce an internal copy of its motor output called the 'efference copy'(Sperry, 1950; Von Holst and Mittelstaedt, 1950) and is thought to be used as a reference to generate predictions of the sensory consequences of the action. When the auditory feedback detected is the same as expected feedback, there is what we call 'speaking-induced suppression' (SIS) or reduced neural activity in the auditory cortex (Creutzfeldt *et al.*, 1989; Curio *et al.*, 2000; Houde *et al.*, 2002; Ventura *et al.*, 2009; Flinker *et al.*, 2010). The SFC model states that this suppression occurs because the sensory outcome prediction of a given motor command attenuates the sensitivity of the auditory cortex to that particular outcome. Importantly, it has been shown that SIS does not take all production variability into account. There is greater suppression if the auditory feedback is closer to the median production and lesser suppression if the feedback deviates from the median production (Niziolek *et al.*, 2013), a phenomenon called SIS-falloff. The SFC model supposes that feedforward predictions should reflect any variability in feedback arising due to variations in motor cortex activity. Thus, SIS fall-off is thought to reflect variability in the speech production network downstream from the motor cortex, i.e. subcortical brain regions like the nucleus ambiguus, the

peripheral nervous system and the musculature. It can be seen from Fig. 3.3B that there is greater baseline pitch production variability in patients with LD than in controls. Looking at SIS modulation in patients and comparing it with that in controls would also help predict where exactly variability is introduced in the speech motor control system.

In conclusion, we found that patients with LD not only have impaired sensory feedback processing during phonation, as found in previous studies, but also impaired feedforward or preparatory activity even before the laryngeal muscles start moving as well as an abnormal response to somatosensory feedback after glottal movement onset. Patients with LD seem to exhibit abnormalities in neural activity in almost every node of the cortical speech motor control network and the cerebellum. These abnormalities arise at or around particular phonatory events, thus highlighting that speech motor control deficits in patients with LD possess temporal specificity along with a pan-network involvement.

As mentioned before, treatment options for LD are limited and focus on symptom alleviation. Modulation of activity in impacted brain regions using non-invasive brain stimulation tools like Transcranial Magnetic Stimulation (TMS) (Barker *et al.*, 1985) or Transcranial Direct Current Stimulation (tDCS) (Brunoni *et al.*, 2012) may be a viable path for long-term therapeutic purposes.

3.6 References

Abbasi O, Gross J. Beta-band oscillations play an essential role in motor–auditory interactions. Human brain mapping 2020; 41(3): 656-65.

Ackermann H. Cerebellar contributions to speech production and speech perception: psycholinguistic and neurobiological perspectives. Trends in neurosciences 2008; 31(6): 265-72.

Ackermann H, Mathiak K, Riecker A. The contribution of the cerebellum to speech production and speech perception: clinical and functional imaging data. The cerebellum 2007; 6(3): 202-13.

Ali SO, Thomassen M, Schulz GM, Hosey LA, Varga M, Ludlow CL, et al. Alterations in CNS Activity Induced by Botulinum Toxin Treatment in Spasmodic Dysphonia: An H215O PET Study. Journal of Speech, Language, and Hearing Research 2006; 49(5): 1127-46.

Barker AT, Jalinous R, Freeston IL. Non-invasive magnetic stimulation of human motor cortex. The Lancet 1985; 325(8437): 1106-7.

Bianchi S, Fuertinger S, Huddleston H, Frucht SJ, Simonyan K. Functional and structural neural bases of task specificity in isolated focal dystonia. Movement Disorders 2019; 34(4): 555-63.

Blitzer A. Spasmodic dysphonia and botulinum toxin: experience from the largest treatment series. European Journal of Neurology 2010; 17: 28-30.

Bloch CS, Hirano M, Gould WJ. Symptom improvement of spastic dysphonia in response to phonatory tasks. The Annals of otology, rhinology, and laryngology 1985; 94(1 Pt 1): 51-4.

Bodenlos ML. Measurement of Laryngeal Diadochokinesis in the Early Adult Population: Citeseer; 2013.

Bouchard KE, Mesgarani N, Johnson K, Chang EF. Functional organization of human sensorimotor cortex for speech articulation. Nature 2013; 495(7441): 327-32.

Brunoni AR, Nitsche MA, Bolognini N, Bikson M, Wagner T, Merabet L, et al. Clinical research with transcranial direct current stimulation (tDCS): challenges and future directions. Brain stimulation 2012; 5(3): 175-95.

Castelon Konkiewitz E, Trender-Gerhard I, Kamm C, Warner T, Ben-Shlomo Y, Gasser T, et al. Service-based survey of dystonia in munich. Neuroepidemiology 2002; 21(4): 202-6.

Creutzfeldt O, Ojemann G, Lettich E. Neuronal activity in the human lateral temporal lobe. II. Responses to the subjects own voice. Experimental brain research 1989; 77(3): 476-89.

Curio G, Neuloh G, Numminen J, Jousmäki V, Hari R. Speaking modifies voice-evoked activity in the human auditory cortex. Human brain mapping 2000; 9(4): 183-91.

Dalal SS, Guggisberg AG, Edwards E, Sekihara K, Findlay AM, Canolty RT, et al. Five-dimensional neuroimaging: localization of the time–frequency dynamics of cortical activity. NeuroImage 2008; 40(4): 1686-700.

Dalal SS, Zumer JM, Guggisberg AG, Trumpis M, Wong DD, Sekihara K, et al. MEG/EEG source reconstruction, statistical evaluation, and visualization with NUTMEG. Computational intelligence and neuroscience 2011; 2011: 758973.

Daliri A, Heller Murray ES, Blood AJ, Burns J, Noordzij JP, Nieto-Castanon A, et al. Auditory feedback control mechanisms do not contribute to cortical hyperactivity within the voice production network in adductor spasmodic dysphonia. Journal of Speech, Language, and Hearing Research 2020; 63(2): 421-32.

Flinker A, Chang EF, Kirsch HE, Barbaro NM, Crone NE, Knight RT. Single-trial speech suppression of auditory cortex activity in humans. Journal of Neuroscience 2010; 30(49): 16643-50.

Ghosh SS, Tourville JA, Guenther FH. A neuroimaging study of premotor lateralization and cerebellar involvement in the production of phonemes and syllables. Journal of Speech, Language, and Hearing Research 2008.

Gosseries O, Demertzi A, Noirhomme Q, Tshibanda J, Boly M, Op de Beeck M, et al. Functional neuroimaging (fMRI, PET and MEG): what do we measure? Revue médicale de Liège 2008; 63(5-6): 231-7.

Guenther FH. Neural control of speech; 2016.

Haslinger B, Erhard P, Dresel C, Castrop F, Roettinger M, Ceballos-Baumann AO. "Silent event-related" fMRI reveals reduced sensorimotor activation in laryngeal dystonia. Neurology 2005; 65(10): 1562-9.

Hinkley LB, Marco EJ, Brown EG, Bukshpun P, Gold J, Hill S, et al. The Contribution of the Corpus Callosum to Language Lateralization. The Journal of neuroscience : the official journal of the Society for Neuroscience 2016; 36(16): 4522-33.

Hinkley LBN, Dale CL, Cai C, Zumer J, Dalal S, Findlay A, et al. NUTMEG: Open Source Software for M/EEG Source Reconstruction. Frontiers in Neuroscience 2020; 14: 710.

Houde JF, Chang EF. The cortical computations underlying feedback control in vocal production. Current opinion in neurobiology 2015; 33: 174-81.

Houde JF, Gill JS, Agnew Z, Kothare H, Hickok G, Parrell B, et al. Abnormally increased vocal responses to pitch feedback perturbations in patients with cerebellar degeneration. The Journal of the Acoustical Society of America 2019; 145(5): EL372.

Houde JF, Nagarajan SS. Speech production as state feedback control. Frontiers in human neuroscience 2011; 5: 82.

Houde JF, Nagarajan SS, Sekihara K, Merzenich MM. Modulation of the auditory cortex during speech: an MEG study. Journal of cognitive neuroscience 2002; 14(8): 1125-38.

Jacobson BH, Johnson A, Grywalski C, Silbergleit A, Jacobson G, Benninger MS, et al. The voice handicap index (VHI) development and validation. American journal of speech-language pathology 1997; 6(3): 66-70.

Khosravani S, Mahnan A, Yeh IL, Watson PJ, Zhang Y, Goding G, et al. Atypical somatosensory-motor cortical response during vowel vocalization in spasmodic dysphonia. Clinical Neurophysiology 2019; 130(6): 1033-40.

Kirke DN, Battistella G, Kumar V, Rubien-Thomas E, Choy M, Rumbach A, et al. Neural correlates of dystonic tremor: a multimodal study of voice tremor in spasmodic dysphonia. Brain imaging and behavior 2017; 11(1): 166-75.

Kiyuna A, Kise N, Hiratsuka M, Kondo S, Uehara T, Maeda H, et al. Brain activity in patients with adductor spasmodic dysphonia detected by functional magnetic resonance imaging. Journal of Voice 2017; 31(3): 379. e1-. e11.

Kort NS, Cuesta P, Houde JF, Nagarajan SS. Bihemispheric network dynamics coordinating vocal feedback control. Human brain mapping 2016; 37(4): 1474-85.

Kort NS, Nagarajan SS, Houde JF. A bilateral cortical network responds to pitch perturbations in speech feedback. NeuroImage 2014; 86: 525-35.

Kostic VS, Agosta F, Sarro L, Tomić A, Kresojević N, Galantucci S, et al. Brain structural changes in spasmodic dysphonia: a multimodal magnetic resonance imaging study. Parkinsonism & related disorders 2016; 25: 78-84.

Lametti DR, Smith HJ, Freidin PF, Watkins KE. Cortico-cerebellar Networks Drive Sensorimotor Learning in Speech. Journal of cognitive neuroscience 2018; 30(4): 540-51.

Lombard L, Solomon NP. Laryngeal diadochokinesis across the adult lifespan. Journal of Voice 2019.

Ludlow CL. Spasmodic dysphonia: a laryngeal control disorder specific to speech. Journal of neuroscience 2011; 31(3): 793-7.

Ludlow CL, Adler CH, Berke GS, Bielamowicz SA, Blitzer A, Bressman SB, et al. Research priorities in spasmodic dysphonia. Otolaryngology--head and neck surgery : official journal of American Academy of Otolaryngology-Head and Neck Surgery 2008; 139(4): 495-505.

Ludlow CL, Connor NP. Dynamic aspects of phonatory control in spasmodic dysphonia. Journal of speech and hearing research 1987; 30(2): 197-206.

Manto M, Bower JM, Conforto AB, Delgado-García JM, Da Guarda SNF, Gerwig M, et al. Consensus paper: roles of the cerebellum in motor control—the diversity of ideas on cerebellar involvement in movement. The Cerebellum 2012; 11(2): 457-87.

Nash EA, Ludlow CL. Laryngeal muscle activity during speech breaks in adductor spasmodic dysphonia. The Laryngoscope 1996; 106(4): 484-9.

Naunheim ML, Yung KC, Schneider SL, Henderson-Sabes J, Kothare H, Hinkley LB, et al. Cortical networks for speech motor control in unilateral vocal fold paralysis. The Laryngoscope 2019; 129(9): 2125-30.

Niziolek CA, Nagarajan SS, Houde JF. What does motor efference copy represent? Evidence from speech production. The Journal of neuroscience : the official journal of the Society for Neuroscience 2013; 33(41): 16110-6.

Parnes SM, Lavorato AS, Myers EN. Study of spastic dysphonia using videofiberoptic laryngoscopy. Annals of Otology, Rhinology & Laryngology 1978; 87(3): 322-6.

Parrell B, Agnew Z, Nagarajan S, Houde J, Ivry RB. Impaired Feedforward Control and Enhanced Feedback Control of Speech in Patients with Cerebellar Degeneration. The Journal of neuroscience : the official journal of the Society for Neuroscience 2017; 37(38): 9249-58.

Parrell B, Houde J. Modeling the role of sensory feedback in speech motor control and learning. Journal of Speech, Language, and Hearing Research 2019; 62(8S): 2963-85.

Parsons TW. Voice and speech processing: McGraw-Hill College; 1987.

Price CJ. A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. NeuroImage 2012; 62(2): 816-47.

Ranasinghe KG, Kothare H, Kort N, Hinkley LB, Beagle AJ, Mizuiri D, et al. Neural correlates of abnormal auditory feedback processing during speech production in Alzheimer's disease. Sci Rep 2019; 9(1): 5686.

Samargia S, Schmidt R, Kimberley TJ. Shortened cortical silent period in adductor spasmodic dysphonia: evidence for widespread cortical excitability. Neuroscience letters 2014; 560: 12-5.

Samargia S, Schmidt R, Kimberley TJ. Cortical silent period reveals differences between adductor spasmodic dysphonia and muscle tension dysphonia. Neurorehabilitation and neural repair 2016; 30(3): 221-32.

Simonyan K, Horwitz B. Laryngeal motor cortex and control of speech in humans. The Neuroscientist 2011; 17(2): 197-208.

Simonyan K, Ludlow CL. Abnormal activation of the primary somatosensory cortex in spasmodic dysphonia: an fMRI study. Cerebral cortex 2010; 20(11): 2749-59.

Simonyan K, Ludlow CL. Abnormal structure-function relationship in spasmodic dysphonia. Cerebral cortex 2012; 22(2): 417-25.

Simonyan K, Ludlow CL, Vortmeyer AO. Brainstem pathology in spasmodic dysphonia. The Laryngoscope 2010; 120(1): 121-4.

Sörös P, Doñamayor N, Wittke C, Al-Khaled M, Brüggemann N, Münte TF. Increase in Beta-Band Activity during Preparation for Overt Speech in Patients with Parkinson's Disease. Frontiers in human neuroscience 2017; 11: 371.

Sperry RW. Neural basis of the spontaneous optokinetic response produced by visual inversion. Journal of comparative and physiological psychology 1950; 43(6): 482.

Steeves TD, Day L, Dykeman J, Jette N, Pringsheim T. The prevalence of primary dystonia: a systematic review and meta-analysis. Movement disorders : official journal of the Movement Disorder Society 2012; 27(14): 1789-96.

Subramaniam K, Kothare H, Mizuiri D, Nagarajan SS, Houde JF. Reality Monitoring and Feedback Control of Speech Production Are Related Through Self-Agency. Frontiers in human neuroscience 2018; 12(82).

Tourville JA, Reilly KJ, Guenther FH. Neural mechanisms underlying auditory feedback control of speech. NeuroImage 2008; 39(3): 1429-43.

Ventura MI, Nagarajan SS, Houde JF. Speech target modulates speaking induced suppression in auditory cortex. BMC Neurosci 2009; 10: 58.

Verdolini K, Palmer PM. Assessment of a" profiles" approach to voice screening. Journal of Medical Speech Language Pathology 1997; 5: 217-32.

Von Holst E, Mittelstaedt H. The reafference principle: interaction between the central nervous system and the periphery. Selected Papers of Erich von Holst: The Behavioural Physiology of Animals and Man. London: Methuen (From German) 1950; 1: 39-73.

Waugh JL, Kuster JK, Levenstein JM, Makris N, Multhaupt-Buell TJ, Sudarsky LR, et al. Thalamic volume is reduced in cervical and laryngeal dystonias. PloS one 2016; 11(5): e0155302.

Zraick RI, Kempster GB, Connor NP, Thibeault S, Klaben BK, Bursac Z, et al. Establishing validity of the consensus auditory-perceptual evaluation of voice (CAPE-V). American journal of speech-language pathology 2011.

Chapter 4: Cortical dynamics of speech motor control in the non-fluent variant of Primary Progressive Aphasia

4.1 Abstract

Primary Progressive Aphasia (PPA) is a clinical syndrome in which patients progressively lose speech and language abilities. The non-fluent variant of PPA (nfvPPA) is characterised by impaired motor speech and agrammatism. These deficits are associated with left fronto-insular-striatal atrophy and imaging studies further suggest impaired activity and functional connectivity of brain regions implicated in speech production. However, to date no study has either directly examined speech motor control in nfvPPA or documented the dynamics of the recruitment of the speech motor control network during vocal production. Using high-temporal magnetoencephalographic (MEG) imaging, we investigated vocal motor control in 18 patients with nfvPPA and 17 controls. Participants were prompted to phonate the vowel /ɑ/ for ~2.4s while the pitch of their feedback was shifted either up or down by 100 cents for a period of 400ms mid-utterance. Participants were unaware of but nevertheless responded to these pitch feedback perturbations by changes in vocal pitch. Task-induced neural oscillations were examined in the alpha (8-12 Hz) and beta bands (13-30 Hz) during pitch feedback perturbation. Nonparametric statistical tests were performed to look at neural activity differences in patients compared to healthy controls. In vocal pitch responses, nfvPPA patients showed a smaller compensation response to pitch perturbation than controls ($p<0.05$). Pre-perturbation variations in vocal pitch variability did not differ significantly between the two groups, indicating that reduced vocal compensation cannot simply be attributed to insufficient voice range in patients. Patients also exhibited reduced task-induced alpha-band neural activity in the right superior temporal gyrus, right superior temporal sulcus, right middle temporal gyrus and the right

temporoparietal junction (FDR-corrected and cluster-corrected at 20 voxels and $p < 0.01$) from 250ms to 750ms after pitch perturbation onset. Patients also showed increased task-induced beta-band activity in the left dorsal sensorimotor cortex, left premotor cortex and the left supplementary motor area (FDR-corrected and cluster-corrected at 20 voxels and $p < 0.01$) from 50ms to 150ms after pitch perturbation onset. Reduction in alpha-band power could predict speech motor impairment in patients with nfvPPA ($\beta = 3.41$, $F = 8.31$, $p = 0.0128$) whereas the increase in beta-band power could not ($\beta = -1.75$, $F = 1.72$, $p = 0.2123$). Collectively, these results suggest significant disruption in sensorimotor integration and control during vocal production in patients with nfvPPA.

4.2 Introduction

Primary progressive aphasia (PPA) is a clinical syndrome characterised by progressive loss of speech and language abilities (Mesulam, 2003). Currently accepted clinical classification (Gorno-Tempini *et al.*, 2011) identifies three subtypes of PPA, each with a characteristic predominant impairment in speech and language function, namely: logopenic variant of PPA (lvPPA), semantic variant of PPA (svPPA) and non-fluent variant of PPA (nfvPPA). Among these three variants, the nfvPPA subtype includes patients who predominantly exhibit motor-speech deficits. Structural magnetic resonance imaging (MRI) often used to complement clinical characterisation shows a signature pattern of neuronal atrophy in patients with nfvPPA over the left posterior fronto-insular region - encompassing areas involved in the production of speech and language like the inferior frontal gyrus (IFG), insula, premotor cortex and supplementary motor area (SMA) (Gunawardena *et al.*, 2010; Grossman, 2012; Leyton *et al.*, 2016). Furthermore, functional neuroimaging techniques have also shown significant alterations in

functional and structural connectivity of the cortical regions involved in speech production (Galantucci *et al.*, 2011; Mandelli *et al.*, 2014; Mandelli *et al.*, 2018). However, little is known about the dynamics of the recruitment of these cortical regions during speech production in patients with nfvPPA.

Current models of speech production posit that ongoing execution of speech engages a distributed network of frontal, parietal, and temporal cortical areas and describe the online control of speech as a process that involves continuous integration of sensory feedback with motor output (Houde and Nagarajan, 2011; Tourville and Guenther, 2011; Guenther, 2016; Parrell and Houde, 2019; Parrell *et al.*, 2019). Under normal speaking conditions, the sensory feedback matches closely with the motor predictions, hence the feedback prediction error or the difference between predictions and feedback registered by the sensorimotor integration network is negligible. The goal of the speech motor control system is to minimise this prediction error. If auditory feedback is perturbed so that feedback predictions are deliberately mismatched, an exaggerated prediction error is generated. The speaker then modifies their motor output to compensate for the perturbation and reduce the prediction error.

Given the speech motor control system's goal to reduce the prediction error, if the pitch of the auditory feedback is perturbed during speaking, the speaker's resulting behavioural vocal response to compensate for the prediction error is involuntary and is thus called the pitch perturbation reflex (Burnett *et al.*, 1998). This pitch perturbation reflex is an experimental paradigm that has been widely studied to examine sensorimotor integration in a number of neurological disorders (Mollaei *et al.*, 2016; Ranasinghe *et al.*, 2017a; Houde *et al.*, 2019). The

well-described anatomical and psychophysical details of the pitch perturbation reflex makes it an ideal tool to study the dynamics of the recruitment of speech motor cortical regions in patients with nfvPPA during speech production. The specific behavioural and neural deficits of speech motor integration, if any, will lend more insight on the functional deficits of speech production in nfvPPA.

Here, we examined the behavioural response during pitch perturbation reflex and the neural activity patterns of the underlying sensorimotor integration in patients with nfvPPA. Specifically, we compared vocal and neural responses induced by perturbing the pitch of the auditory feedback signal in patients with nfvPPA to those in healthy controls. We used the millisecond-precise temporal resolution of magnetoencephalography (MEG), to investigate the spatial and temporal patterns of the cortical speech motor network during the pitch perturbation reflex. We hypothesised that given the pathophysiological processes in nfvPPA that target brain regions involved in speech production, nfvPPA patients will exhibit abnormal behavioural and neural activity patterns during pitch feedback perturbation. We further predicted that the severity of speech motor impairment in nfvPPA patients would be associated with the degree of abnormal pitch perturbation response.

4.3 Materials and methods

4.3.1 Participants

For this study, 20 patients and 18 healthy controls (see Table 4.1 for demographic details) were recruited from the Memory and Aging Center at the University of California, San Francisco (UCSF). 17 patients and 14 controls took part in the pitch perturbation experiment with

simultaneous MEG imaging. Three patients and four controls took part in the behavioural pitch perturbation experiment without neuroimaging. Patients' PPA diagnosis and classification as nfvPPA was conducted by a team of expert neurologists. Eligibility criteria for controls included: no structural brain abnormalities, normal cognitive performance, absence of neurological or psychiatric disorders. Participants (or their assigned surrogate decision makers) provided informed consent before taking part in the study. This study was approved by the Committee on Human Research of the University of California, San Francisco.

4.3.2 Neuropsychological assessment

All participants underwent a Mini-Mental State Examination (MMSE). A Clinical Dementia Rating (CDR) score was also calculated for all participants after an interview with the participants and their caregivers (Morris, 1993). 19 of the 20 patients also underwent a comprehensive neuropsychological evaluation to test their performance on a number of speech and language tasks as described elsewhere (Ranasinghe *et al.*, 2017b) (see Table 4.1 for more details). The Western Aphasia Battery was used to evaluate speech and syntactic production (Kertesz, 2007). Apraxia of speech and dysarthria were rated using Motor Speech Evaluation (Wertz *et al.*, 1984). To measure long and short syntax comprehension, sentences were read aloud by the examiner and the patient had to select the picture that best matched the sentence from two options. Confrontation naming was assessed using a short form version (15 items) of the Boston Naming Test (Mack *et al.*, 1992). Patients were asked to name as many animals as possible in 60 seconds to assess category fluency (Spreen and Benton, 1977). A subset of 16 items from the Peabody Picture Vocabulary Test (PPVT; 4 items each: verbs, descriptive,

animate and inanimate) where patients were asked to match a word with one of four picture choices was used to test word comprehension (Dunn and Dunn, 2007).

Table 4.1 Participant demographics and neuropsychological test performance (speech and language battery) in patients with nfvPPA

Demographics	Controls (n = 18)	Patients with nfvPPA (n = 20)	p-value
Age in years	64.81 ± 5.76	67.79 ± 8.02	0.1936
Females, number (%)	13 (72.22)	14 (70)	1
White race, number (%)	18 (100)	19 (95)	1
Education in years	18.00 (16.75 - 18.25)	16.00 (14.00 - 17.5)	0.0083
Right handedness, number (%)	18 (100)	17 (85)	0.2319
MMSE	30.0 (29.0 - 30.0)	28.0 (24.0 - 29.0)	<0.0001
CDR Total	0.00 (0.00 - 0.00)	0.75 (0.00 - 1.50)	<0.0001

Values for Age are means ± standard deviation, age ranges are 56.21-76.55 for controls and 56.75-82.14 for patients.
Values for Education, MMSE, CDR Total are median and the lower and upper quartiles.
Statistical tests: unpaired two-tailed t-test for age; Fisher's exact test for sex, race and handedness; Wilcoxon rank-sum test for education, MMSE and CDR Total.

Neuropsychological test performance in patients

Test	Score
Western Aphasia Battery – Fluency (out of 10)	6 (5 - 9)
Western Aphasia Battery – Repetition (out of 100)	90 (82 - 96)
Western Aphasia Battery – Sequential Command (out of 80)	80 (68 - 80)
Apraxia of Speech rating (out of 7)	3 (2 - 4)
Dysarthria rating (out of 7)	2 (0 - 3)
Long Syntax Comprehension (Percentage)	93.75 (68.75 - 100)
Short Syntax Comprehension (Percentage)	100 (90.625 - 100)
Boston Naming Test (out of 15)	14 (13 - 15)
Category Fluency (Animals named / minute)	14 (8 - 19)
Peabody Picture Vocabulary Test (out of 16)	15 (14 - 16)

All scores are median (lower quartile - upper quartile).
n = 19 for all scores except for Long Syntax Comprehension (n = 17) and Short Syntax Comprehension (n = 18).
For Apraxia of Speech rating and Dysarthria rating, higher scores indicate greater impairment. For all other scores, lower scores indicate greater impairment.

4.3.3 Structural MRI acquisition

Participants who underwent an MEG scan also took part in a structural brain imaging protocol in a 3T Siemens MRI scanner at the Neuroscape MRI Lab at UCSF. T1-weighted structural MRI images were acquired using a T1-MPRAGE sequence (repetition time = 2300ms, echo time =

2.98ms, inversion time = 900ms, slice thickness = 1mm, field of view = 240mm x 256mm). These structural MRIs were used to generate head models for source-space reconstruction of MEG sensor data and to calculate grey matter volume estimates for statistical correction of grey-matter atrophy in the MEG results. Individual structural MRIs were spatially normalised to a standard Montreal Neurological Institute (MNI) template using SPM8 (https://www.fil.ion.ucl.ac.uk/spm/software/spm8/).

4.3.4 MEG Imaging

A 275-channel whole-head MEG scanner (CTF Inc., Coquitlam, BC, Canada) was used for MEG-imaging. Participants were scanned in supine position. Signals were acquired at a sampling rate of 1200Hz and three fiducial coils (one at the nasion and two at preauricular points on both sides) were used to record head position relative to the sensor array. These fiducial locations were co-registered to participants' structural MRI images and head shapes were generated for each individual.

4.3.5 Pitch perturbation experiment

The pitch perturbation experiment consisted of 120 trials. In every trial, participants were prompted to vocalise the vowel /ɑ/ for as long as they saw a green dot on the screen (duration ~ 2.4s) (see Fig. 4.1A). Participants' vocal output was recorded using an optical microphone (Phone-Or Ltd., Or-Yehuda, Israel) and they could simultaneously hear themselves through insert earphones (ER-3A, Etymotic Research, Inc., Elk Grove Village, IL). During every trial, between 200-500ms after voice onset was detected, the pitch of the participants' auditory feedback was perturbed either upwards or downwards by 100 cents ($1/12^{th}$ of an octave) for a

duration of 400ms using a digital signal processor called Feedback Utility for Speech Processing (FUSP) also used in previous studies (Katseff *et al.*, 2012; Ranasinghe *et al.*, 2017a; Schuerman *et al.*, 2017; Naunheim *et al.*, 2018). The direction of shift was randomly determined for each trial with an equal number of trials with upward and downward shifts.

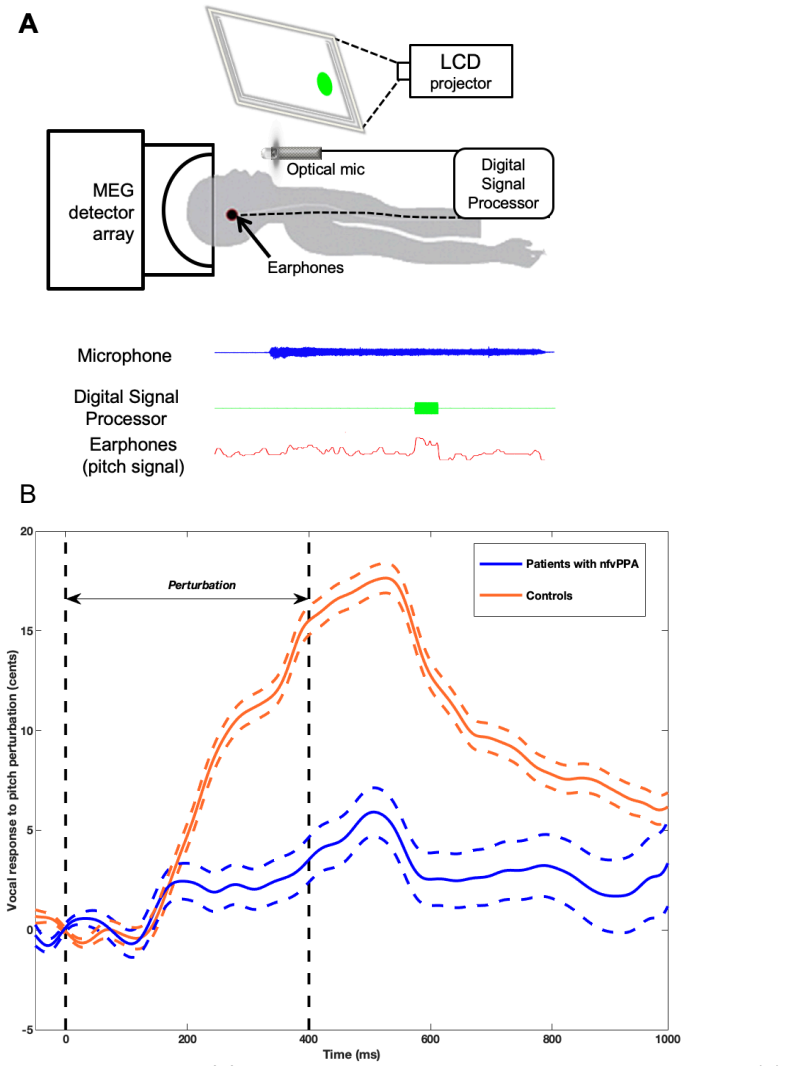


Fig. 4.1 Task description and behavioural results
(A) *During every trial, participants started to vocalise the vowel /a/ into a microphone upon seeing a green dot on the display. Participants also wore earphones through which they could hear themselves throughout the trial. Between 200-500ms after the start of vocalisation, a digital signal processing unit shifted the pitch of the participants' speech signal either up or down for 400ms and sent this shifted signal to the participants' earphones.* ***(B)*** *On average, patients with*

nfvPPA (n = 18) have a smaller compensation response to pitch perturbation as compared to controls (n = 17). Dashed lines represent standard error of the mean.

4.3.6 Data Analysis

4.3.6.1 Audio Data Processing

Participants' speech and the pitch-altered feedback were both recorded at 11,025 Hz. For every trial, the time course of the pitch track was determined using an autocorrelation-based pitch tracking method ((Parsons, 1987). Pitch tracks were aligned from 200ms prior to perturbation onset to 1000ms after perturbation onset. All trials with pitch tracking errors and incomplete utterances were marked bad and excluded from further analysis. For all good trials, pitch values were converted from Hertz to cents using the following formula:

$$Pitch_{cents}(t) = 1200\ log_2(\frac{Pitch_{Hz}(t)}{Pitch_{ref}})$$

where $Pitch_{Hz}(t)$ is the pitch value in Hertz at timepoint t and $Pitch_{ref}$ is the reference pitch calculated as the mean pitch over a window spanning from 50ms prior to perturbation onset to 50ms after perturbation onset. Participants' responses to pitch perturbation were expressed as deviations from their baseline pitch track. For each participant, responses to both upward and downward perturbations were calculated and pooled together. Downward responses to upward perturbations were flipped and combined with upward responses to downward perturbations, thus making all compensatory responses positive.

For statistical analysis, the pitch track for each individual trial was divided into bins of 50ms each and pitch values were averaged within these bins. Group differences in compensatory pitch responses between patients and controls were calculated for each of these bins using trial means from both groups and running a one-way analysis of variance (anova1 function, MATLAB,

MathWorks, Natick, MA). To control for Type I error, Bonferroni thresholds were applied to p-values for $\alpha = 0.05$.

4.3.6.2 MEG Data Processing

We performed third-order gradient noise and direct current (DC) offset correction on the MEG sensor data. Pitch perturbation onset was marked for every trial and subsequent neural analyses were locked to this perturbation onset marker. All trials with noisy signal due to head movement, dental artefact, eye blinks or saccades were excluded from the analysis. We looked at neural responses in the theta (4-7Hz), alpha (8-12Hz), beta (13-30Hz), low gamma (30-55Hz) and high gamma (65-150Hz) frequency bands but we found cohort differences only in alpha and beta bands which are known to contain robust electrophysiological signatures of sensorimotor integration (Hari, 2006; Jenson *et al.*, 2014). Hence, we focus on these two bands in the results and the discussion.

Source localisation for induced alpha-band and beta-band activity in each participant was performed using time-frequency-optimised adaptive spatial filtering (8mm lead field) in the Neurodynamic Utility Toolbox for MEG (NUTMEG: http://nutmeg.berkeley.edu) (Dalal *et al.*, 2008; Hinkley *et al.*, 2020). Voxel-by-voxel estimates of neural activity were generated using a linear combination of a spatial weight matrix and a sensor data matrix. Active windows were defined as the time period after perturbation onset. Pseudo-F statistics were computed by comparing the active window to a control window prior to perturbation onset. Using the NUTMEG toolbox (Dalal *et al.*, 2011), both within-group and between-group statistical analyses were performed using statistical non-parametric mapping methods. For within-group contrasts, we used

a 5% False Discovery Rate (FDR) to correct for multiple comparisons across time and space and corrected p-value thresholds were calculated for $\alpha = 0.01$. To minimise the possibility of observing spurious effects due to noise, cluster correction was also performed to exclude clusters with less than 40 congruent voxels.

For across-group contrasts, cortical grey matter atrophy in patients was taken into account. Grey matter (GM) maps were generated by tissue-type segmentation of the T1-weighted MRIs of patients and controls using the DARTEL pipeline (Ashburner, 2007). These GM maps were then spatially normalised to a custom group template (n = 100) in MNI space using the same dimensions as the MEG data (79x95x68 matrix). With the voxel-by-voxel MEG activity and the GM maps, aligned, voxel-wise group effects corrected for atrophy were estimated using an analysis of covariance (ANCOVA) model. In this model, voxel-wise change in oscillatory power (in dB) was the predictor variable and GM intensity at the same voxel was a covariate. We used the same FDR correction parameters as in the within-group analyses but the threshold used for cluster correction was 20 voxels.

4.3.7 Correlation of average peak neural activity with speech motor impairment

To explore possible relationships between neural responses to pitch perturbation and speech motor impairment in patients, we first calculated a Speech Motor Composite Score (SMCS), for the 16 patients having both neural data and neuropsychological assessment scores, using the following formula:

$$SMCS = \frac{[z(WAB\ Fluency\ score) + (-AoS\ rating) + (-Dysarthria\ rating)]}{3}$$

where z stands for z-score normalisation with respect to the normative score, WAB = Western Aphasia Battery, AoS = Apraxia of Speech. The values for AoS rating and Dysarthria rating were negated because we wanted lower scores to depict greater impairment and in both these ratings higher values indicate greater impairment. AoS and Dysarthria ratings are already normalised and hence a z-score was not calculated for them.

For the voxel and timepoints showing peak significant differences between patients and controls, the average alpha power and the average beta power were calculated for each of the 16 patients.

Then, for each frequency band (alpha and beta), a generalised linear model (GLM) was fit in SAS 9.4 (SAS Institute Inc., Cary, NC) to investigate whether the average peak power in the frequency band predicted speech motor impairment.

4.4 Results

4.4.1 Participant characteristics

Patients with nfvPPA showed mild cognitive impairment (median value for MMSE = 28, see Table 4.1). The control and patient cohorts did not differ in terms of age, sex, handedness and race. However, controls and patients did differ when it came to average years of education (median = 18 for controls and 16 for patients). Patients' average neuropsychological test performance reflected the profile of a cohort that showed motor speech difficulties but was not severely impaired.

4.4.2 Behavioural response to pitch perturbation

Figure 4.1B shows both groups' vocal response to pitch perturbation. Patients with nfvPPA and controls started responding to perturbation at ~150 ms and followed a similar trajectory until ~200 ms. From 200ms onwards, control participants continued to respond along the same incline until they reached a peak at 526ms (peak response =17.64 ± 0.75 cents). Patients with nfvPPA, in sharp contrast, showed a different trajectory after ~200ms where they followed a less-steep slope and reached a smaller peak at 505ms (peak response = 5.91 ±1.22 cents). From 200ms until 950ms after perturbation onset, patients with nfvPPA showed a significantly reduced behavioural response compared to controls (Figure 4.1B).

To evaluate whether the smaller responses in patients was due to a limited vocal motor range, we quantified vocal motor output capacity as pitch variability in patients and controls in a 200ms pre-perturbation baseline window as done in previous studies (Naunheim *et al.*, 2018; Houde *et al.*, 2019). Variability in patients with nfvPPA did not differ significantly from that in controls, for both within-trial and across-trial analyses (Supp. Fig. C.1).

4.4.3 Neural response to pitch perturbation

Next, we examined neural activity patterns during the pitch perturbation response. As the behavioural response started at ~150 ms (Figure 4.1B), the neural activity before this point in time can be presumed to be involved in feedback error detection and preparation towards motor correction. Neural activity after 150ms until the peak behavioural response is reached (i.e., around 500ms) may reflect the sensorimotor integration processing of auditory feedback. Any neural activity after peak compensation would be indicative of the return to baseline vocalisation.

We focused on the neural responses within alpha (8-12 Hz) and beta (13-30 Hz) oscillatory bands, in nfvPPA and controls, because we found significant cohort differences in these two bands.

4.4.3.1 Patients with nfvPPA have impaired right-hemispheric alpha-band activity during speech motor integration processing

Controls showed increased alpha-band activity in bilateral posterior cortices (Figure 4.2A) throughout the analytical time window of 0-900ms after perturbation onset. This response was larger in the right hemisphere than the left. The posterior superior parietal cortex was involved in both hemispheres while in the right hemisphere, the occipital cortex, temporo-parietal junction and the inferior temporal cortex were also involved. Controls also showed reduced alpha-band power in the left anterior temporal lobe from 250-550ms.

Patients with nfvPPA also showed increased alpha-band activity in the posterior regions of the brain bilaterally and throughout the analytical window (Figure 4.2A). However, unlike the controls, patients did not show right hemispheric dominance. Patients showed bilateral posterior superior hyperactivity with no additional regional involvement in the right hemisphere. A direct contrast between neural activity in controls and patients emphasised this group distinction and identified significantly reduced alpha-band activity in the right temporo-parieto-occipital junction in patients with nfvPPA compared to controls. This difference persisted from 250ms to 750ms after perturbation onset, a window corresponding to neural processing of speech-motor-integration.

Patients with nfvPPA also showed reduced activity in the anterior temporal cortex and frontal cortex (350-550ms) in the left hemisphere and the inferior frontal cortex (50-550ms) in the right hemisphere. These patterns however were not statistically different when compared to controls.

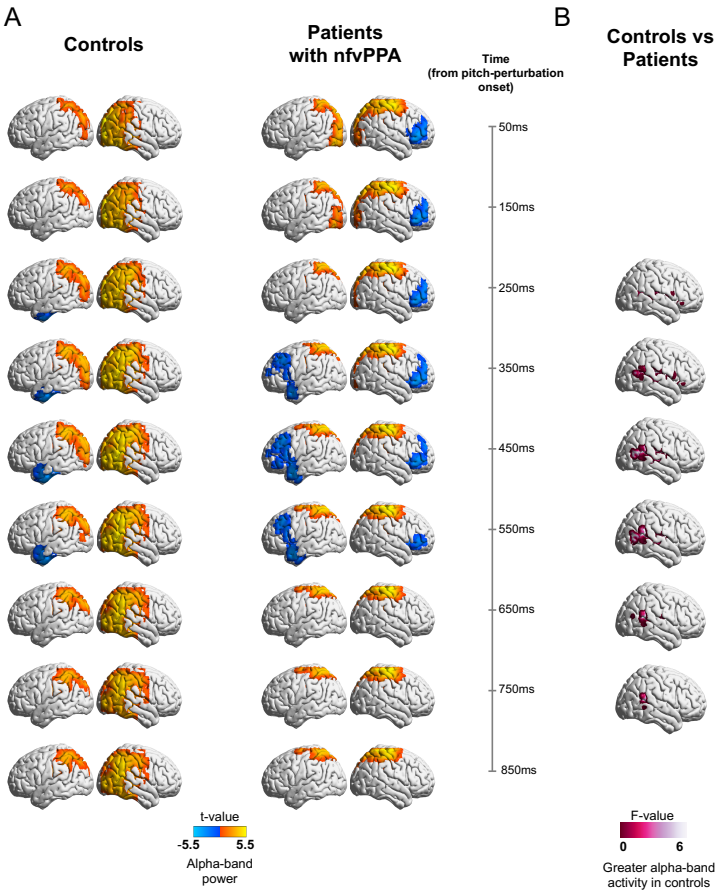


Fig. 4.2 Neural activity during pitch feedback perturbation in alpha band (8 - 12 Hz) ***(A)*** *Neural activity in controls and patients with nfvPPA in the alpha band (8-12 Hz) locked to pitch perturbation onset.* ***(B)*** *Patients with nfvPPA (n = 17) show significantly lesser alpha-band activity in the posterior right temporal lobe and the right temporo-parieto-occipital junction as compared to controls (n = 14). Neural activity shown was corrected for grey-matter atrophy and was False Discovery Rate (FDR)-corrected for multiple frequency bands and time points. Cluster correction was also performed at a threshold of 20 voxels and $p < 0.01$. Timepoint 0 corresponds to perturbation onset.*

4.4.3.2 Patients with nfvPPA have increased left dorsal sensorimotor beta-band activity during feedback error detection and corrective motoric preparation

Control participants showed increased beta-band activity (Figure 4.3A) in both hemispheres involving the posterior parietal regions during most of the analytical window (0-700ms post perturbation onset). This increased activity in the parietal cortex in both hemispheres seemingly spread along the primary somatosensory and primary motor cortices from 150 to 550ms. Apart from this robust and consistent activity over the posterior parietal cortices, there was transient early involvement of left frontal regions (50-150 ms), left occipital regions (50-150 ms), and right temporal cortex (350-450 ms). Patients with nfvPPA also showed consistently increased beta-band activity over the parietal cortex however, with a more dorsally-dominant spatial distribution compared to controls (Figure 4.3A). The increased beta activity in patients was more consistent over time in the right hemisphere, which persisted throughout the analytical window, while it only persisted from 50-350 in the left hemisphere. In a direct comparison between controls and nfvPPA patients, we found that beta-band activity in the left dorsal parietal region in patients was significantly higher than controls (Figure 4.3B).

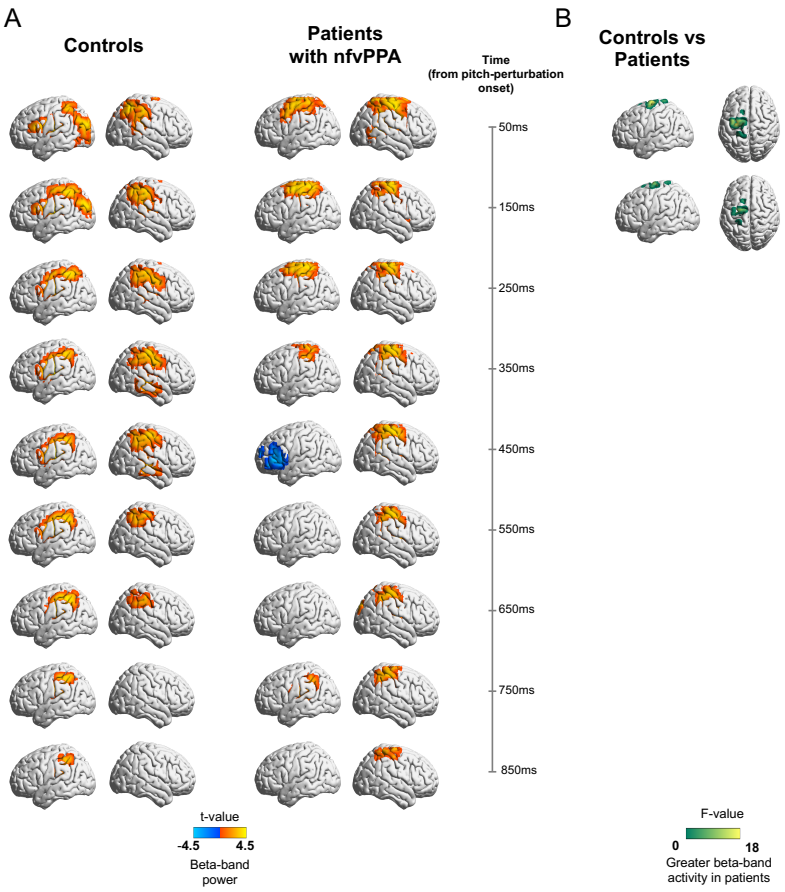


Fig. 4.3 Neural activity during pitch feedback perturbation in beta band (13 - 30 Hz) ***(A)*** *Neural activity in controls and patients with nfvPPA in the beta band (13-30 Hz) locked to pitch perturbation onset* ***(B)*** *Patients with nfvPPA (n = 17) show greater beta-band activity in the dorsal sensorimotor and premotor cortices as compared to controls (n = 14). Neural activity shown was corrected for grey-matter atrophy and was False Discovery Rate (FDR)-corrected for multiple frequency bands and time points. Cluster correction was also performed at a threshold of 20 voxels and p < 0.01. Timepoint 0 corresponds to perturbation onset.*

4.4.4 Neural correlates of speech motor impairment during compensation for pitch perturbation in patients

Next, we examined the associations between the abnormal neural indices of speech motor control processing identified in the above analyses and the clinical measures of speech motor function in patients with nfvPPA. To this end we used a generalised linear model (GLM) and examined the

associations between the composite speech motor score derived from clinical testing and the regional neural activity within regions-of-interest (ROI) as defined by the group contrast analyses above. The GLM between the composite speech motor score and the average peak alpha power within the right temporo-parietal junction ROI, across the timepoints where patients showed significantly reduced alpha-band activity showed a significant positive association, indicating that patients with nfvPPA with lower surplus alpha activity are poor performers in speech motor functional tests (Figure 4.4A; β = 3.41, F = 8.31, p = 0.0128). A similar GLM for the associations between speech motor composite scores and the average peak beta power in the left dorsal parietal ROI, across the timepoints where patients showed significantly reduced alpha-band activity did not reveal a significant relationship (Figure 4.4B; β = -1.75, F = 1.72, p = 0.2123).

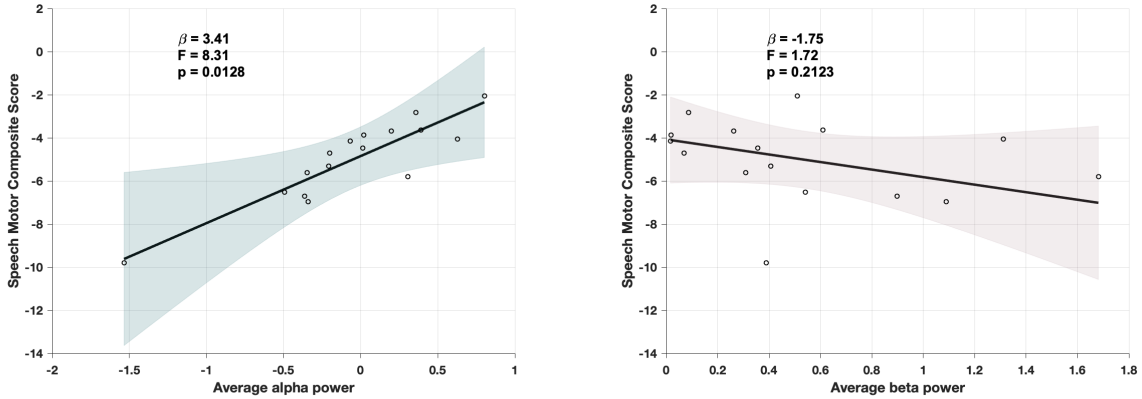


Fig. 4.4 Alpha-band power predicts speech motor impairment
In a generalised linear model, patients' speech motor composite score was significantly positively correlated with the average of peak alpha power (from 250-750ms after pitch perturbation onset) but there was no relationship with the average of peak beta power (50-150ms after pitch perturbation onset). Thus implying that the lower the induced alpha power changes due to pitch perturbation in a patient with nfvPPA, the greater is their speech motor impairment.

4.5 Discussion

We investigated the cortical dynamics of the speech production network in patients with nfvPPA using the widely-studied pitch perturbation reflex as a tool to examine speech motor control. Patients showed significantly reduced vocal behaviour compensation for altered feedback as compared to controls (only ~33% of the peak compensation seen in controls). Neural activity from MEG imaging revealed that patients with nfvPPA had significantly reduced alpha-band activity in the right posterior temporal lobe and the right temporo-parieto-occipital junction when responding to pitch perturbation. Patients showed significantly increased activity in the beta band in the dorsal sensorimotor and premotor cortices during early response to pitch perturbation. Reduced alpha-band activity in patients predicted speech motor impairment in patients whereas increased beta-band activity did not correlate with speech motor impairment in patients. Together, the results point towards significant impairments in the sensorimotor integration process during speech production in patients with nfvPPA. These sensorimotor impairments may be playing a contributory role in the speech motor deficits and other characteristic loss of speech abilities associated with nfvPPA.

4.5.1 Reduced vocal response to pitch perturbation

Patients with the non-fluent variant of Primary Progressive Aphasia primarily exhibit effortful, non-fluent speech (Grossman, 2012) and motor speech deficits (Wilson *et al.*, 2010) like apraxia of speech (Rohrer *et al.*, 2010) and dysarthria (Ogar *et al.*, 2007). These deficits along with neuronal atrophy (Gunawardena *et al.*, 2010) and impaired structural connectivity of brain regions involved in speech production (Galantucci *et al.*, 2011; Mandelli *et al.*, 2014; Mandelli *et al.*, 2018) suggest that the functional recruitment of the speech motor control network will be

impacted in nfvPPA. Models of speech motor control posit that speech production involves continuous integration of sensory feedback to adjust for sensory feedback errors and regulate motor behaviour (Houde and Nagarajan, 2011; Tourville and Guenther, 2011; Guenther, 2016; Parrell and Houde, 2019; Parrell *et al.*, 2019). The functional efficacy of the speech motor control network can be tested by altering sensory feedback during speech production and evaluating the network's ability to compensate for this alteration. Our findings demonstrated that patients with nfvPPA have a significantly reduced vocal response to pitch feedback alteration, compared to controls. This is in sharp contrast to the increased vocal response to feedback alteration shown in patients with Alzheimer's disease (Ranasinghe *et al.*, 2017a; Ranasinghe *et al.*, 2019), Parkinson's disease (Liu *et al.*, 2012; Chen *et al.*, 2013; Huang *et al.*, 2016; Mollaei *et al.*, 2016) and cerebellar degeneration (Parrell *et al.*, 2017; Houde *et al.*, 2019), when compared to healthy controls. The reduced vocal response in nfvPPA patients cannot be attributed to a limited vocal output range because patients' within-trial and across-trials baseline pitch variability (Supp. Fig. C.1) is large enough to achieve compensation. What then might be responsible for a reduced compensatory response in nfvPPA patients? It is important to look at what triggers a compensatory response to altered auditory feedback. When auditory feedback is externally perturbed, it causes a mismatch between the predicted feedback and the actual feedback, thus generating a feedback error. Compensating for altered feedback reduces this feedback error. Any difficulty in the detection or auditory processing of this feedback error may explain a weak compensatory response to external feedback perturbation. We looked for clues suggesting this possible explanation in the neuroimaging results.

4.5.2 Reduced alpha-band neural activity in patients' posterior temporal and temporo-parietal regions and prediction of speech motor impairment

Previous studies have shown that posterior temporal regions are highly sensitive to pitch perturbation (McGuire *et al.*, 1996; Zarate and Zatorre, 2005; Fu *et al.*, 2006). It is postulated that these regions are involved in auditory error detection (Tourville *et al.*, 2008; Chang *et al.*, 2013; Houde and Chang, 2015). Moreover, activity in the alpha band is thought to be associated with suppression of irrelevant sensory information (Strauß *et al.*, 2014) suggesting alpha-modulated noise suppression in speech processing. Higher alpha-band activity indicates successful selective inhibition while less alpha-band activity would be indicative of an inability to tune out irrelevant information. Patients show significantly lesser alpha-band activity than controls in the posterior right temporal lobe and the right temporo-parieto-occipital junction from 250ms (when patient's vocal response deviates from that of controls) to 750ms, suggesting that impaired suppression of information that is irrelevant to the task at hand may be a potential mechanism contributing to their reduced vocal response to pitch perturbation. Given that the impaired reduction in alpha activity started closer to the perturbation onset and persisted until after the perturbation offset, it is likely that this abnormal neural activity hinders the processing of the auditory feedback error. If patients with nfvPPA are unable to process the altered pitch feedback they would subsequently be unable to integrate the error into their corrective motor output. Reduced compensatory response to pitch perturbation in patients with nfvPPA may thus be an indicator of impaired sensory feedback processing and provide a useful tool to quantify such impairment.

We also found that reduced alpha-band power over the right posterior temporal region is significantly correlated with the clinical measures of speech motor impairment in patients with nfvPPA (Figure 4.4). Patients who showed more hypoactivity in the cluster in the posterior right temporal lobe showed poorer scores in a speech motor composite measure as determined by neuropsychological tests. Although the mechanisms connecting motor speech production abnormalities as captured in neuropsychological testing and sensorimotor integration of auditory feedback response via posterior temporal region are yet to be determined, our results suggest that sensorimotor integration abnormalities, in addition to direct motor production deficits, also contribute to the behavioural motor speech impairments found in patients with nfvPPA. Future studies with longitudinal assessments will determine the temporal evolution and interdependencies of these sensorimotor and motor components of the speech dysfunction in nfvPPA and how they progress in the disease course.

4.5.3 Increased beta-band neural activity in patients' left dorsal sensorimotor and premotor regions

At timepoints immediately succeeding perturbation onset (50ms and 150ms), we observed an increase in beta-band neural activity in the left dorsal sensorimotor cortex and the left dorsal premotor cortex which are together involved in motor planning and motor control. It can be seen that patients' vocal response at these timepoints aligns temporally with the behavioural vocal response in controls. Thus, we believe that this increase in activity in patients did not contribute towards the reduced vocal response in patients. Patients with nfvPPA exhibit cortical atrophy in the ventral regions of the motor cortex or the laryngeal motor cortex (see Supp. Fig. C.2) and the dorsal hyperactivity may be the manifestation of an increased effort by patients in motor

planning to respond to an auditory feedback error, perhaps a compensatory strategy to overcome the atrophy in regions that are normally active during laryngeal control. Indeed it has recently been shown that the conventional idea of neuronal organisation as per the 'motor homunculus' may not be rigid and that neurons in the dorsal motor cortex may be recruited for speech production (Stavisky *et al.*, 2019). Moreover, the beta-band hyperactivity did not predict speech motor impairment in patients (Figure 4.4) which further suggests that it did not contribute to abnormal sensorimotor integration. Differences in beta-band activity may be a signature of neural plasticity due to significant damage to the cortex and white matter tracts observed in nfvPPA.

As stated before, from the perspective of computational models of speech motor control, neural activity before the onset of a behavioural response (~150ms after pitch feedback perturbation onset) is thought to be involved in the detection and registration of a prediction error and preparation of a motoric response to this error. From 50-150ms after pitch feedback perturbation onset, beta-band hyperactivity is observed in patients with nfvPPA in brain regions involved in generating a corrective motor in response to a prediction error (sensorimotor cortex and premotor cortex) but not in regions associated with the registration of a prediction error (higher-order association cortices). Neural activity after the onset of a behavioural response (after 150ms post-perturbation onset) is thought to be involved in online sensorimotor integration processing of auditory feedback. Therefore, alpha-band hypoactivity in patients with nfvPPA (from 250-750ms after pitch feedback perturbation onset) in brain regions involved in continuous monitoring of auditory prediction errors points to a disruption in sensorimotor integration.

4.6 Conclusion

Magnetoencephalographic neuroimaging during speech production provides an insight into disrupted sensorimotor integration in patients with nfvPPA. This knowledge will help further elucidate the mechanisms behind speech motor control and how they are affected by pathophysiological processes of neurodegeneration in primary progressive aphasia variants.

4.7 References

Ashburner J. A fast diffeomorphic image registration algorithm. NeuroImage 2007; 38(1): 95-113.

Burnett TA, Freedland MB, Larson CR, Hain TC. Voice F0 responses to manipulations in pitch feedback. The Journal of the Acoustical Society of America 1998; 103(6): 3153-61.

Chang EF, Niziolek CA, Knight RT, Nagarajan SS, Houde JF. Human cortical sensorimotor network underlying feedback control of vocal pitch. Proceedings of the National Academy of Sciences of the United States of America 2013; 110(7): 2653-8.

Chen X, Zhu X, Wang EQ, Chen L, Li W, Chen Z, et al. Sensorimotor control of vocal pitch production in Parkinson's disease. Brain research 2013; 1527: 99-107.

Dalal SS, Guggisberg AG, Edwards E, Sekihara K, Findlay AM, Canolty RT, et al. Five-dimensional neuroimaging: localization of the time–frequency dynamics of cortical activity. NeuroImage 2008; 40(4): 1686-700.

Dalal SS, Zumer JM, Guggisberg AG, Trumpis M, Wong DD, Sekihara K, et al. MEG/EEG source reconstruction, statistical evaluation, and visualization with NUTMEG. Computational intelligence and neuroscience 2011; 2011: 758973.

Dunn LM, Dunn DM. Peabody Picture Vocabulary Test, 4 edn. NCS Pearson. Inc; 2007.

Fu CHY, Vythelingum GN, Brammer MJ, Williams SCR, Amaro Jr E, Andrew CM, et al. An fMRI study of verbal self-monitoring: neural correlates of auditory verbal feedback. Cerebral cortex 2006; 16(7): 969-77.

Galantucci S, Tartaglia MC, Wilson SM, Henry ML, Filippi M, Agosta F, et al. White matter damage in primary progressive aphasias: a diffusion tensor tractography study. Brain : a journal of neurology 2011; 134(10): 3011-29.

Gorno-Tempini ML, Hillis AE, Weintraub S, Kertesz A, Mendez M, Cappa SF, et al. Classification of primary progressive aphasia and its variants. Neurology 2011; 76(11): 1006-14.

Grossman M. The non-fluent/agrammatic variant of primary progressive aphasia. The Lancet Neurology 2012; 11(6): 545-55.

Guenther FH. Neural control of speech: MIT Press; 2016.

Gunawardena D, Ash S, McMillan C, Avants B, Gee J, Grossman M. Why are patients with progressive nonfluent aphasia nonfluent? Neurology 2010; 75(7): 588-94.

Hari R. Action–perception connection and the cortical mu rhythm. Progress in brain research 2006; 159: 253-60.

Hinkley LBN, Dale CL, Cai C, Zumer J, Dalal S, Findlay A, et al. NUTMEG: Open Source Software for M/EEG Source Reconstruction. Frontiers in Neuroscience 2020; 14: 710.

Houde JF, Chang EF. The cortical computations underlying feedback control in vocal production. Current opinion in neurobiology 2015; 33: 174-81.

Houde JF, Gill JS, Agnew Z, Kothare H, Hickok G, Parrell B, et al. Abnormally increased vocal responses to pitch feedback perturbations in patients with cerebellar degeneration. The Journal of the Acoustical Society of America 2019; 145(5): EL372.

Houde JF, Nagarajan SS. Speech production as state feedback control. Frontiers in human neuroscience 2011; 5: 82.

Huang X, Chen X, Yan N, Jones JA, Wang EQ, Chen L, et al. The impact of parkinson's disease on the cortical mechanisms that support auditory–motor integration for voice control. Human brain mapping 2016; 37(12): 4248-61.

Jenson D, Bowers AL, Harkrider AW, Thornton D, Cuellar M, Saltuklaroglu T. Temporal dynamics of sensorimotor integration in speech perception and production: independent component analysis of EEG data. Frontiers in psychology 2014; 5(656).

Katseff S, Houde J, Johnson K. Partial compensation for altered auditory feedback: a tradeoff with somatosensory feedback? Language and speech 2012; 55(Pt 2): 295-308.

Kertesz A. Western aphasia battery: Revised: Pearson; 2007.

Leyton CE, Britton AK, Hodges JR, Halliday GM, Kril JJ. Distinctive pathological mechanisms involved in primary progressive aphasias. Neurobiology of aging 2016; 38: 82-92.

Liu H, Wang EQ, Metman LV, Larson CR. Vocal responses to perturbations in voice auditory feedback in individuals with Parkinson's disease. PloS one 2012; 7(3): e33629.

Mack WJ, Freed DM, Williams BW, Henderson VW. Boston Naming Test: shortened versions for use in Alzheimer's disease. Journal of gerontology 1992; 47(3): P154-P8.

Mandelli ML, Caverzasi E, Binney RJ, Henry ML, Lobach I, Block N, et al. Frontal white matter tracts sustaining speech production in primary progressive aphasia. Journal of Neuroscience 2014; 34(29): 9754-67.

Mandelli ML, Welch AE, Vilaplana E, Watson C, Battistella G, Brown JA, et al. Altered topology of the functional speech production network in non-fluent/agrammatic variant of PPA. Cortex; a journal devoted to the study of the nervous system and behavior 2018; 108: 252-64.

McGuire PK, Silbersweig DA, Frith CD. Functional neuroanatomy of verbal self-monitoring. Brain : a journal of neurology 1996; 119(3): 907-17.

Mesulam MM. Primary progressive aphasia—a language-based dementia. New England Journal of Medicine 2003; 349(16): 1535-42.

Mollaei F, Shiller DM, Baum SR, Gracco VL. Sensorimotor control of vocal pitch and formant frequencies in Parkinson's disease. Brain research 2016; 1646: 269-77.

Morris JC. The Clinical Dementia Rating (CDR). Current version and scoring rules 1993; 43(11): 2412--a.

Naunheim ML, Yung KC, Schneider SL, Henderson-Sabes J, Kothare H, Mizuiri D, et al. Vocal motor control and central auditory impairments in unilateral vocal fold paralysis. The Laryngoscope 2018.

Ogar JM, Dronkers NF, Brambati SM, Miller BL, Gorno-Tempini ML. Progressive nonfluent aphasia and its characteristic motor speech deficits. Alzheimer Disease & Associated Disorders 2007; 21(4): S23-S30.

Parrell B, Agnew Z, Nagarajan S, Houde J, Ivry RB. Impaired Feedforward Control and Enhanced Feedback Control of Speech in Patients with Cerebellar Degeneration. The Journal of neuroscience : the official journal of the Society for Neuroscience 2017; 37(38): 9249-58.

Parrell B, Houde J. Modeling the role of sensory feedback in speech motor control and learning. Journal of Speech, Language, and Hearing Research 2019; 62(8S): 2963-85.

Parrell B, Ramanarayanan V, Nagarajan S, Houde J. The FACTS model of speech motor control: Fusing state estimation and task-based control. PLoS computational biology 2019; 15(9): e1007321.

Parsons TW. Voice and speech processing: McGraw-Hill College; 1987.

Ranasinghe KG, Gill JS, Kothare H, Beagle AJ, Mizuiri D, Honma SM, et al. Abnormal vocal behavior predicts executive and memory deficits in Alzheimer's disease. Neurobiology of aging 2017a; 52: 71-80.

Ranasinghe KG, Hinkley LB, Beagle AJ, Mizuiri D, Honma SM, Welch AE, et al. Distinct spatiotemporal patterns of neuronal functional connectivity in primary progressive aphasia variants. Brain : a journal of neurology 2017b; 140(10): 2737-51.

Ranasinghe KG, Kothare H, Kort N, Hinkley LB, Beagle AJ, Mizuiri D, et al. Neural correlates of abnormal auditory feedback processing during speech production in Alzheimer's disease. Sci Rep 2019; 9(1): 5686.

Rohrer JD, Rossor MN, Warren JD. Apraxia in progressive nonfluent aphasia. Journal of neurology 2010; 257(4): 569-74.

Schuerman WL, Nagarajan S, McQueen JM, Houde J. Sensorimotor adaptation affects perceptual compensation for coarticulation. The Journal of the Acoustical Society of America 2017; 141(4): 2693-704.

Spreen FO, Benton AL. Manual of instructions for the neurosensory center comprehensive examination for aphasia. Victoria: University of Victoria 1977.

Stavisky SD, Willett FR, Wilson GH, Murphy BA, Rezaii P, Avansino DT, et al. Neural ensemble dynamics in dorsal motor cortex during speech in people with paralysis. Elife 2019; 8: e46015.

Strauß A, Wöstmann M, Obleser J. Cortical alpha oscillations as a tool for auditory selective inhibition. Frontiers in human neuroscience 2014; 8: 350.

Tourville JA, Guenther FH. The DIVA model: A neural theory of speech acquisition and production. Language and cognitive processes 2011; 26(7): 952-81.

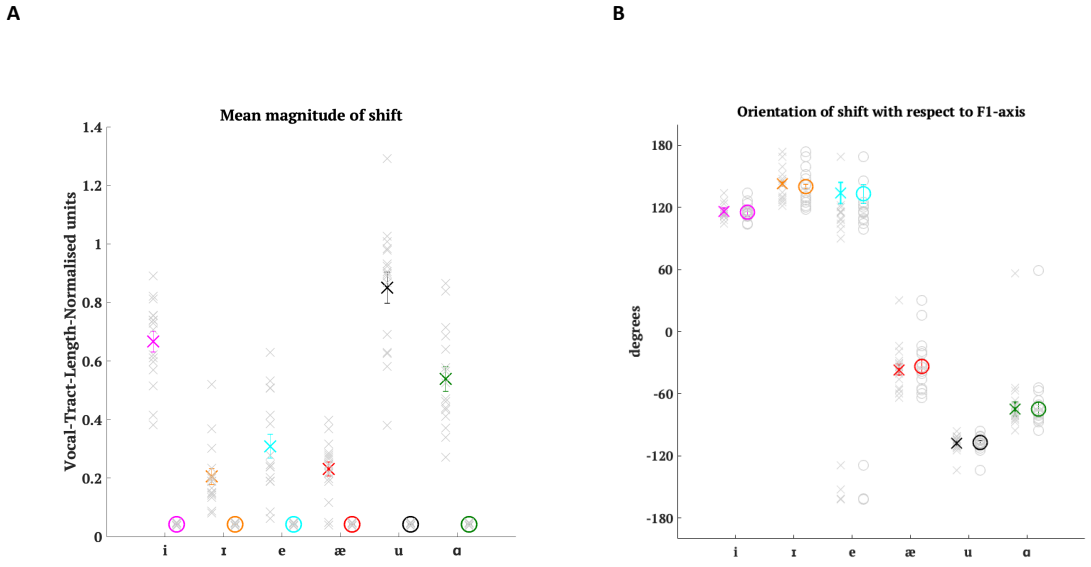
Tourville JA, Reilly KJ, Guenther FH. Neural mechanisms underlying auditory feedback control of speech. NeuroImage 2008; 39(3): 1429-43.

Wertz RT, LaPointe LL, Rosenbek JC. Apraxia of speech in adults: The disorder and its management: Grune & Stratton Orlando, FL; 1984.

Wilson SM, Henry ML, Besbris M, Ogar JM, Dronkers NF, Jarrold W, et al. Connected speech production in three variants of primary progressive aphasia. Brain : a journal of neurology 2010; 133(7): 2069-88.

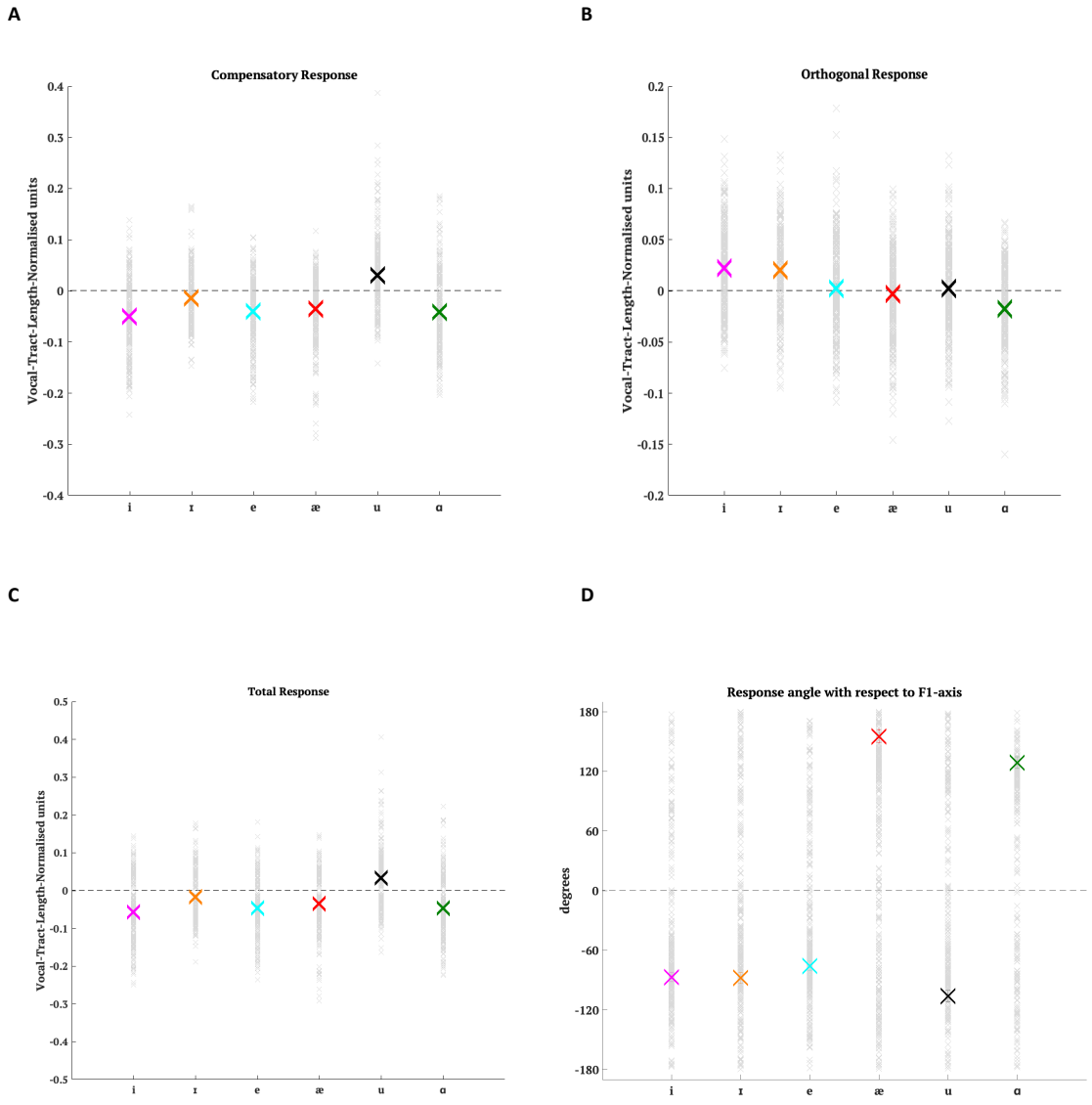
Zarate JM, Zatorre RJ. Neural substrates governing audiovocal integration for vocal pitch regulation in singing. Annals of the New York Academy of Sciences 2005; 1060(1): 404-8.

Appendix A: Supplementary Material to Chapter 2



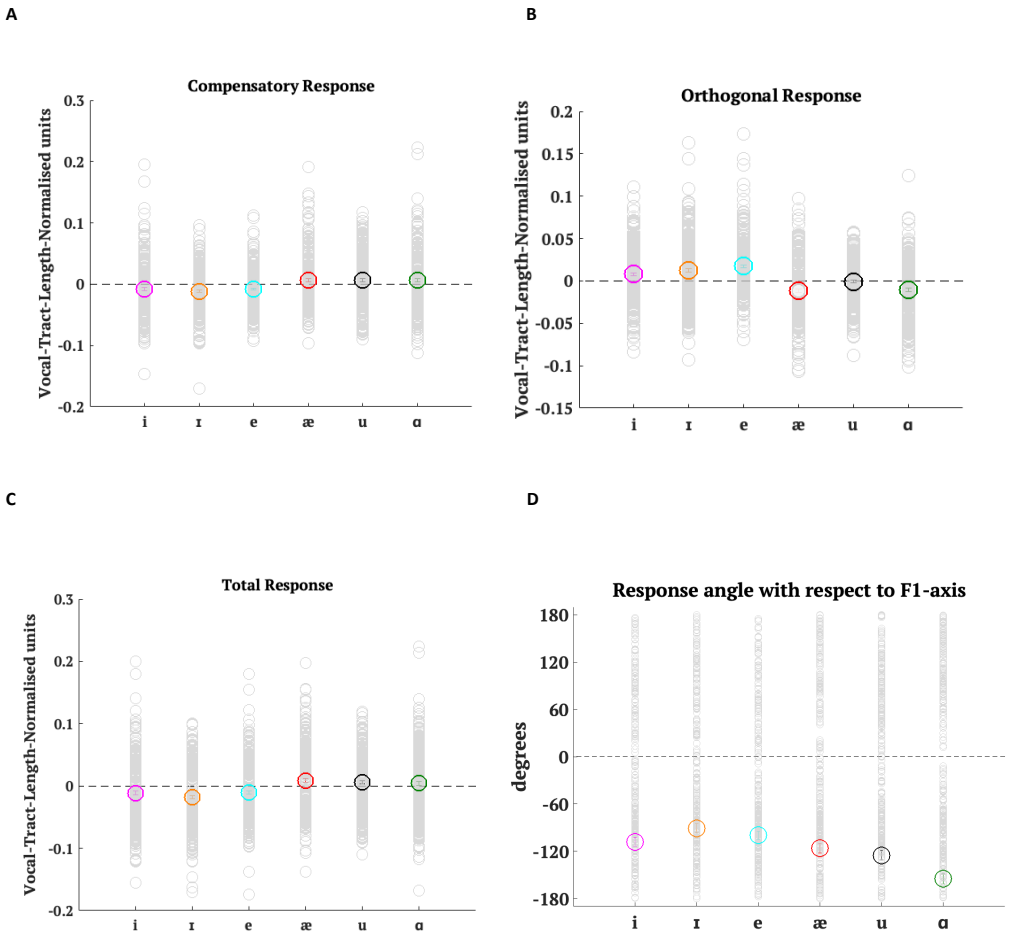
Supp. Fig. A.1 Individual data points for magnitudes of shift vectors and angles between the shift vectors and the F1-axis

A. *Mean magnitude of the shift vector (in vocal-tract-length-normalised units) for all 6 cases in Experiment 1 (crosses) and Experiment 2 (circles) averaged across participants along with data points for individual subjects (grey) in the background. Error bars depict ± 1 standard error of mean.* ***B.*** *Circular mean angle between the shift vector and the F1-axis (in degrees) for all 6 cases in Experiment 1 (crosses) and Experiment 2 (circles) averaged across participants along with data points for individual subjects (grey) in the background. Positive values indicate that the angle is measured clockwise from a line parallel to the F1-axis to the response vector. Negative values indicate that the angle is measured anticlockwise from a line parallel to the F1-axis to the response vector. Error bars depict ± 1 standard error of mean.*



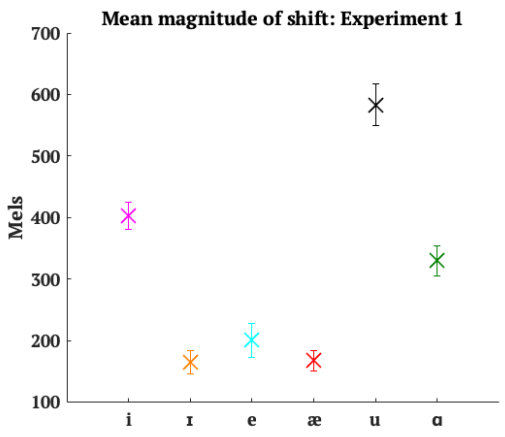
Supp. Fig. A.2 Individual data points for Experiment 1 results

Mean results for all 6 cases in Experiment 1 averaged across participants along with all data points (grey) in the background. Error bars depict ± 1 standard error of mean. ***A.*** *Mean compensatory response (in vocal-tract-length-normalised units).* ***B.*** *Mean orthogonal response magnitude (in vocal-tract-length-normalised units).* ***C.*** *Mean total response (in vocal-tract-length-normalised units).* ***D.*** *Circular mean angle between the response vector and the F1-axis (in degrees).*

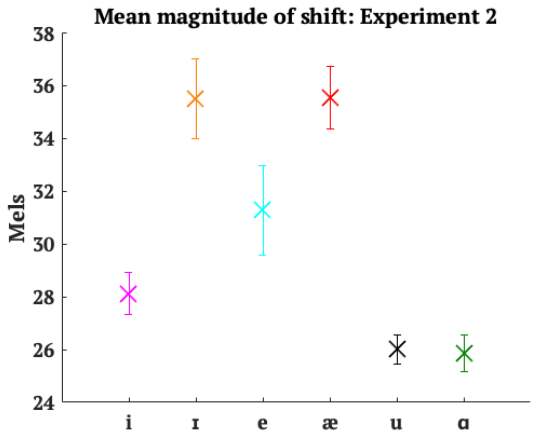


Supp. Fig. A.3 Individual data points for Experiment 2 results
Mean results for all 6 cases in Experiment 2 averaged across participants along with all data points (grey) in the background. ***A.*** *Mean compensatory response (in vocal-tract-length-normalised units).* ***B.*** *Mean orthogonal response magnitude (in vocal-tract-length-normalised units).* ***C.*** *Mean total response (in vocal-tract-length-normalised units).* ***D.*** *Circular mean angle between the response vector and the F1-axis (in degrees).*

A

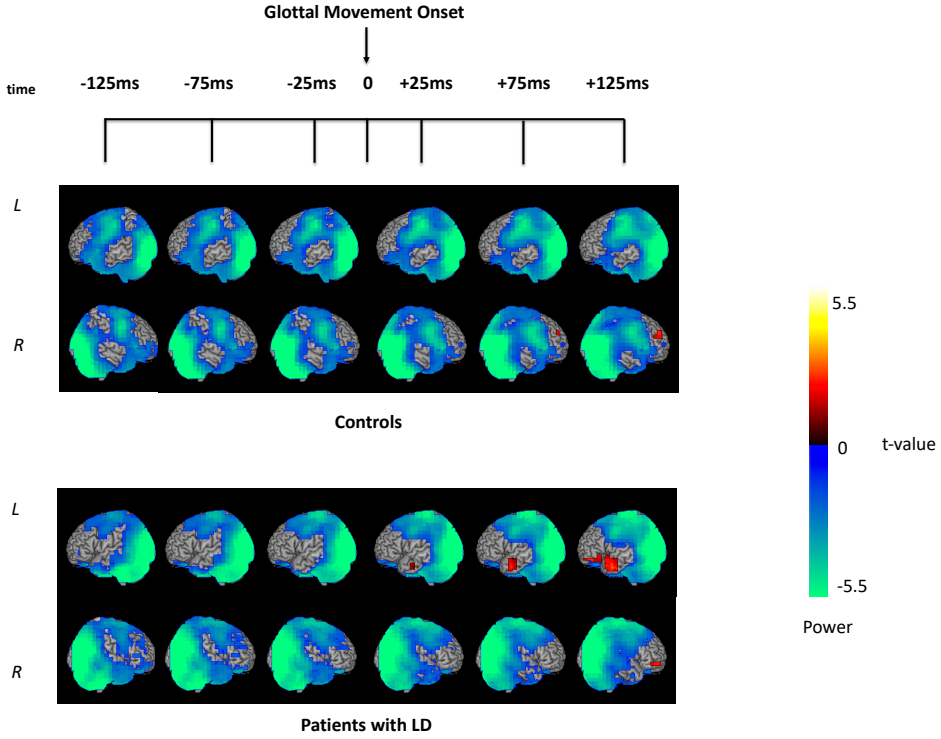


B

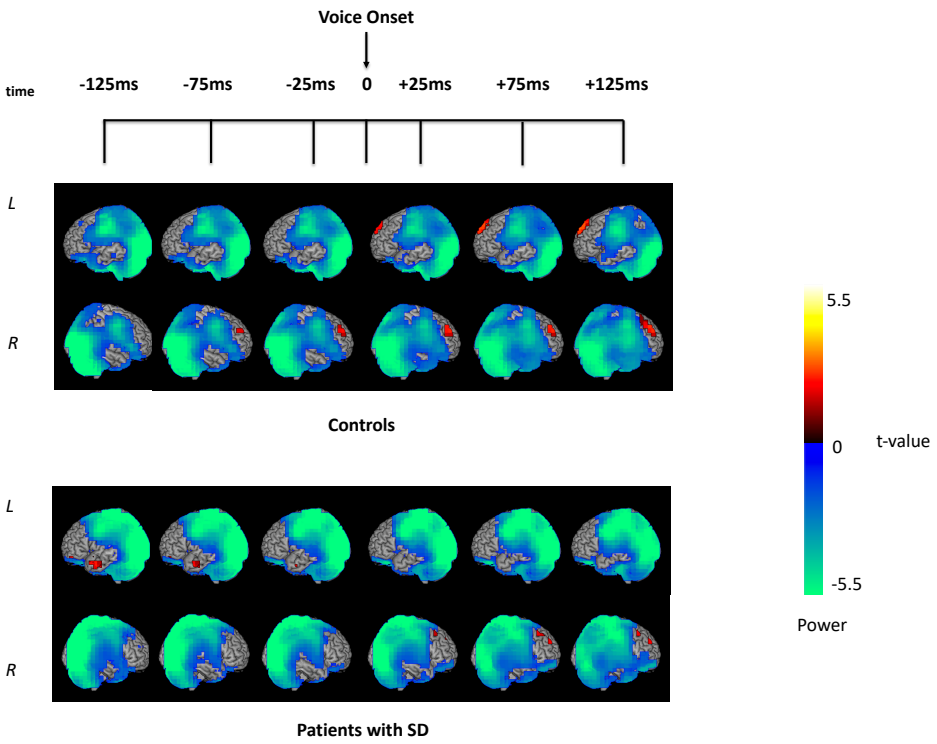


Supp. Fig. A.4 Mean magnitudes of shift vectors in mels
A. *Mean magnitude of the shift vector (in mels) for all 6 cases in Experiment 1.* ***B.*** *Mean magnitude of the shift vector (in mels) for all 6 cases in Experiment 2.*

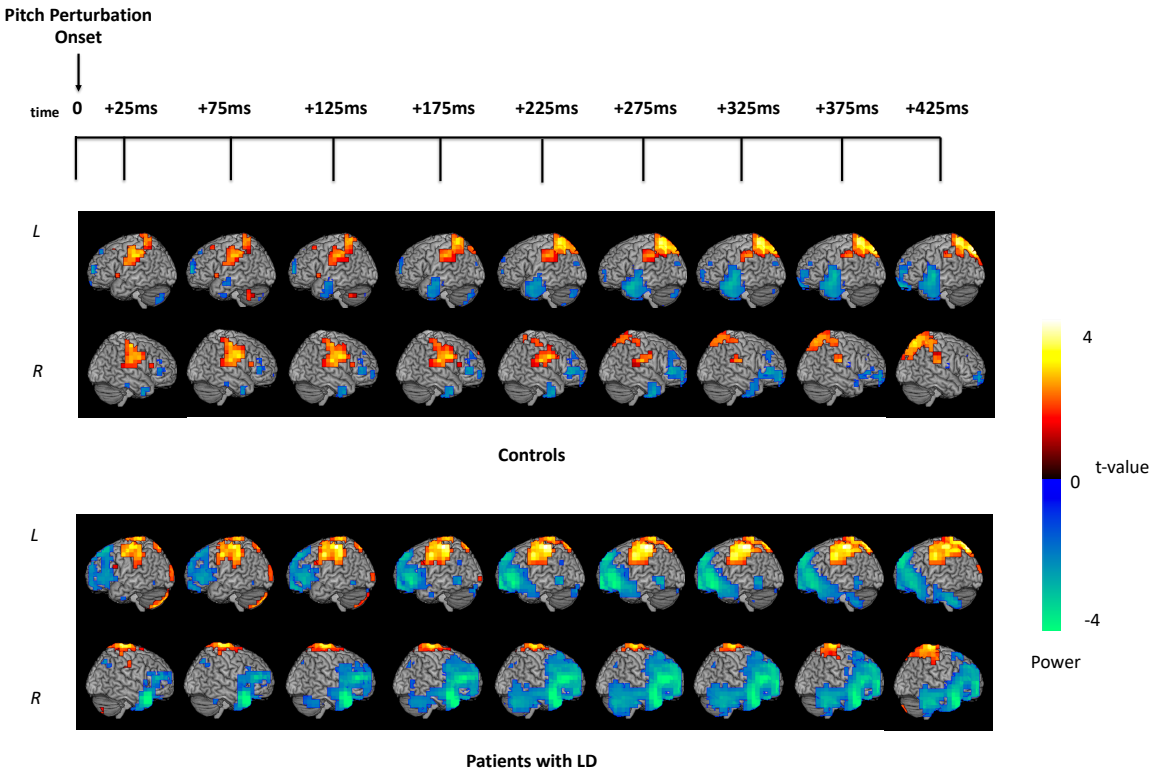
Appendix B: Supplementary Material to Chapter 3



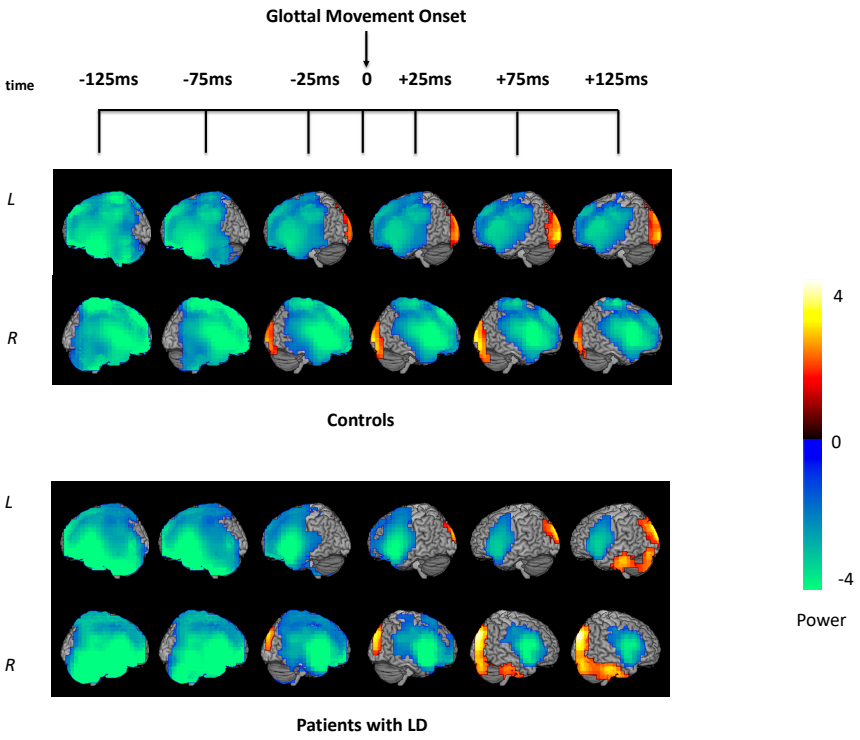
Supp. Fig. B.1 Neural activity in controls and patients with LD in the beta band (12 - 30 Hz) locked to glottal movement onset



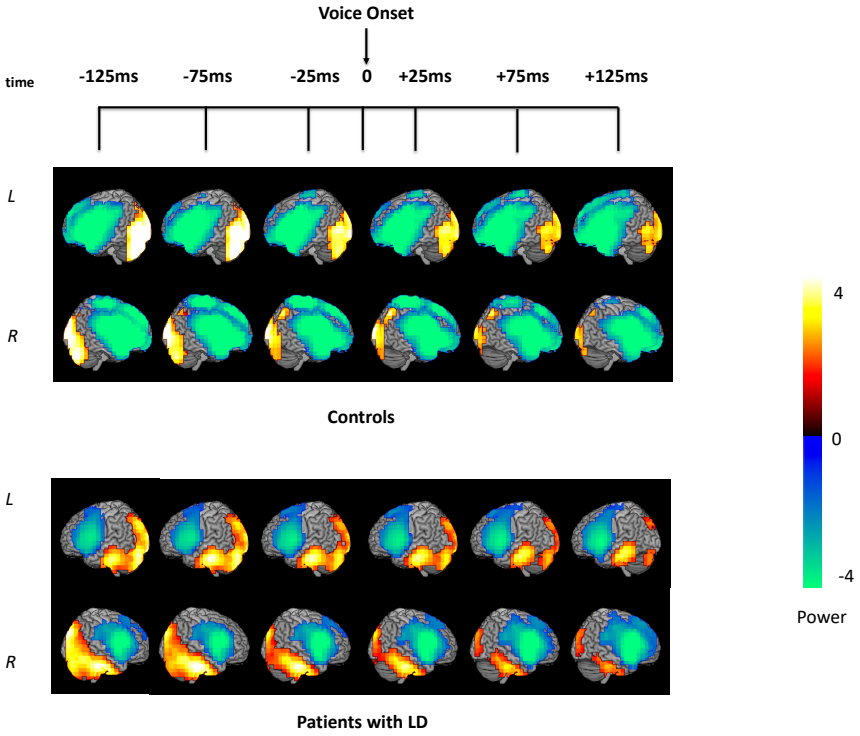
Supp. Fig. B.2 Neural activity in controls and patients with LD in the beta band (12 - 30 Hz) locked to voice onset



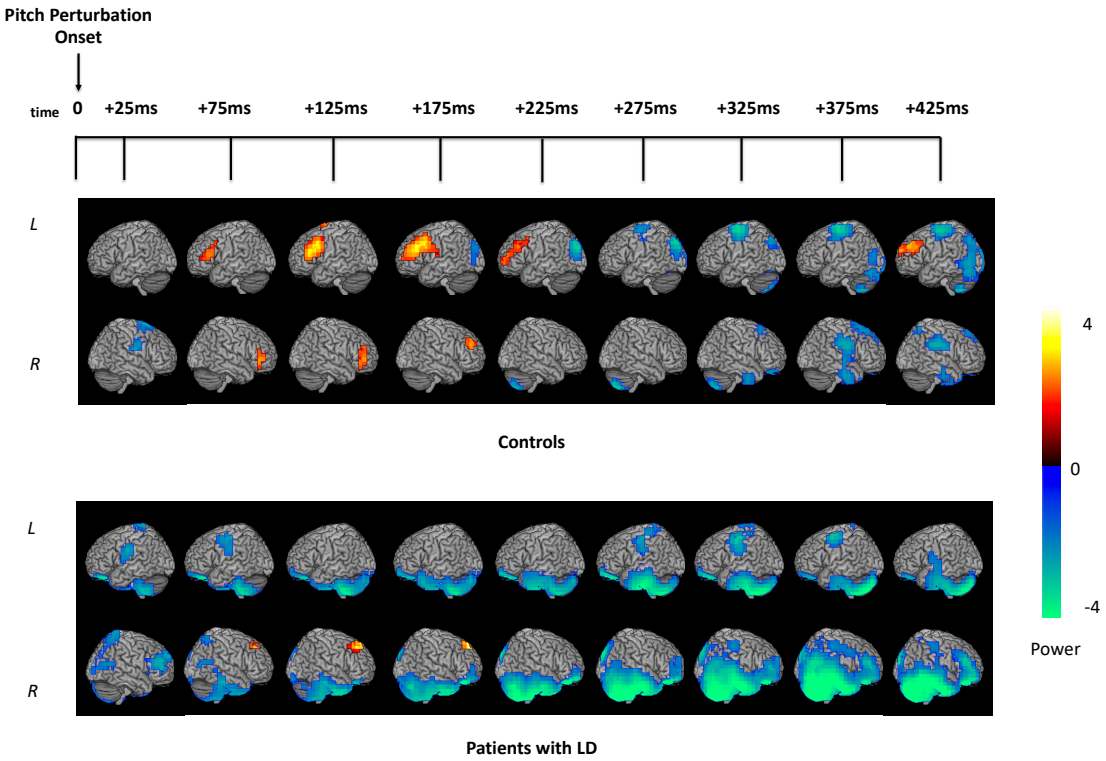
Supp. Fig. B.3 Neural activity in controls and patients with LD in the beta band (12 - 30 Hz) locked to pitch perturbation onset



Supp. Fig. B.4 Neural activity in controls and patients with LD in the high gamma band (65-150 Hz) locked to glottal movement onset



Supp. Fig. B.5 Neural activity in controls and patients with LD in the high gamma band (65-150 Hz) locked to voice onset



Supp. Fig. B.6 Neural activity in controls and patients with LD in the high gamma band (65-150 Hz) locked to pitch perturbation onset

Supp. Table B.1 Meta-analysis of studies of the CNS in patients with Adductor LD

Study	Modality	Task	Regions with increase in activity	Regions with decrease in activity	Regions with increase in connectivity (Seed - Target)	Regions with decrease in connectivity (Seed - Target)
Haslinger B, et al. (2005), Neurology; 65(10): 1562-9.	Silent event-related fMRI	Prolonged vowel phonation		**L superior M1 and S1, R inferior M1 and S1, ACC, mesial and L SMA, L and R dPMC, L SFG, R IFG, L MFG, R superior parietal,R occipital, L fusiform gyrus, L parahippocampal gyrus, L hemisphere of the cerebellum**		
		Whispered speech		**L superior M1, R inferior postcentral, L and R inferior parietal, mesial frontal, R STG, R MTG, R fusiform gyrus, R parieto-occipital**		
Ali et al. (2006), Journal of Speech, Language, and Hearing Research; Vol. 49 1127-1146	$H_2{}^{15}O$ PET	Narrative speech	**Cerebellum, dorsal and ventral precentral gyrus, anterior insula, ACC, A1, SII, posterior auditory association cortex**	**Posterior SMG, posterior MTG, dorsal postcentral gyrus, anterior auditory association cortex, PAG, SMA, anterior MTG**		
Simonyan & Ludlow (2010), Cerebral Cortex; 20:2749--2759	fMRI	Symptomatic syllable production	**M1, S1, Insula, STG, MTG, Cerebellum, Operculum, Basal Ganglia, Thalamus,**	**Midbrain**		
		Asymptomatic whimper	**M1, S1, Operculum, Insula, MCC, MTG**	**Insula, Thalamus, Basal Ganglia, Cerebellum, SMA**		
		Coughing	**M1, S1, Operculum, anterior insula, STG, MCC, MTG, Midbrain**	**Cerebellum**		
		Voluntary breathing	**M1, S1**	**Insula, Midbrain, Cerebellum, SMG**		
Kiyuna, A., et al. (2017), J Voice; **31**(3): p. 379 e1-379 e11.	fMRI	Reading five-digit numbers	**L MTG, L thalamus, L and R precentral gyrus, L and R postcentral gyrus, R insula, R Cerebellum VIII and IX, R putamen, L Cerebellum I-IV, R SMA**	**L Cerebellum Crus 1 and 2, R Cerebellum Crus 1, L STG, L Cerebellum VI**		
		Resting state			**L Thalamus - L Caudate; R precentral gyrus - L ITG, L temporal pole, L MTG; L postcentral gyrus - R frontal pole; L inferior operculum - R precentral and postcentral gyri; Cerebellum (vermis I, II) - R lateral occipital cortex, R superior parietal lobule; R Cerebellum (IX) - R precentral and postcentral gyri**	**L insula - R angular gyrus, R lateral occipital cortex; R thalamus - L MFG, L IFG, pars triangularis; L precuneus - L and R lingual gyrus; R precentral gyrus - R occipital pole**
Khosravani S, et al. (2019), Clinical Neurophysiology; 130(6): 1033-40.	EEG	Vowel vocalisation	**L somatosensory-premotor cortices (late vocalisation, gamma band)**	**L motor cortex (early vocalisation, alpha band)**		
Daliri A, et al. (2020), Journal of Speech, Language, and Hearing Research; 63(2): 421-32.	fMRI	Normal sentence production and Sentence production under masking noise	**L ventral sensorimotor cortex, L anterior planum temporale, L posterior STG / planum temporale**			
		Resting state			**L mid-Rolandic cortex - L Heschl's gyrus, L posterior STG, R Heschl's gyrus; L ventral Rolandic cortex - L posterior STG; R mid-Rolandic cortex - L posterior STG**	

Abbreviations used: fMRI = Functional Magnetic Resonance Imaging, PET = Positron Emission Tomography, EEG = Electroencephalography, L = left, R = right, M1 = primary motor cortex, S1 = primary somatosensory cortex, ACC = anterior cingulate cortex, SMA = supplementary motor area, dPMC = dorsal premotor cortex, SFG = superior frontal gyrus, IFG = inferior frontal gyrus, MFG = middle frontal gyrus, STG = superior temporal gyrus, MTG = middle temporal gyrus, A1 = primary auditory cortex, SII = secondary somatosensory cortex, SMG = supramarginal gyrus, PAG = periaqueductal grey, MCC = middle cingulate cortex, ITG = inferior temporal gyrus.

Supp. Table B.2 Peak voxels with significant beta-band activity differences with respect to glottal movement onset

Hemisphere	Peak	MNI Coordinates	Anatomical Labels	Time with respect to glottal movement onset
Left	1	[-41.2 -78.1 32.2]	Angular Gyrus (Brodmann Area 39)	-125 to +125ms
	2	[-40.4 -81.9 28.3]	Superior Occipital Gyrus (Brodmann Area 19)	-125 to +125ms
	3	[-41.9 -70.3 25.9]	Middle Temporal Gyrus (Brodmann Area 39)	-125ms
	4	[-23.3 -35.8 -25.7]	Cerebellar Anterior Lobe	-125ms
	5	[-59.0 -3.8 22.0]	Precentral Gyrus (Brodmann Area 4)	-125 to +25ms
	6	[-59.0 2.5 22.0]	Precentral Gyrus (Brodmann Area 6)	-125 to +75ms
	7	[-63.7 -29.0 27.5]	Inferior Parietal Lobule (Brodmann Area 40)	-125 to -25ms
	8	[-55.2 19.6 -18.4]	Superior Temporal Gyrus (Brodmann Area 38)	-125 to +125ms
	9	[-55.2 28.1 -8.9]	Inferior Frontal Gyrus (Brodmann Area 47)	-125 to +125ms
	10	[-55.9 4.1 20.4]	Inferior Frontal Gyrus (Brodmann Area 6)	-125 to +75ms
	11	[-48.2 -76.4 26.7]	Middle Temporal Gyrus (Brodmann Area 39)	-75 to +125ms
	12	[-58.3 3.3 15.6]	Inferior Frontal Gyrus (Brodmann Area 44)	+75 to +125ms
	13	[-51.3 21.9 8.5]	Inferior Frontal Gyrus (Brodmann Area 45)	+75ms
	14	[-47.4 43.5 -17.6]	Inferior Frontal Gyrus (Brodmann Area 47)	+125ms
	15	[-14.8 -48.6 -57.2]	Cerebellar Tonsil	+125ms
	16	[-59.0 5.7 12.5]	Precentral Gyrus (Brodmann Area 44)	+125ms
	17	[-52.0 12.0 -24.7]	Superior Temporal Gyrus (Brodmann Area 38)	+125ms
	18	[-52.0 9.5 -30.2]	Middle Temporal Gyrus (Brodmann Area 38)	+125ms
	19	[-30.3 17.3 60.7]	Middle Frontal Gyrus (Brodmann Area 6)	+125ms
Right	1	[41.9 -69.1 34.6]	Precuneus (Brodmann Area 39)	-125 to +25ms
	2	[41.9 -69.1 39.3]	Inferior Parietal Lobule (Brodmann Area 39)	-125 to +125ms
	3	[49.0 -72.2 37.0]	Angular Gyrus (Brodmann Area 39)	-125 to +25ms
	4	[60.1 -27.9 -27.1]	Inferior Temporal Gyrus (Brodmann Area 20)	-125ms
	5	[24.5 25.2 62.3]	Superior Frontal Gyrus (Brodmann Area 8)	-25 to +125ms
	6	[16.6 -77.8 34.6]	Precuneus (Brodmann Area 19)	+75ms
	7	[8.0 -83.0 47.0]	Precuneus (Brodmann Area 7)	+75 to +125ms
	8	[59.3 23.6 -1.0]	Inferior Frontal Gyrus (Brodmann Area 45)	+75 to +125ms
	9	[48.2 52.1 -8.9]	Middle Frontal Gyrus (Brodmann Area 10)	+75 to +125ms
	10	[31.6 -84.1 -14.4]	Middle Occipital Gyrus (Brodmann Area 19)	+75 to +125ms
	11	[23.7 -77.0 -24.7]	Cerebellar Posterior Lobe	+75 to +125ms
	12	[22.1 -75.4 -9.7]	Lingual Gyrus	+75 to +125ms
	13	[49.8 -58.0 -16.8]	Fusiform Gyrus	+125ms
	14	[39.5 -53.2 44.9]	Inferior Parietal Lobule	+125ms

Supp. Table B.3 Peak voxels with significant high-gamma-band activity differences with respect to glottal movement onset

Hemisphere	Peak	MNI Coordinates	Anatomical Labels	Time with respect to glottal movement onset
Left	1	[-64.0 -27.0 39.0]	Postcentral Gyrus (Brodmann Area 40)	+25 to +125ms
	2	[-65.3 -25.5 -8.4]	Middle Temporal Gyrus (Brodmann Area 21)	+125ms
	3	[-61.4 5.6 -2.9]	Superior Temporal Gyrus (Brodmann Area 22)	+125ms
Right	1	[46.6 -74.2 -57]	Cerebellar Posterior Lobe	-125 to -75ms
	2	[48.0 37.0 -9.0]	Inferior Frontal Gyrus (Brodmann Area 47)	+75 to +125ms
	3	[46.6 4.4 -40.6]	Middle Temporal Gyrus (Brodmann Area 21)	+75 to +125ms
	4	[62.1 -26.9 8.2]	Superior Temporal Gyrus (Brodmann Area 41)	+125ms

Supp. Table B.4 Peak voxels with significant beta-band activity differences with respect to voice onset

Hemisphere	Peak	MNI Coordinates	Anatomical Labels	Time with respect to voice onset
Left	1	[-54.4 14.2 -27.1]	Superior Temporal Gyrus (Brodmann Area 38)	-125 to -25ms
	2	[-54.4 9.5 -32.6]	Middle Temporal Gyrus (Brodmann Area 38)	-125 to -75ms
	3	[-33.4 42.0 -17.6]	Middle Frontal Gyrus (Brodmann Area 47)	-125ms
	4	[-55.2 20.4 14.0]	Inferior Frontal Gyrus (Brodmann Area 45)	-125 to -75ms
	5	[-31.8 11.1 60.7]	Middle Frontal Gyrus (Brodmann Area 6)	-125 to +25ms
	6	[-14.8 -13.7 68.6]	Superior Frontal Gyrus (Brodmann Area 6)	-125 to -75ms
	7	[-38.8 -27.5 45.7]	Postcentral Gyrus (Brodmann Area 2)	-125 to -75ms
	8	[-39.6 -76.5 38.6]	Precuneus (Brodmann Area 19)	-125 to -75ms
	9	[-39.6 -83.5 27.5]	Superior Occipital Gyrus (Brodmann Area 19)	-125 to -75ms
	10	[-42.7 -79.5 30.6]	Angular Gyrus (Brodmann Area 39)	-125 to -75ms
	11	[-31.8 -27.5 69.4]	Precentral Gyrus (Brodmann Area 4)	-25 to +125ms
	12	[-52.8 -26.7 58.3]	Postcentral Gyrus (Brodmann Area 1)	-25 to +125ms
	13	[-31.1 -85.0 -11.2]	Inferior Occipital Gyrus (Brodmann Area 18)	-25 to +125ms
	14	[-15.5 -58.7 -57.9]	Cerebellar Tonsil	-25 to +125ms
Right	1	[22.9 -66.7 62.3]	Superior Parietal Lobule (Brodmann Area 7)	-125 to -75ms
	2	[22.1 -62.0 61.5]	Superior Parietal Lobule (Brodmann Area 7)	-125 to -75ms
	3	[34.0 -35.8 69.4]	Postcentral Gyrus (Brodmann Area 1)	-125 to +125ms
	4	[25.3 -84.9 -16.8]	Middle Occipital Gyrus (Brodmann Area 19)	-125 to +125ms
	5	[56.9 -59.6 -17.6]	Inferior Temporal Gyrus (Brodmann Area 37)	-25 to +125ms
	6	[47.4 -50.9 -50.0]	Cerebellar Tonsil	-25 to +125ms
	7	[34.0 -28.7 67.0]	Precentral Gyrus (Brodmann Area 4)	+25 to +125ms
	8	[6.3 -3.3 67.0]	Superior Frontal Gyrus (Brodmann Area 6)	+25 to +75ms
	9	[53.7 20.4 13.3]	Inferior Frontal Gyrus (Brodmann Area 44)	+75 to +125ms
	10	[41.1 22.0 37.0]	Middle Frontal Gyrus (Brodmann Area 8)	+75 to +125ms
	11	[6.3 3.8 69.4]	Superior Frontal Gyrus (Brodmann Area 6)	+125ms

Supp. Table B.5 Peak voxels with significant high-gamma-band activity differences with respect to voice onset

Hemisphere	Peak	MNI Coordinates	Anatomical Labels	Time with respect to voice onset
Left	1	[-8.0 53.0 39.0]	Superior Frontal Gyrus (Brodmann Area 9)	-125ms
	2	[-64.0 -19.0 23.0]	Postcentral Gyrus (Brodmann Area 1)	-125 to -25ms
	3	[-16.0 61.0 -17.0]	Superior Frontal Gyrus (Brodmann Area 11)	-125 to +125ms
	4	[-56.0 -11.0 -17.0]	Middle Temporal Gyrus (Brodmann Area 21)	-125 to +125ms
Right	1	[41.3 6.5 -42.2]	Inferior Temporal Gyrus (Brodmann Area 20)	-125 to +25ms
	2	[48.0 -19.0 -33.0]	Inferior Temporal Gyrus (Brodmann Area 20)	-125 to +25ms
	3	[64.0 -27.5 8.2]	Superior Temporal Gyrus (Brodmann Area 41)	-125 to +125ms
	4	[8.0 -67.0 39.0]	Precuneus (Brodmann Area 7)	-125 to -75ms
	5	[46.8 -68.5 45.2]	Inferior Parietal Lobule (Brodmann Area 39)	+75ms

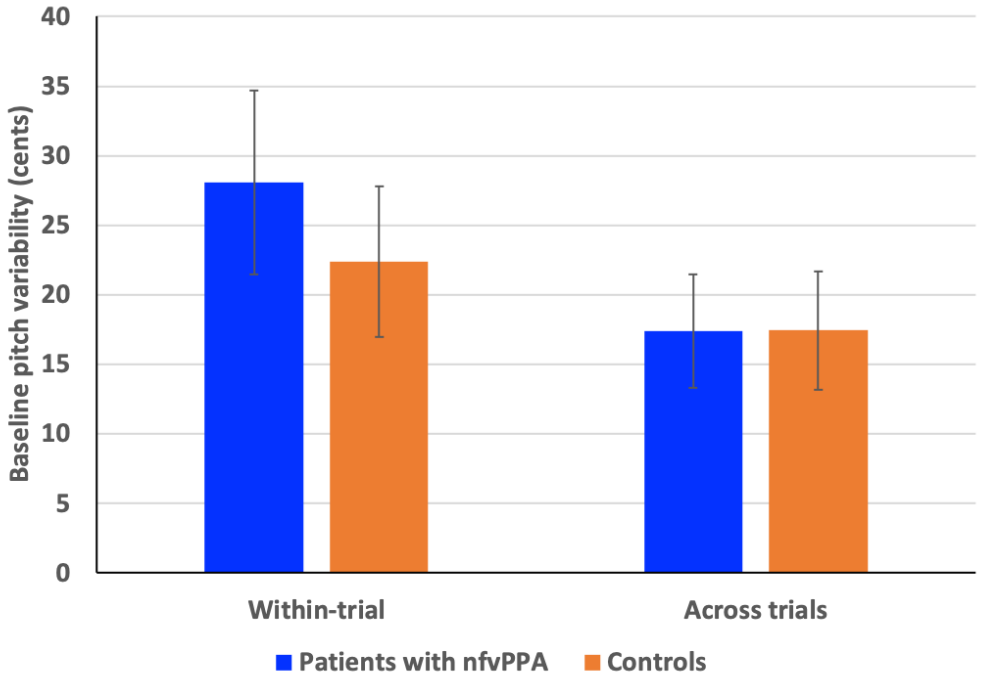
Supp. Table B.6 Peak voxels with significant beta-band activity differences with respect to pitch perturbation onset

Hemisphere	Peak	MNI Coordinates	Anatomical Labels	Time with respect to voice onset
Left	1	[-39.6 -75.0 -50.8]	Cerebellar Inferior Semi-Lunar Lobule	+25 to +225ms
	2	[-48.2 20.4 4.6]	Inferior Frontal Gyrus (Brodmann Area 45)	+25 to +175ms
	3	[-38.8 23.5 44.1]	Middle Frontal Gyrus (Brodmann Area 8)	+25ms
	4	[-25.6 -59.0 44.9]	Superior Parietal Lobule (Brodmann Area 7)	+175 to +375ms
	5	[-52.8 27.0 -11.2]	Inferior Frontal Gyrus (Brodmann Area 47)	+225ms
	6	[-25.6 -51.9 43.3]	Precuneus (Brodmann Area 7)	+225 to +375ms
	7	[-46.6 34.9 -11.2]	Inferior Frontal Gyrus (Brodmann Area 47)	+275ms
Right	1	[48.0 21.0 -25.0]	Superior Temporal Gyrus (Brodmann Area 38)	+25 to +425ms
	2	[62.4 26.0 10.1]	Inferior Frontal Gyrus (Brodmann Area 9)	+25 to +425ms
	3	[54.5 20.4 14.8]	Inferior Frontal Gyrus (Brodmann Area 44)	+25 to +325ms
	4	[49.8 -3.3 33.8]	Precentral Gyrus (Brodmann Area 6)	+75 to +275ms
	5	[49.8 3.8 38.6]	Middle Frontal Gyrus (Brodmann Area 6)	+125 to +275ms
	6	[55.3 3.8 31.4]	Inferior Frontal Gyrus (Brodmann Area 6)	+125 to +275ms
	7	[55.3 26.0 14.0]	Inferior Frontal Gyrus (Brodmann Area 9)	+125 to +325ms
	8	[55.3 -67.5 -41.3]	Cerebellar Posterior Lobe	+175 to +325ms
	9	[18.2 52.9 38.6]	Superior Frontal Gyrus (Brodmann Area 9)	+225 to +325ms
	10	[15.0 62.4 7.7]	Superior Frontal Gyrus (Brodmann Area 10)	+375 to +425ms
	11	[55.3 -52.5 -43.7]	Cerebellar Tonsil	+375 to +425ms
	12	[45.8 -66.7 45.7]	Inferior Parietal Lobule (Brodmann Area 39)	+375 to +425ms
	13	[63.2 -12.1 -16.0]	Middle Temporal Gyrus (Brodmann Area 21)	+425ms

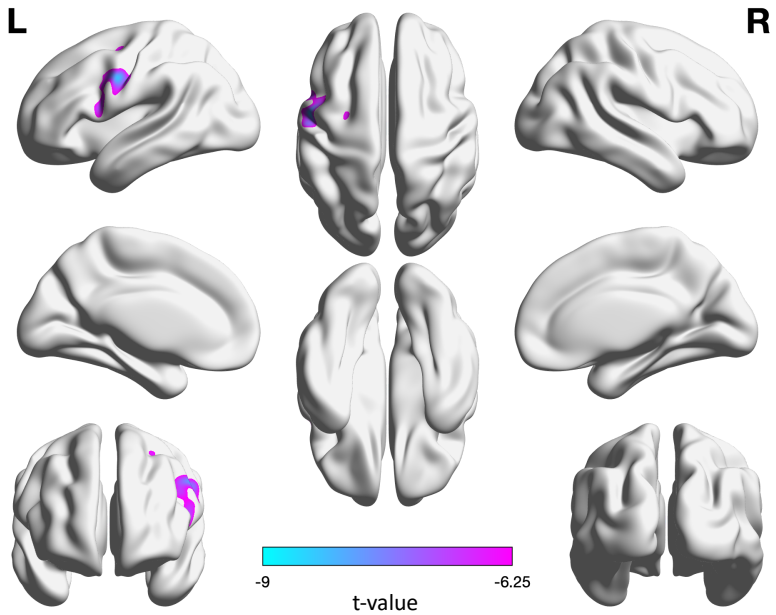
Supp. Table B.7 Peak voxels with significant high-gamma-band activity differences with respect to pitch perturbation onset

Hemisphere	Peak	MNI Coordinates	Anatomical Labels	Time with respect to pitch perturbation onset
Left	1	[-32.0 29.0 7.0]	Inferior Frontal Gyrus (Brodmann Area 45)	+25 to +125ms
	2	[10/9 35.1 -26.5]	Rectal Gyrus (Brodmann Area 11)	+25 to +425ms
	3	[-38.1 -42.0 -54.0]	Cerebellar Tonsil	+25 to +425ms
	4	[-31.1 -81.9 38.1]	Angular Gyrus (Brodmann Area 39)	+175 to +275ms
	5	[-48.0 -20.7 -34.3]	Inferior Temporal Gyrus (Brodmann Area 20)	+225 to +325ms
Right	1	[52.3 37.2 17.6]	Middle Frontal Gyrus (Brodmann Area 46)	+25 to +225ms
	2	[14.6 29.3 -25.0]	Orbital Gyrus (Brodmann Area 47)	+125 to +425ms
	3	[56.0 -27.5 -24.9]	Inferior Temporal Gyrus (Brodmann Area 20)	+125 to +425ms
	4	[48.0 45.0 23.0]	Middle Frontal Gyrus (Brodmann Area 46)	+125 to +225ms
	5	[56.3 -74.0 -33.0]	Cerebellar Posterior Lobe	+175 to +425ms
	6	[8.0 -91.0 31.0]	Cuneus (Brodmann Area 19)	+175 to +275ms
	7	[46.0 -84.3 -16.2]	Inferior Occipital Gyrus (Brodmann Area 18)	+225 to +425ms
	8	[24.0 33.3 54.6]	Superior Frontal Gyrus (Brodmann Area 8)	+325ms

Appendix C: Supplementary Material to Chapter 4



Supp. Fig. C.1 Baseline vocal range
Baseline pitch variability (in a 200ms pre-perturbation window) in patients with nfvPPA (n = 18) does not differ from that in controls (n = 17) both within-trial and across trials.



Supp. Fig. C.2 Cortical atrophy map
The 20 patients participating in this study showed cortical atrophy in regions typically known to be atrophied in nfvPPA: left ventral motor and left ventral premotor cortices. A voxel-based morphometry two-sample t-test on grey matter volume was performed between the 20 patients and 100 age and sex-matched healthy controls (family-wise error corrected for $p < 0.05$).

Publishing Agreement

It is the policy of the University to encourage open access and broad distribution of all theses, dissertations, and manuscripts. The Graduate Division will facilitate the distribution of UCSF theses, dissertations, and manuscripts to the UCSF Library for open access and distribution. UCSF will make such theses, dissertations, and manuscripts accessible to the public and will take reasonable steps to preserve these works in perpetuity.

I hereby grant the non-exclusive, perpetual right to The Regents of the University of California to reproduce, publicly display, distribute, preserve, and publish copies of my thesis, dissertation, or manuscript in any form or media, now existing or later derived, including access online for teaching, research, and public service purposes.

DocuSigned by:
Hardik Kothare
C30B5E2D913D469...
Author Signature

12/1/2020
Date