Sensory modulation during speech movement planning in stuttering and nonstuttering adults

Elise LeBovidge

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington
2024

Reading Committee:
Ludo Max, Chair
David J. Perkel
Richard Wright
Chet Moritz

Program authorized to offer degree:
Speech and Hearing Sciences

University of Washington

ABSTRACT

Sensory modulation during speech movement planning
in stuttering and nonstuttering adults

Elise LeBovidge

Chair of the Supervisory Committee:
Ludo Max
Department of Speech and Hearing Sciences

Stuttering is a complex neurodevelopmental communication disorder with an approximate prevalence of one percent, affecting about eighty million people worldwide. To date, the specific neural mechanisms associated with stuttering and its overt symptoms are still largely unknown, despite evidence of speech and nonspeech sensorimotor differences and extensive sensorimotor network involvement. This dissertation presents novel results from three electrophysiological experiments that aimed to improve our understanding of the sensorimotor mechanisms underlying stuttering. First, I extended our previously established pre-speech auditory modulation (PSAM) paradigm to investigate the modulatory influence of speech movement planning on auditory processing (assessed with auditory evoked potentials) prior to the production of full sentences rather than just isolated words. Second, I investigated whether pre-speech sensory modulation can also be detected in the somatosensory domain. For this study, sensory processing during speech movement planning was assessed by means of somatosensory-evoked potentials in response to experimentally applied orofacial skin stretches. Third, I explored the functional relevance of PSAM in a preliminary study evaluating the correlation between speakers' extent of PSAM and kinematics-based estimates of their reliance on feedforward versus feedback control. The results from these experiments replicated key findings from our laboratory's prior work and provide several new insights into the phenomenon of pre-speech sensory modulation in stuttering and nonstuttering speakers. Specifically, *Study 1* revealed substantial auditory modulation during speech movement planning for both words and sentences. *Study 2* found no pre-speech somatosensory modulation (PSSM) for either typically fluent speakers or speakers who stutter. Thus, the specific cutaneous mechanoreceptors stimulated with the present experimental technique may not be subject to modulation by pre-motor systems involved in speech planning. *Study 3* provided preliminary evidence that PSAM is associated with neural mechanisms directly involved in feedforward mechanisms rather than preparing auditory cortex for optimized feedback monitoring. Taken together, the overall findings confirm that adults who stutter show atypical motor-to-auditory interactions during speech movement planning and suggest that this observation relates to a fundamental limitation in the use of feedforward control mechanisms for speech production. The results suggest new testable hypotheses for further research into the neural mechanisms underlying stuttering and speech sensorimotor control in general.

ACKNOWLEDGMENTS

This endeavor simply could not have been possible without the hundreds of research participants that I have worked with over the past six years. In particular, I am deeply appreciative of the stuttering community in the greater Seattle area. Many research participants went above and beyond to help advance our understanding of stuttering, and I cannot fully express my gratitude for every participant’s patience and willingness to do even the weirdest of experimental tasks.

I have learned so much from each interaction, and I have cherished each conversation.

Thank you for sharing your lives with me.

This document is divided into six sections.

The **first section** provides a brief overview of stuttering,
included to supplement the reader's understanding of this disorder of speech fluency
and to help situate the experiments conducted for this dissertation in a broader context.

The **second section** provides an overview of the organization of the remainder of the
dissertation, serving as a general introduction to the included research program.

The subsequent **three sections** (*Studies 1-3*) comprise the empirical research conducted
for this dissertation, each section formatted as a stand-alone manuscript.

The **final section** summarizes the conducted experiments and integrates the work
into the current literature on stuttering with an emphasis on future directions.

Table of Contents

Section I: A Brief Introduction to Stuttering 10

1. Symptomatology 11
2. Onset and spontaneous recovery 13
3. Theories on etiology: A brief overview 13
4. Theories on etiology: Impaired sensorimotor function 17
5. Treatment options for children and adults 19
6. References 22

Section II: General Introduction 31

1. Framing the dissertation 31
2. References 34

Section III: Pre-speech auditory modulation prior to sentence-level speech (*Study 1*) 37

1. ABSTRACT 38
2. INTRODUCTION 39
3. METHODOLOGY 41
    - *3.1. Participants 41*
    - *3.2. Experimental paradigm 42*
    - *3.3. Data analysis: EEG processing 46*
    - *3.4. Statistical analyses 49*
4. RESULTS 50
    - *4.1. N1 peak amplitude 50*
    - *4.2. N1 peak latency 51*
    - *4.3. Pre-speech auditory modulation 51*
5. DISCUSSION AND CONCLUSIONS 51
6. ARTWORK 56
7. ACKNOWLEDGEMENTS 60
8. AUTHOR CONTRIBUTIONS 60
9. FUNDING SOURCES 60
10. REFERENCES 60

Section IV: Cortical somatosensory responses evoked by orofacial skin stretch experimentally applied during speech movement planning in stuttering and nonstuttering adults (*Study 2*) **63**

1. ABSTRACT 64
2. NEW AND NOTEWORTHY 65
3. INTRODUCTION 65
4. MATERIALS AND METHODS 69
    - *4.1. Participants* *69*
    - *4.2. Experimental paradigm* *71*
    - *4.3. "No stretch" auditory control group* *74*
    - *4.4. Data analysis* *75*
    - *4.5. Statistical analyses* *78*
5. RESULTS 79
    - *5.1. N1 peak amplitude* *79*
    - *5.2. N1 peak latency* *80*
    - *5.3. Pre-speech somatosensory modulation (PSSM) and stuttering severity* *80*
6. DISCUSSION 80
7. CONCLUSIONS AND FUTURE DIRECTIONS 84
8. DATA AVAILABILITY 85
9. ACKNOWLEDGMENTS 85
10. GRANTS 85
11. DISCLOSURES 86
12. AUTHOR CONTRIBUTIONS 86
13. REFERENCES 86
14. FIGURE LEGENDS 93
15. TABLES 99

Section V: Preliminary insights into the functional relevance of pre-speech auditory modulation: Correlations with kinematic estimates of feedback-based control in adults who stutter (*Study 3*) **100**

1. ABSTRACT 101
2. INTRODUCTION 102
3. METHODOLOGY 106

3.1. Participants 106
3.2. Overall experimental paradigm 108
3.3. Auditory Modulation task 109
3.4. Kinematic task 112
3.5. Data analysis: Auditory Modulation task 114
3.6. Data analysis: Kinematic task 117
3.7. Statistical analyses 120
4. RESULTS 122
4.1. Auditory Modulation task 122
4.1.1. N1 peak amplitude 122
4.1.2. N1 peak latency 122
4.1.3. Pre-speech auditory modulation (PSAM) and stuttering severity 123
4.2. Kinematic task 123
4.2.1. Correlations between peak kinematic parameters and movement extent 125
4.2.2. Feedforward vs. feedback control estimates 127
4.2.3. Correlating PSAM with feedback vs. feedforward control estimates 129
5. DISCUSSION 129
6. CONCLUSION 136
7. ACKNOWLEDGEMENTS 136
8. REFERENCES 137
9. FIGURES 142
10. TABLES 157

Section VI: General Discussion 159
1. Conclusions and future directions 159
2. References 162

Section I: A Brief Introduction to Stuttering

“Actually, there is not enough information about the essential nature of stuttering to allow more than a general outline of its patterning. But if we must fail, let us fail forward! Others will mend and build upon the insecure foundations on which this precarious monument to human benightedness is so hesitantly erected.”
(Van Riper, 1971)

“First, it must be clear that a considerable amount is known about stuttering. Persons who describe it as ‘a mysterious and ill-understood disorder’ are merely confessing that they have not read the recent literature.”
(Andrews et al., 1983)

“The subject of stuttering is widely acknowledged to be confusing and difficult. A substantial part of this condition is due to the extensive literature about the disorder, so much of which is not scientifically oriented or supported, but is instead replete with conjecture, fabrication, and an extensive body of substance best described as lore.”
(Wingate, 2002)

“There are, thus, no symptoms of stuttering; stuttering is stuttering!”
(Wingate, 2002)

Stuttering is a neurodevelopmental communication disorder with an approximate prevalence of one percent, affecting around 80 million people worldwide (Bloodstein & Bernstein Ratner, 2008). For many individuals who stutter (IWS), quality of life is negatively affected by the hallmark symptoms of the disorder: overt speech behaviors such as involuntary repetitions and prolongations of sounds and syllables and covert behaviors such as substituting words and avoiding social interactions (Bloodstein & Bernstein Ratner, 2008; Craig et al., 2009). In addition to impacting day-to-day communication, stuttering can also lead to negative emotions, anxiety (Iverach et al., 2009; Kraaimaat et al., 2002), suicidal ideation (Tichenor et al., 2023), and stigmatization, affecting self-esteem (Blood et al., 2011), quality of life (Craig et al., 2009), communication skills and attitudes (Blood et al., 2001), academic and occupational achievement (Craig & Calver, 1991; Crichton-Smith, 2002; Klein & Hood, 2004; Williams et al., 1969), and interpersonal relationships (Davis et al., 2002; Linn & Caruso, 1998; Van Borsel et al., 2011).

It is important to note that there are distinct types of stuttering that are beyond the scope of this review. Here, *stuttering* will be considered synonymous with what is more accurately described as *persistent developmental stuttering*. In other words, stuttering that is neurodevelopmental in nature, appearing during the acquisition of speech in early childhood. This is in contrast to stuttering that may first appear due to an acute neurological injury (*acquired neurogenic stuttering*) or a specific psychological condition (*psychogenic stuttering*).

1. Symptomatology

Over the years, clinicians and researchers have struggled to formally define stuttering (for an extensive overview, see chapter 2 of Wingate, 2002). Leading stuttering experts in the mid-20th century were often imprecise, describing stuttering as “an anticipatory, apprehensive, hypertonic avoidance reaction” (Johnson, 1967), “an anticipatory struggle” (Bloodstein, 1958), and “words that are improperly patterned in time and the speaker’s reactions thereto” (Van Riper, 1971).

Today, the overt speech symptoms that experts consider to be stuttering are always within-word dysfluencies (Conture, 1990). These include both audible and inaudible sound prolongations (e.g., *ssssssstutter* and ___*table*, respectively, where ___ indicates that no sound is being produced, but the articulators are in the proper posture), syllable and sound repetitions (e.g., *te-te-te-telescope* and *m-m-m-m-my*), and broken words (e.g., *schoo-ool*). These dysfluencies often occur together with physical movements that are known as *secondary characteristics.* Common examples of these behaviors include excessive or forceful eye blinking, finger or foot tapping, and tongue or lip clicking. It is generally believed that these accessory behaviors are learned over time through classical conditioning (Guitar, 2019).

In addition to the overt symptoms of stuttering, there are covert symptoms that can be even more disruptive to an individual's psychological well-being. Many individuals who stutter substitute entire words to avoid syllables or sounds that they find particularly difficult and/or anticipate struggling with in a given situation (Jackson et al., 2015; Vanryckeghem et al., 2004). In fact, even though typical speakers also use coping responses during speech production (e.g., interjecting sounds and avoiding eye contact), many of the most reported coping responses by IWS were not even among the top ten reported by typical speakers (Vanryckeghem et al., 2004). In addition, more extreme avoidance behaviors are often seen as many IWS report avoiding social interactions to eschew judgment and unknown listener reactions –behaviors that can develop in early childhood (Crichton-Smith, 2002; Plexico et al., 2009; Solomon, 1939).

Presently, the most common definition of stuttering is founded in the disorder's well-known speech symptoms. Some have, however, suggested that stuttering should be more inclusively defined to consider the covert components of the disorder, as IWS will always be monitoring and subtly adjusting their speech (Jackson et al., 2015). Furthermore, when adults who stutter were asked to define *stuttering* in a large, open-ended survey (N = 430), shared definitional themes included the loss of control, the emotional, cognitive, and behavioral reactions, the limitations and impact of the disorder, and the perceived influence of listeners (Tichenor & Yaruss, 2019). To individuals who stutter, the disorder clearly encompasses a wide collection of experiences and behaviors beyond the overt speech symptoms perceived by listeners. There has therefore been a recent push to specifically distinguish between stuttering moments themselves (i.e., the overt *symptoms* of the disorder) and the unique lived experience of the stuttering *condition* (Tichenor et al., 2022), extending an earlier call-to-action made by researcher Dean Williams (Williams, 1957).

2. Onset and spontaneous recovery

Children begin to stutter during typical speech development, as early as age 2 (Yairi et al., 1993). A majority of the children who begin stuttering will naturally "outgrow" the disorder on their own, or with the help of speech-language pathologists (Yairi & Ambrose, 2013). Studies investigating the frequency of this so-called *spontaneous recovery* – including retrospective studies relying predominately on parental memory – vary in several key methodological aspects (e.g., how they define "stuttering") which impact the statistics that they report. However, it is widely regarded that about 80% of children who stutter will go on to spontaneously recover (Yairi et al., 1993; Yairi & Ambrose, 1992a, 1992b). Thus, only about a quarter of these children go on to stutter as adults. Despite the close to 1:1 sex ratio near onset, more biological females tend to spontaneously recover from stuttering, leading to the large proportion of biological males who stutter as adults (Seider et al., 1983). Interestingly, there are rare cases that report spontaneous recovery after early childhood (Anderson & Felsenfeld, 2003), but these anecdotes are exceedingly rare and under-researched.

3. Theories on etiology: A brief overview

Of course, of great interest to clinicians, researchers, and IWS across the globe is a specific, targetable etiology for stuttering. For millennia, humans have sought to isolate a particular underlying cause (Rieber & Wollock, 1977; Van Riper, 1971), but unfortunately, this neurodevelopmental disorder is complex, evading all "quick and easy" fixes that have been suggested. Most early hypotheses focused on some aspect of the tongue, including Francis Bacon's suggestion to use hot wine to re-establish the tongue's optimal moisture and warmth (Rieber & Wollock, 1977; Van Riper, 1971). More contemporary suggestions as to the fundamental cause of stuttering have included listener reactions to typical dysfluencies (i.e., the diagnosogenic theory; Johnson, 1942), as well as deficits in psychological fortitude (Solomon, 1939), timing and sequencing (Etchell et al., 2014), auditory processing (Molt & Guilford, 1979), basal ganglia function (Alm, 2004), auditory feedback weighting (Civier et al., 2010), and sensorimotor control (Hickok et al., 2011; Max, 2004; Max et al., 2004). A few of these theories will be briefly reviewed here.

Stuttering has often been considered to be a disorder of timing and/or sequencing (Cooper & Allen, 1977; Etchell et al., 2014; Mackay & Macdonald, 1984; Van Riper, 1971), and there has undoubtedly been a recent uptick in the popularity of this hypothesis (see, e.g., Chang et al., 2019; Garnett et al., 2019). This suggestion may have initially stemmed from the observation that speaking under certain rhythmic conditions (e.g., with a metronome or a second speaker) can drastically improve fluency for IWS. However, early findings of differences in orofacial movement sequencing patterns (Caruso et al., 1988) have come under question, as IWS seem to show similar sequencing patterns as typical speakers for both speech and non-speech movements (Max et al., 2003). Furthermore, no differences have been seen for IWS in their ability to synchronize orofacial speech, orofacial nonspeech, or finger movements to a metronome when compared to typical speakers (Max & Yudman, 2003). Children who stutter also do not differ from typically developing kids with respect to bimanual motor timing, as evidenced by similar rates of clapping and inter-clap intervals (Hilger et al., 2016). Studies have certainly varied in their methods of inquiry and analysis techniques, but repeated findings of typical abilities to perform rapid, rhythmic movements of various effectors seem to suggest that stuttering is not associated with deficits in this realm (Strother & Kriegman, 1943).[1] In fact, there

---

[1] For a detailed discussion, see “Organicity in Stuttering” in van Riper (1971) and “The Person Who Stutters: Motor Abilities” in Bloodstein & Bernstein Ratner (2008).

seems to be large inter-individual variability for timing skills even amongst typical speakers, with ranges overlapping between those who stutter and those who do not (Cooper & Allen, 1977). Not only are IWS able to properly time and coordinate movements, but their learning of speech sequences is not impaired compared to typical speakers, although their movement is often slower (e.g., Masapollo et al., 2021). In sum, stuttering does not fundamentally appear to be a problem with motor coordination or sequence learning, although neural data does exist that suggests network-level asynchronization or neural miscommunication (Chang et al., 2019; Sengupta et al., 2016, 2019).

Observing the high heritability of stuttering within families, researchers have also investigated the possibility that stuttering is fundamentally a genetic disorder. Early research relied on large family histories, but due to the increased accessibility of genetic sequencing, genetic research has now shifted to large-scale, genome-wide linkage and association studies (e.g., Riaz et al., 2005; Shaw et al., 2021). It was recently discovered, following such methods, that mutations in a lysosomal enzyme-targeting pathway have been associated with stuttering in a fraction of individuals (Kang et al., 2010). Naturally, researchers have been intrigued by this finding ever since. It still remains unclear how mutations in a so-called “housekeeping gene” can yield the dysfluencies we know and label as stuttering. However, after the identification of this gene, animal models with similar genetic mutations were developed to begin systematically studying the role it plays in vocalization and neurodevelopment. Two such studies were able to produce vocalization patterns in mice that seemed to approximate stuttering in humans (Barnes et al., 2016; Han et al., 2019). These behaviors co-occurred with abnormal astrocyte development, particularly in the corpus callosum (Han et al., 2019). Interestingly, abnormal neural development in children who stutter has also been linked to gene expression data. The expression of two intracellular trafficking genes known to be associated with stuttering was spatially correlated with gray matter volume differences (relative to typically fluent children) in a recent analysis (Chow et al., 2020). Future research into the genetic basis of stuttering is clearly much needed, as identifying potential genotypes that are directly linked to stuttering may allow for the exploration of novel gene therapy options for treatment.

Bolstered in part by increased technological accessibility in recent decades, researchers have also turned to neuroimaging techniques to try to identify the neural underpinnings of stuttering. As early as 1996, positron emission tomography (PET) scans during reading showed that IWS were over-recruiting the right inferior premotor regions and under-recruiting auditory cortex compared to individuals who do not stutter (Fox et al., 1996). Subsequent meta-analyses on neuroimaging findings in stuttering corroborated this early finding by showing that functional magnetic resonance imaging (fMRI) also consistently highlights right frontal overactivation and left fronto-parieto-temporal *under*-activation in IWS (Belyk et al., 2017; Neef et al., 2015). Another notable and consistent neuroimaging finding in stuttering is that of decreased white matter integrity along the left dorsal stream, and often also in the interhemispheric connections between sensorimotor cortices (Chang et al., 2008; Neef et al., 2015). However, new neuroimaging studies with IWS are being published at an astonishing rate, yet conclusions are often hard to interpret and synthesize with previous findings as many papers highlight extensive network-level differences compared to typical speakers (e.g., Gracco et al., 2022). In short, it is perhaps not surprising that many significant differences have been measured across cortical and subcortical structures in a population that has adapted to a highly stigmatized communication disorder (often for *decades*, considering the sample demographics of recent studies).

4. Theories on etiology: Impaired sensorimotor function

Supplementing the neuroimaging findings that underscore diverse structural and functional differences in key sensorimotor areas, behavioral studies have found that individuals who stutter show subtle sensorimotor differences. Specifically, researchers have shown that IWS are not able to use their sensory feedback to adjust their subsequent movements the same way as typical speakers (Daliri et al., 2018; Daliri & Max, 2018; Kim et al., 2020; Sengupta et al., 2016). In speech auditory-motor learning tasks, participants produce short, monosyllabic words while experimenters manipulate their auditory feedback either suddenly or gradually over time. A typical design involves incrementally shifting one or more of the formants (i.e., the resonant frequencies of the vocal tract that uniquely identify vowels) of the speech signal in near real-time. Typical speakers are able to use this altered feedback to adjust their subsequent productions in an attempt to offset the applied perturbation and bring their sensory feedback back to what they believe is more ideal. This motor learning behavior is quite automatic, subconscious, and effortless, and thus believed to be highly implicit in nature (Kim & Max, 2021; Shiller et al., 2020). However, IWS show a reduced capacity to integrate their sensory feedback into subsequent corrections, marked by a significant decrease in speech auditory-motor adaptation compared to typical speakers (Daliri et al., 2018; Daliri & Max, 2018; Kim et al., 2020; Sengupta et al., 2016). Related paradigms in which participants receive unexpected perturbations of auditory feedback during extended periods of phonation also show limitations for IWS compared to typical speakers, suggesting a broad impairment in integrating auditory feedback during speech production (Cai et al., 2012, 2014; Loucks et al., 2012).

Sensorimotor limitations are also seen in nonspeech motor learning tasks. A visuomotor analog of the above task can be performed where participants repeatedly reach their arm toward visual targets. This reaching movement is obscured from the participant's view so that the only feedback they are receiving is a visual representation of their hand's location on a screen. Gradually or abruptly, experimenters can visually rotate the participant's cursor to induce a sensory prediction error (i.e., the participant sees the cursor move in a slightly different direction than their veridical feedback). Just as in speech production, typical participants will use the altered visual feedback to adjust subsequent reaching movements and minimize this perceived error (for a detailed overview, see Krakauer et al., 2019). However, adults who stutter are also impaired at adapting in this visuomotor analog, showing a significant difference in their rate of learning compared to typically fluent adults (Kim et al., 2020). These sensorimotor adaptation differences across effector systems may suggest that IWS have difficulty learning the neural representations that map motor commands to their sensory consequences. There is extensive literature discussing sensorimotor learning paradigms and their theoretical motor control ramifications (e.g., Krakauer & Mazzoni, 2011; Shadmehr et al., 2010), so a more intensive discussion of these findings will not occur here. However, as briefly mentioned above, there is also research that suggests atypical weighting of sensory feedback in stuttering (Civier et al., 2010; Daliri et al., 2013; De Nil et al., 2001; Fukawa et al., 1988).

Another sensorimotor difference observed in stuttering is a motor-to-sensory modulation of auditory processing, a phenomenon our laboratory has named *pre-speech auditory modulation* (PSAM). We have discovered differences in auditory cortex responses immediately before speech onset in IWS as compared with typically fluent adults (Daliri & Max, 2015a, 2015b, 2018; Max & Daliri, 2019). Specifically, whereas typical speakers show a reduced auditory cortex response to probe tones presented during speech movement planning (i.e., PSAM), most IWS do not. Recent investigations in our laboratory have led to the hypothesis that this phenomenon may reflect ongoing communication between cortical areas involved in preparing outgoing motor commands (Li et al, 2024). Specifically, if auditory cortical neurons are recruited by (pre)motor areas during speech movement planning to aid in the fine-tuning and selection of the most appropriate motor command to be sent out, fewer auditory neuronal populations would be available to respond to the probe tones presented in our PSAM paradigm. The following sections (and *Study 3* in particular) will discuss this phenomenon and its hypothesized relation to feedforward control in more detail.

These recent results have all reinforced the hypothesis that stuttering may be due to a fundamental deficit in processes related to sensorimotor integration. Further evidence for a sensorimotor difference is found in the ability of certain sensory manipulations (e.g., delaying auditory feedback) to drastically decrease stuttering frequency (e.g., Johnson & Rosen, 1937; Andrews et al., 1982; Howell et al., 1987). At their core, these behavioral findings could reflect a poor mapping between sensory and motor systems (Hickok et al., 2011) or minimally activated internal models (Max et al., 2004). Of course, neuroimaging findings consistently highlight structural and functional differences in areas known to be relevant for sensorimotor processes (e.g., along the left dorsal stream), so these deficits may underlyingly be due to structural differences that limit functional behavior.

5. Treatment options for children and adults

Given that the percentage of children who spontaneously recover from stuttering (i.e., “outgrow” the disorder) has been reported to be as high as 80% (Yairi & Ambrose, 1992a), a large body of literature is devoted to debating whether or not young children who stutter should receive any therapy at onset at all (Curlee & Yairi, 1997; Ingham & Cordes, 1998; Onslow & Kelly, 2020; Onslow & Millard, 2012; Onslow & Yaruss, 2007). However, despite the high rate of recovery for children, it seems prudent to intervene as early as possible when therapy is accessible. This is particularly wise considering the evidence showing overall reduced quality of life and psychosocial ramifications due to stuttering that persists into adulthood (e.g., Blood et al., 2011; Craig et al., 2009; Iverach et al., 2009; Linn & Caruso, 1998; Van Borsel et al., 2011; Williams et al., 1969). Furthermore, the majority of children who recover will do so within approximately three years after stuttering onset (Yairi & Ambrose, 1999), suggesting that multiple years of speech and social development can still be affected in these cases. Given the evidence that communication attitudes are already impacted by dysfluency at such a young age (Vanryckeghem et al., 2005), therapy is clearly warranted.

There are two main camps of stuttering treatment for young children: *direct* and *indirect* treatment. The former involves active, speech-targeting therapy with the child themselves, while the latter is mainly composed of environmental adjustments made by the child’s caregivers. Indirect treatment, although incredibly popular amongst clinicians, is often considered inferior to a more direct approach (for further discussion, see Ingham & Cordes, 1998). One highly successful direct treatment program for young children is known as the Lidcombe program (Onslow et al., 2001). This program was developed by Australian researchers in the early 2000s and, as of September 2024, has 69 related publications on PubMed. Decades of research into this treatment program has shown high rates of success, in which “success” is defined as a significant reduction in stuttering frequency that is maintained over time (Jones et al., 2005; O’Brian et al., 2013). In other words, the majority of the children who receive this treatment will remain entirely fluent for the rest of their lives (Lincoln & Onslow, 1997). Children can attain this fluency after only a small number of sessions (Jones et al., 2000) and it is difficult for listeners to differentiate between typically-developing children and children who have been through the Lidcombe program (Lincoln et al., 1997).

Notably, there is no known cure for stuttering once it persists into adulthood. Treatment options for adults rely on managing the symptoms themselves, as stuttering will undoubtedly continue for the rest of the individual's life. Current treatment approaches for adults include modifying the dysfluent speech patterns (i.e., *stuttering modification* therapy), completely altering overall speech patterns to minimize the probability of stuttering (i.e., *fluency shaping* therapy), psychological interventions targeting anxiety, and even pharmacological intervention (for a detailed discussion of stuttering therapy for both children and adults, see both Bloodstein & Bernstein Ratner, 2008 and Guitar, 2019). All of these approaches aim to minimize the overt speech symptoms of the disorder, yet most have a high rate of relapse and often do not significantly improve fluency or quality of life for IWS (Craig & Calver, 1991; Craig & Hancock, 1995; O'Brian et al., 2003). In fact, a recent meta-analysis showed that no stuttering intervention for adults reduced the frequency of dysfluencies to typical levels (Brignell et al., 2020).

Even as early as the 1930s, it has been emphasized that therapy should involve more than just the speech symptoms themselves (albeit often for the wrong reasons, e.g., Goldiamond, 1965; Solomon, 1939). A therapeutic approach that also addresses psychological impact is indeed warranted for such a complex disorder, particularly in light of the documented mental health ramifications of stuttering (see *Symptomatology* above). Following the nascent movement to celebrate neurodiversity, recent publications have even emphasized the importance of moving away from therapy that focuses on reducing the frequency of stuttering moments at all, inherently pathologizing a simple variation in speech behavior that exists along a spectrum (Constantino et al., 2022; Tichenor et al., 2022; Tichenor & Yaruss, 2019).

> *"…to teach the stutterer to talk without stuttering [is] basically repressive in nature, doing little to relieve the emotional and relationship problems… such methods are unwise, for not only are their results short-lived, but subsequent failure engenders feelings of guilt, frustration and helplessness in the stutterer."*
>
> Andrews & Harris (1964),
> discussing Van Riper's approach of stuttering modification therapy

---

Many individuals who stutter may hold beliefs that fall somewhere in the middle: not wanting to completely dissociate themselves from their stuttering identity, but also not being content with the current treatment options. Still, some are desperately searching for a cure. There is thus an obvious need for novel therapies that are more successful over longer periods and across dimensions (i.e., actively involving the stuttering community in treatment development). Based on findings of reduced cortical activity during speech production in IWS, research into the use of neurostimulation to supplement behavioral therapeutic intervention is currently ongoing (e.g., Chesters et al., 2018; Garnett et al., 2019; Moein et al., 2022). However, the neural mechanisms associated with the core symptoms themselves are still poorly understood. Thus, the long-term goal of neurophysiological research in stuttering is to improve therapeutic options by directly targeting the neurological deficit itself or bolstering a more fluent network for speech production.

6. References

Alm, P. A. (2004). Stuttering and the basal ganglia circuits: A critical review of possible relations. *Journal of Communication Disorders*, *37*(4), 325–369. https://doi.org/10.1016/j.jcomdis.2004.03.001

Anderson, T. K., & Felsenfeld, S. (2003). A thematic analysis of late recovery from stuttering. *American Journal of Speech-Language Pathology*, *12*(2), 243–253. https://doi.org/10.1044/1058-0360(2003/070)

Andrews, G., Howie, P. M., Dozsa, M., & Guitar, B. E. (1982). Stuttering: Speech pattern characteristics under fluency-inducing conditions. *Journal of Speech and Hearing Research*, 25, 208–216. https://doi.org/10.1044/jshr.2502.208

Andrews, G., Craig, A., Feyer, A., Hoddinott, S., Howie, P., & Neilson, M. (1983). Stuttering: A review of research findings and theories circa 1982. *Journal of Speech and Hearing Disorders*, *48*(3), 226–246. https://doi.org/10.1044/JSHD.4803.226

Andrews, G., & Harris, M. (1964). *The syndrome of stuttering*. The Spastics Society Medical Education and Information Unit in association with William Heinemann, Medical Books Ltd.

Barnes, T. D., Wozniak, D. F., Gutierrez, J., Han, T. U., Drayna, D., & Holy, T. E. (2016). A mutation associated with stuttering alters mouse pup ultrasonic vocalizations. *Current Biology*, *26*(8), 1009–1018. https://doi.org/10.1016/j.cub.2016.02.068

Belyk, M., Kraft, S. J., & Brown, S. (2017). Stuttering as a trait or a state revisited: Motor system involvement in persistent developmental stuttering. *European Journal of Neuroscience*, *45*(4), 622–624. https://doi.org/10.1111/EJN.13512

Blood, G. W., Blood, I. M., Michael Tramontana, G., Sylvia, A. J., Boyle, M. P., & Motzko, G. R. (2011). Self-reported experience of bullying of students who stutter: Relations with life satisfaction, life orientation, and self-esteem. *Perceptual and Motor Skills*, *113*(2), 353–364. https://doi.org/10.2466/07.10.15.17.PMS.113.5.353-364

Blood, G. W., Blood, I. M., Tellis, G., & Gabel, R. (2001). Communication apprehension and self-perceived communication competence in adolescents who stutter. *Journal of Fluency Disorders*, *26*(3), 161–178. https://doi.org/10.1016/S0094-730X(01)00097-3

Bloodstein, O. (1958). Stuttering as an anticipatory struggle reaction. In J. Eisenson (Ed.), *Stuttering: A Symposium* (pp. 3–69). Harper & Row.

Bloodstein, O., & Bernstein Ratner, N. (2008). *A Handbook on Stuttering*. Thomson Delmar Learning.

Brignell, A., Krahe, M., Downes, M., Kefalianos, E., Reilly, S., & Morgan, A. T. (2020). A systematic review of interventions for adults who stutter. *Journal of Fluency Disorders*, *64*, 105766. https://doi.org/10.1016/J.JFLUDIS.2020.105766

Cai, S., Beal, D. S., Ghosh, S. S., Guenther, F. H., & Perkell, J. S. (2014). Impaired timing adjustments in response to time-varying auditory perturbation during connected speech production in persons who stutter. *Brain and Language*, *129*(1), 24–29. https://doi.org/10.1016/j.bandl.2014.01.002

Cai, S., Beal, D. S., Ghosh, S. S., Tiede, M. K., Guenther, F. H., & Perkell, J. S. (2012). Weak responses to auditory feedback perturbation during articulation in persons who stutter: Evidence for abnormal auditory-motor transformation. *PLoS ONE*, *7*(7), 1–13. https://doi.org/10.1371/journal.pone.0041830

Caruso, A. J., Abbs, J. H., & Gracco, V. L. (1988). Kinematic analysis of multiple movement coordination during speech in stutterers. *Brain*, *111*, 439–455. https://doi.org/10.1093/brain/111.2.439

Chang, S.-E., Erickson, K. I., Ambrose, N. G., Hasegawa-Johnson, M. A., & Ludlow, C. L. (2008). Brain anatomy differences in childhood stuttering. *NeuroImage*, *39*(3), 1333–1344. https://doi.org/10.1016/j.neuroimage.2007.09.067

Chang, S.-E., Garnett, E. O., Etchell, A., & Chow, H. M. (2019). Functional and neuroanatomical bases of developmental stuttering: current insights. *The Neuroscientist*, *25*(6), 566–582. https://doi.org/10.1177/1073858418803594

Chang, S.-E., Zhu, D. C., Choo, A. L., & Angstadt, M. (2015). White matter neuroanatomical differences in young children who stutter. *Brain*, *138*(3), 694–711. https://doi.org/10.1093/brain/awu400

Chesters, J., Möttönen, R., & Watkins, K. E. (2018). Transcranial direct current stimulation over left inferior frontal cortex improves speech fluency in adults who stutter. *Brain*, *141*(4), 1161–1171. https://doi.org/10.1093/brain/awy011

Chow, H. M., Garnett, E. O., Li, H., Etchell, A., Sepulcre, J., Drayna, D., Chugani, D., & Chang, S.-E. (2020). Linking lysosomal enzyme targeting genes and energy metabolism with altered gray matter volume in children with persistent stuttering. *Neurobiology of Language*, *1*(3), 365–380. https://doi.org/10.1162/NOL_A_00017

Civier, O., Tasko, S. M., & Guenther, F. H. (2010). Overreliance on auditory feedback may lead to sound/syllable repetitions: Simulations of stuttering and fluency-inducing conditions with a neural model of speech production. *Journal of Fluency Disorders*, *35*(3), 246–279. https://doi.org/10.1016/j.jfludis.2010.05.002

Constantino, C., Campbell, P., & Simpson, S. (2022). Stuttering and the social model. *Journal of Communication Disorders*, 106200. https://doi.org/10.1016/J.JCOMDIS.2022.106200

Conture, E. G. (1990). Childhood stuttering: What is it and who does it? In J. Cooper (Ed.), *Research Needs in Stuttering: Roadblocks and Future Directions* (Vol. 18, pp. 2–14). American Speech-Language-Hearing Association.

Cooper, M. H., & Allen, G. D. (1977). Timing control accuracy in normal speakers and stutterers. *Journal of Speech and Hearing Research*, *20*(1), 55–71. https://doi.org/10.1044/JSHR.2001.55

Craig, A., Blumgart, E., & Tran, Y. (2009). The impact of stuttering on the quality of life in adults who stutter. *Journal of Fluency Disorders*, *34*(2), 61–71. https://doi.org/10.1016/j.jfludis.2009.05.002

Craig, A., & Calver, P. (1991). Following up on treated stutterers: Studies of perceptions of fluency and job status. *Journal of Speech and Hearing Research*, *34*(2), 279–284. https://doi.org/10.1044/jshr.3402.279

Craig, A., & Hancock, K. (1995). Self-reported factors related to relapse following treatment for stuttering. *Australian Journal of Human Communication Disorders*, *23*(1), 48–60. https://doi.org/10.3109/asl2.1995.23.issue-1.04

Crichton-Smith, I. (2002). Communicating in the real world: Accounts from people who stammer. *Journal of Fluency Disorders*, *27*(4), 333–352. https://doi.org/10.1016/S0094-730X(02)00161-4

Curlee, R. F., & Yairi, E. (1997). Early intervention with early childhood stuttering. *American Journal of Speech-Language Pathology*, *6*(2), 8–18. https://doi.org/10.1044/1058-0360.0602.08

Daliri, A., & Max, L. (2015a). Electrophysiological evidence for a general auditory prediction deficit in adults who stutter. *Brain and Language*, *150*, 37–44. https://doi.org/10.1016/j.bandl.2015.08.008

Daliri, A., & Max, L. (2015b). Modulation of auditory processing during speech movement planning is limited in adults who stutter. *Brain and Language*, *143*, 59–68. https://doi.org/10.1016/j.bandl.2015.03.002

Daliri, A., & Max, L. (2018). Stuttering adults' lack of pre-speech auditory modulation normalizes when speaking with delayed auditory feedback. *Cortex*, *99*, 55–68. https://doi.org/10.1016/j.cortex.2017.10.019

Daliri, A., Prokopenko, R. A., & Max, L. (2013). Afferent and efferent aspects of mandibular sensorimotor control in adults who stutter. *Journal of Speech, Language, and Hearing Research*, *56*(6), 1774–1788. https://doi.org/10.1044/1092-4388(2013/12-0134)

Daliri, A., Wieland, E. A., Cai, S., Guenther, F. H., & Chang, S.-E. (2018). Auditory-motor adaptation is reduced in adults who stutter but not in children who stutter. *Developmental Science*, *21*(2), e12521. https://doi.org/10.1111/desc.12521

Davis, S., Howell, P., & Cooke, F. (2002). Sociodynamic relationships between children who stutter and their non-stuttering classmates. *Journal of Child Psychology and Psychiatry*, *43*(7), 939–947. https://doi.org/10.1111/1469-7610.00093

De Nil, L. F., Kroll, R. M., & Houle, S. (2001). Functional neuroimaging of cerebellar activation during single word reading and verb generation in stuttering and nonstuttering adults. *Neuroscience Letters*, *302*(2–3), 77–80. https://doi.org/10.1016/S0304-3940(01)01671-8

Etchell, A. C., Johnson, B. W., & Sowman, P. F. (2014). Behavioral and multimodal neuroimaging evidence for a deficit in brain timing networks in stuttering: A hypothesis and theory. *Frontiers in Human Neuroscience*, *8*, 467. https://doi.org/10.3389/fnhum.2014.00467

Fox, P. T., Ingham, R. J., Ingham, J. C., Hirsch, T. B., Downs, J. H., Martin, C., Jerabek, P., Glass, T., & Lancaster, J. L. (1996). A PET study of the neural systems of stuttering. *Nature*, *382*(6587), 158–162. https://doi.org/10.1038/382158a0

Fukawa, T., Yoshioka, H., Ozawa, E., & Yoshida, S. (1988). Difference of susceptibility to delayed auditory feedback between stutterers and nonstutterers. *Journal of Speech and Hearing Research*, *31*(3), 475–479. https://doi.org/10.1044/jshr.3103.475

Garnett, E. O., Chow, H. M., Choo, A. L., & Chang, S.-E. (2019). Stuttering severity modulates effects of non-invasive brain stimulation in adults who stutter. *Frontiers in Human Neuroscience*, *13*, 411. https://doi.org/10.3389/fnhum.2019.00411

Garnett, E. O., Chow, H. M., Limb, S., Liu, Y., & Chang, S.-E. (2022). Neural activity during solo and choral reading: A functional magnetic resonance imaging study of overt continuous speech production in adults who stutter. *Frontiers in Human Neuroscience*, *16*, 894676. https://doi.org/10.3389/fnhum.2022.894676

Goldiamond, I. (1965). Stuttering and fluency as manipulatable operable response classes. In L. Krasner & L. P. Ullmann (Eds.), *Research in Behavior Modification: New Developments and Implications* (pp. 106–156). Holt, Rinehart and Winston, Inc.

Gracco, V. L., Sares, A. G., & Koirala, N. (2022). Structural brain network topological alterations in stuttering adults. *Brain Communications*. https://doi.org/10.1093/BRAINCOMMS/FCAC058

Guitar, B. (2019). *Stuttering: An Integrated Approach to Its Nature and Treatment*. Lippincott Williams & Wilkins.

Han, T. U., Root, J., Reyes, L. D., Huchinson, E. B., Du Hoffmann, J., Lee, W. S., Barnes, T. D., & Drayna, D. (2019). Human GNPTAB stuttering mutations engineered into mice cause vocalization deficits and astrocyte pathology in the corpus callosum. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(35), 17515–17524. https://doi.org/10.1073/pnas.1901480116

Hickok, G., Houde, J., & Rong, F. (2011). Sensorimotor integration in speech processing: Computational basis and neural organization. *Neuron*, *69*(3), 407–422. https://doi.org/10.1016/j.neuron.2011.01.019

Hilger, A. I., Zelaznik, H., & Smith, A. (2016). Evidence that bimanual motor timing performance is not a significant factor in developmental stuttering. *Journal of Speech, Language, and Hearing Research, 59*(4), 674–685. https://doi.org/10.1044/2016_JSLHR-S-15-0172

Howell, P., El-Yaniv, N., & Powell, D. J. (1987). Factors affecting fluency in stutterers when speaking under altered auditory feedback. In H. F. M. Peters & W. Hulstijn (Eds.), *Speech Motor Dynamics in Stuttering* (pp. 361–369).

Ingham, R. J., & Cordes, A. K. (1998). Treatment decisions for young children who stutter: further concerns and complexities. *American Journal of Speech-Language Pathology*, *7*(3), 10–19. https://doi.org/10.1044/1058-0360.0703.10

Iverach, L., O'Brian, S., Jones, M., Block, S., Lincoln, M., Harrison, E., Hewat, S., Menzies, R. G., Packman, A., & Onslow, M. (2009). Prevalence of anxiety disorders among adults seeking speech therapy for stuttering. *Journal of Anxiety Disorders*, *23*(7), 928–934. https://doi.org/10.1016/j.janxdis.2009.06.003

Jackson, E. S., Yaruss, J. S., Quesal, R. W., Terranova, V., & Whalen, D. H. (2015). Responses of adults who stutter to the anticipation of stuttering. *Journal of Fluency Disorders*, *45*, 38–51. https://doi.org/10.1016/J.JFLUDIS.2015.05.002

Johnson, W. (1942). A study of the onset and development of stuttering. *Journal of Speech Disorders*, 251–257. https://doi.org/10.1044/jshd.0703.251

Johnson, W. (1967). *Speech handicapped school children* (W. Johnson & D. Moeller, Eds.). Harper & Row.

Johnson, W., & Rosen, L. (1937). Studies in the psychology of stuttering: VII. Effect of certain changes in speech pattern upon frequency of stuttering. *Journal of Speech Disorders*, 2(2), 105–110. https://doi.org/10.1044/jshd.0202.105

Jones, M., Onslow, M., Harrison, E., & Packman, A. (2000). Treating stuttering in young children: Predicting treatment time in the Lidcombe Program. *Journal of Speech, Language, and Hearing Research*, *43*(6), 1440–1450. https://doi.org/10.1044/JSLHR.4306.1440

Jones, M., Onslow, M., Packman, A., Williams, S., Ormond, T., Schwarz, I., & Gebski, V. (2005). Randomised controlled trial of the Lidcombe programme of early stuttering intervention. *BMJ*, *331*(7518), 659. https://doi.org/10.1136/BMJ.38520.451840.E0

Kang, C., Riazuddin, S., Mundorff, J., Krasnewich, D., Friedman, P., Mullikin, J. C., & Drayna, D. (2010). Mutations in the lysosomal enzyme–targeting pathway and persistent stuttering. *The New England Journal of Medicine*, *362*(8), 677–685. https://doi.org/10.1056/NEJMOA0902630

Kim, K. S., Daliri, A., Flanagan, J. R., & Max, L. (2020). Dissociated development of speech and limb sensorimotor learning in stuttering: Speech auditory-motor learning is impaired in both children and adults who stutter. *Neuroscience*, *451*, 1–21. https://doi.org/10.1016/j.neuroscience.2020.10.014

Kim, K. S., & Max, L. (2021). Speech auditory-motor adaptation to formant-shifted feedback lacks an explicit component: Reduced adaptation in adults who stutter reflects limitations in implicit sensorimotor learning. *European Journal of Neuroscience*, *53*(9), 3093–3108. https://doi.org/10.1111/ejn.15175

Klein, J. F., & Hood, S. B. (2004). The impact of stuttering on employment opportunities and job performance. *Journal of Fluency Disorders*, *29*(4), 255–273. https://doi.org/10.1016/J.JFLUDIS.2004.08.001

Kraaimaat, F. W., Vanryckeghem, M., & Van Dam-Baggen, R. (2002). Stuttering and social anxiety. *Journal of Fluency Disorders*, *27*(4), 319–331. https://doi.org/10.1016/S0094-730X(02)00160-2

Krakauer, J. W., Hadjiosif, A. M., Xu, J., Wong, A. L., & Haith, A. M. (2019). Motor Learning. In *Comprehensive Physiology* (Vol. 9, pp. 613–663). https://doi.org/10.1002/cphy.c170043

Krakauer, J. W., & Mazzoni, P. (2011). Human sensorimotor learning: Adaptation, skill, and beyond. *Current Opinion in Neurobiology*, *21*(4), 636–644. https://doi.org/10.1016/J.CONB.2011.06.012

Li, J.J., Daliri, A., Kim, K. S., & Max, L. (2024). Does pre-speech auditory modulation reflect processes related to feedback monitoring or speech movement planning? *Neuroscience Letters*. https://doi.org/10.1016/j.neulet.2024.138025

Lincoln, M. A., & Onslow, M. (1997). Long-term outcome of early intervention for stuttering. *American Journal of Speech-Language Pathology*, *6*(1), 51–58. https://doi.org/10.1044/1058-0360.0601.51

Lincoln, M. A., Onslow, M., & Reed, V. (1997). Social validity of the treatment outcomes of an early intervention program for stuttering. *American Journal of Speech-Language Pathology*, *6*(2), 77–84. https://doi.org/10.1044/1058-0360.0602.77

Linn, G. W., & Caruso, A. J. (1998). Perspectives on the effects of stuttering on the formation and maintenance of intimate relationships. *Journal of Rehabilitation*, *64*(3), 12–15.

Loucks, T. M., Chon, H., & Han, W. (2012). Audiovocal integration in adults who stutter. *International Journal of Language & Communication Disorders*, *47*(4), 451–456. https://doi.org/10.1111/J.1460-6984.2011.00111.X

Mackay, D. G., & Macdonald, M. C. (1984). Stuttering as a sequencing and timing disorder. In W. H. Perkins & R. Curlee (Eds.), *Nature and treatment of stuttering: New directions* (pp. 261–282). College Hill Press.

Masapollo, M., Segawa, J. A., Beal, D. S., Tourville, J. A., Nieto-Castañón, A., Heyne, M., Frankford, S. A., & Guenther, F. H. (2021). Behavioral and neural correlates of speech motor sequence learning in stuttering and neurotypical speakers: An fMRI Investigation. *Neurobiology of Language*, *2*(1), 106–137. https://doi.org/10.1162/nol_a_00027

Max, L. (2004). Stuttering and internal models for sensorimotor control: A theoretical perspective to generate testable hypotheses. In B. Maassen, R. Kent, H. Peters, P. van Lieshout, & W. Hulstijn (Eds.), *Speech Motor Control in Normal and Disordered Speech* (pp. 357–387). Oxford University Press.

Max, L., Caruso, A. J., & Gracco, V. L. (2003). Kinematic analyses of speech, orofacial nonspeech, and finger movements in stuttering and nonstuttering adults. *Journal of Speech, Language, and Hearing Research*, *46*(1), 215–232. https://doi.org/10.1044/1092-4388(2003/017)

Max, L., & Daliri, A. (2019). Limited pre-speech auditory modulation in individuals who stutter: Data and hypotheses. *Journal of Speech, Language, and Hearing Research*, *62*(8S), 3071–3084. https://doi.org/10.1044/2019_JSLHR-S-CSMC7-18-0358

Max, L., Guenther, F. H., Gracco, V. L., Ghosh, S. S., & Wallace, M. E. (2004). Unstable or insufficiently activated internal models and feedback-biased motor control as sources of dysfluency: A theoretical model of stuttering. *Contemporary Issues in Communication Science and Disorders*, *31*, 105–122.

Max, L., & Yudman, E. M. (2003). Accuracy and variability of isochronous rhythmic timing across motor systems in stuttering versus nonstuttering individuals. *Journal of Speech, Language, and Hearing Research*, *46*(1), 146–163. https://doi.org/10.1044/1092-4388(2003/012)

Moein, N., Mohamadi, R., Rostami, R., Nitsche, M., Zomorrodi, R., & Ostadi, A. (2022). Investigation of the effect of delayed auditory feedback and transcranial direct current stimulation (DAF-tDCS) treatment for the enhancement of speech fluency in adults who stutter: A randomized controlled trial. *Journal of Fluency Disorders*, *72*, 105907. https://doi.org/10.1016/J.JFLUDIS.2022.105907

Molt, L. F., & Guilford, A. M. (1979). Auditory processing and anxiety in stutterers. *Journal of Fluency Disorders*, *4*(4), 255–267. https://doi.org/10.1016/0094-730X(79)90002-0

Neef, N. E., Angstadt, M., Koenraads, S. P. C., & Chang, S.-E. (2022). Dissecting structural connectivity of the left and right inferior frontal cortex in children who stutter. *Cerebral Cortex*, bhac328. https://doi.org/10.1093/cercor/bhac328

Neef, N. E., Anwander, A., & Friederici, A. D. (2015). The neurobiological grounding of persistent stuttering: From structure to function. *Current Neurology and Neuroscience Reports*, *15*(9). https://doi.org/10.1007/s11910-015-0579-4

Neef, N. E., & Chang, S.-E. (2024). Knowns and unknowns about the neurobiology of stuttering. *PLOS Biology*. https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.3002492

O'Brian, S., Iverach, L., Jones, M., Onslow, M., Packman, A., & Menzies, R. (2013). Effectiveness of the Lidcombe Program for early stuttering in Australian community clinics. *International Journal of Speech-Language Pathology*, *15*(6), 593–603. https://doi.org/10.3109/17549507.2013.783112

O'Brian, S., Onslow, M., Cream, A., & Packman, A. (2003). The Camperdown Program. *Journal of Speech, Language, and Hearing Research*, *46*(4), 933–946. https://doi.org/10.1044/1092-4388(2003/073)

Onslow, M., & Kelly, E. M. (2020). Temperament and early stuttering intervention: Two perspectives. *Journal of Fluency Disorders*, *64*, 105765. https://doi.org/10.1016/j.jfludis.2020.105765

Onslow, M., Menzies, R. G., & Packman, A. (2001). An operant intervention for early stuttering. The development of the Lidcombe program. *Behavior Modification*, *25*(1), 116–139. https://doi.org/10.1177/0145445501251007

Onslow, M., & Millard, S. (2012). Palin Parent Child Interaction and the Lidcombe Program: Clarifying some issues. *Journal of Fluency Disorders*, *37*(1), 1–8. https://doi.org/10.1016/j.jfludis.2011.10.002

Onslow, M., & Yaruss, J. S. (2007). Differing perspectives on what to do with a stuttering preschooler and why. *American Journal of Speech-Language Pathology*, *16*(1), 65–68. https://doi.org/10.1044/1058-0360(2007/008)

Plexico, L. W., Manning, W. H., & Levitt, H. (2009). Coping responses by adults who stutter: Part I. Protecting the self and others. *Journal of Fluency Disorders*, *34*(2), 87–107. https://doi.org/10.1016/J.JFLUDIS.2009.06.001

Riaz, N., Steinberg, S., Ahmad, J., Pluzhnikov, A., Riazuddin, S., Cox, N. J., & Drayna, D. (2005). Genomewide significant linkage to stuttering on chromosome 12. *The American Journal of Human Genetics*, *76*(4), 647–651. https://doi.org/10.1086/429226

Rieber, R. W., & Wollock, J. (1977). The historical roots of the theory and therapy of stuttering. *Journal of Communication Disorders*, *10*(1–2), 3–24. https://doi.org/10.1016/0021-9924(77)90009-0

Seider, R. A., Kidd, K. K., & Gladstien, K. L. (1983). recovery and persistence of stuttering among relatives of stutterers. *Journal of Speech and Hearing Disorders*, *48*(4), 402–409. https://doi.org/10.1044/jshd.4804.402

Sengupta, R., Shah, S., Gore, K., Loucks, T. M., & Nasir, S. M. (2016). Anomaly in neural phase coherence accompanies reduced sensorimotor integration in adults who stutter. *Neuropsychologia*, *93*, 242–250. https://doi.org/10.1016/j.neuropsychologia.2016.11.004

Sengupta, R., Yaruss, J. S., Loucks, T. M., Gracco, V. L., Pelczarski, K., & Nasir, S. M. (2019). Theta modulated neural phase coherence facilitates speech fluency in adults who stutter. *Frontiers in Human Neuroscience*, *13*, 394. https://doi.org/10.3389/fnhum.2019.00394

Shadmehr, R., Smith, M. A., & Krakauer, J. W. (2010). Error correction, sensory prediction, and adaptation in motor control. *Annual Review of Neuroscience*, *33*(1), 89–108. https://doi.org/10.1146/annurev-neuro-060909-153135

Shaw, D. M., Polikowsky, H. P., Pruett, D. G., Chen, H.-H., Petty, L. E., Viljoen, K. Z., Beilby, J. M., Jones, R. M., Kraft, S. J., & Below, J. E. (2021). Phenome risk classification enables phenotypic imputation and gene discovery in developmental stuttering. *The American Journal of Human Genetics*, *108*(12), 2271–2283. https://doi.org/10.1016/J.AJHG.2021.11.004

Shiller, D. M., Mitsuya, T., & Max, L. (2020). Exposure to auditory feedback delay while speaking induces perceptual habituation but does not mitigate the disruptive effect of delay on speech auditory-motor learning. *Neuroscience*. https://doi.org/10.1016/j.neuroscience.2020.07.041

Solomon, M. (1939). Stuttering as an emotional and personality disorder. *Journal of Speech Disorders*, *4*, 347–357. https://doi.org/10.1044/jshd.0404.347

Strother, C. R., & Kriegman, L. S. (1943). Diadochokinesis in stutterers and non-stutterers. *Journal of Speech Disorders*, *8*, 323–335. https://doi.org/10.1044/jshd.0804.323

Tichenor, S. E., Constantino, C., & Yaruss, J. S. (2022). A point of view about fluency. *Journal of Speech, Language, and Hearing Research*, *13*(7), 1–8. https://doi.org/10.1044/2021_JSLHR-21-00342

Tichenor, S. E., Palasik, S., & Yaruss, J. S. (2023). Understanding the broader impact of stuttering: suicidal ideation. *American Journal of Speech-Language Pathology*, 1–24. https://doi.org/10.1044/2023_AJSLP-23-00007

Tichenor, S. E., & Yaruss, J. S. (2019). Stuttering as defined by adults who stutter. *Journal of Speech, Language, and Hearing Research*, *62*(12), 4356–4369. https://doi.org/10.1044/2019_JSLHR-19-00137

Van Borsel, J., Brepoels, M., & De Coene, J. (2011). Stuttering, attractiveness and romantic relationships: The perception of adolescents and young adults. *Journal of Fluency Disorders*, *36*(1), 41–50. https://doi.org/10.1016/j.jfludis.2011.01.002

Van Riper, C. (1971). *The nature of stuttering*. Prentice-Hall.

Vanryckeghem, M., Brutten, G. J., & Hernandez, L. M. (2005). A comparative investigation of the speech-associated attitude of preschool and kindergarten children who do and do not stutter. *Journal of Fluency Disorders*, *30*(4), 307–318. https://doi.org/10.1016/j.jfludis.2005.09.003

Vanryckeghem, M., Brutten, G. J., Uddin, N., & Borsel, J. V. (2004). A comparative investigation of the speech-associated coping responses reported by adults who do and do not stutter. *Journal of Fluency Disorders*, *29*(3), 237–250. https://doi.org/10.1016/J.JFLUDIS.2004.07.001

Williams, D. E. (1957). A point of view about "stuttering." *Journal of Speech and Hearing Disorders*, *22*(3), 390–397. https://doi.org/10.1044/jshd.2203.390

Williams, D. E., Melrose, B. M., & Woods, C. L. (1969). The relationship between stuttering and academic achievement in children. *Journal of Communication Disorders*, *2*(2), 87–98. https://doi.org/10.1016/0021-9924(69)90032-x

Wingate, M. E. (2002). *Foundations of stuttering*. Academic Press.

Yairi, E., & Ambrose, N. (1992a). A longitudinal study of stuttering in children: A preliminary report. *Journal of Speech and Hearing Research*, *35*(4), 755–760. https://doi.org/10.1044/JSHR.3504.755

Yairi, E., & Ambrose, N. (1992b). Onset of stuttering in preschool children: Selected factors. *Journal of Speech and Hearing Research*, *35*(4), 782–788. https://doi.org/10.1044/JSHR.3504.782

Yairi, E., & Ambrose, N. (2013). Epidemiology of stuttering: 21st century advances. *Journal of Fluency Disorders*, *38*(2), 66–87. https://doi.org/10.1016/j.jfludis.2012.11.002

Yairi, E., & Ambrose, N. G. (1999). Early childhood stuttering I: Persistency and recovery rates. *Journal of Speech, Language, and Hearing Research*, *42*(5), 1097–1112. https://doi.org/10.1044/JSLHR.4205.1097

Yairi, E., Ambrose, N. G., & Niermann, R. (1993). The early months of stuttering: A developmental study. *Journal of Speech and Hearing Research*, *36*(3), 521–528. https://doi.org/10.1044/JSHR.3603.521

Section II: General Introduction

1. Framing the dissertation

In sum, it is evident that stuttering is a significant global health issue with serious individual and societal consequences. To date, the neural mechanisms associated with both the disorder itself and its core symptoms are still not fully understood, despite evidence of speech and nonspeech sensorimotor differences and extensive sensorimotor network involvement (Chang et al., 2015, 2019; Garnett et al., 2022; Kim et al., 2020; Max, 2004; Max et al., 2003, 2004; Neef et al., 2015, 2022). This dissertation presents novel results from three electrophysiological experiments that sought to bolster our understanding of the sensorimotor mechanisms in stuttering by extending our work on a previously established paradigm.

It is well known that the integration of motor and sensory processes is critical for the successful neural control of movement (Shadmehr, 2017; Shadmehr et al., 2010), yet physiological studies investigating the interactions between efferent and afferent neural systems in speech have been sparse. Moreover, despite the fact that stuttering usually occurs at the beginning of a word, such studies have been limited to sensorimotor processing *during* the act of speaking rather than during the preceding planning phase (Beal et al., 2010). To address this, our lab has utilized a unique experimental paradigm to investigate sensory predictions and sensorimotor interactions *prior* to speech production (i.e., during speech movement planning). Using this paradigm, we have discovered that typical speakers modulate auditory processing just prior to the initiation of speech movements (Daliri & Max, 2015a, 2015b, 2016; Max et al., 2008). Furthermore, we found that this modulation of auditory processing during speech movement planning is limited in individuals who stutter, leading our team to suggest an initial hypothesis that stuttering may involve insufficient sensory priming that triggers inappropriate attempts at error correction (Daliri & Max, 2015a, 2015b, 2016, 2018; Max & Daliri, 2019). However, recent work in our laboratory has since led us to a new working hypothesis for the reduced amount of pre-speech auditory modulation (PSAM) observed in stuttering adults, a hypothesis which will be expanded upon in subsequent sections of this dissertation.

Overall, our findings on PSAM are highly robust, and have been replicated in several studies across multiple groups of subjects. Nevertheless, our investigations to date have only measured cortical auditory behavior just prior to the production of monosyllabic words, which are thought to rely mainly on feedforward control. Thus, to further investigate this pre-speech auditory phenomenon, the first experiment presented here (*Section III, Study 1*) involved extending our established paradigm to elicit the production of more complex sentence-level speech. Two matched groups of adult speakers were recruited to perform a delayed response reading task in which they produced either monosyllabic words (N = 15) or entire sentences (N = 15). As in our prior PSAM work, we recorded auditory-evoked potentials in response to probe tones presented just prior to the ‘go’ signal to speak and at an equivalent time during a silent reading control condition. This paradigm allowed us to begin investigating sensorimotor interactions during the movement planning phase for more complex speech that recruits a greater proportion of feedback-based control mechanisms. Ultimately, similar sentence-level paradigms can be used to begin investigating sensorimotor interactions in stuttering speakers just prior to stuttered speech itself.

Furthermore, our prior investigations have focused solely on the *auditory* ramifications of speech movement planning. Thus, a second experiment (*Section IV, Study 2*) aimed to investigate whether deficits in motor-to-sensory modulatory processes in adults who stutter are specific to the auditory domain, or whether they also compromise motor-to-*somatosensory* effects. To accomplish this aim, we recruited a group of adults who stutter (*N* = 15) and a group of age-, handedness-, and sex-matched nonstuttering speakers (*N* = 15). Specifically, we extended our standard delayed-response word reading task and our standard electrophysiological recording procedure to record pre-speech *somatosensory*-evoked potentials (SEPs). Instead of presenting participants with a probe tone during the movement planning phase, we experimentally applied a rapid orofacial skin stretch to elicit SEPs, based on previous work done by Ito and colleagues (Ito, Bai, et al., 2020; Ito et al., 2009, 2015; Ito, Ohashi, et al., 2020; Ito & Ostry, 2010; Ohashi & Ito, 2019). This study allowed us to begin investigating the sensory domain specificity of motor-to-sensory modulatory processes in stuttering.

Finally, we still have an incomplete understanding of the functional relevance of this motor-to-sensory modulation. Thus, the final study herein (*Section V, Study 3*) aimed to further characterize the functional relevance of PSAM by investigating whether it varies with the extent to which speakers rely on feedback-based (vs. feedforward-based) motor control strategies. Previously, we have broadly documented how motor prediction affects auditory processing (Max & Daliri, 2019), allowing us to now ask if this modulatory influence on the auditory system is reflected in control strategies in the motor output. Speakers who stutter (*N* = 6 adults) were matched pairwise with speakers who do not stutter (*N* = 6 adults), allowing us to correlate measures across a larger range of PSAM values. Specifically, in addition to a PSAM task, participants also completed an electromagnetic articulography task in which they produced the same monosyllabic syllables in carrier phrases while we recorded the three-dimensional position of their articulators. The kinematic data was used to estimate each participant's reliance on feedforward vs. feedback-based control, following a statistical modeling procedure originally used for limb reaching movements but recently validated for speech movements as well (Kim &

Max, 2014). This experiment allowed us to start investigating how pre-speech auditory modulation relates to kinematic measures of speech production that reflect the implementation of predictive (rather than reactive) corrections.

The final section of this dissertation (*Section VI*) summarizes and integrates the results of these three experiments, situating the findings in the broader literature and discussing some avenues for future research.

2. References

Beal, D. S., Cheyne, D. O., Gracco, V. L., Quraan, M. A., Taylor, M. J., & De Nil, L. F. (2010). Auditory evoked fields to vocalization during passive listening and active generation in adults who stutter. *NeuroImage*, *52*(4), 1645–1653. https://doi.org/10.1016/j.neuroimage.2010.04.277

Brignell, A., Krahe, M., Downes, M., Kefalianos, E., Reilly, S., & Morgan, A. (2021). Interventions for children and adolescents who stutter: A systematic review, meta-analysis, and evidence map. *Journal of Fluency Disorders*, *70*, 1–21. https://doi.org/10.1016/J.JFLUDIS.2021.105843

Chang, S.-E., Garnett, E. O., Etchell, A., & Chow, H. M. (2019). Functional and neuroanatomical bases of developmental stuttering: current insights. *The Neuroscientist*, *25*(6), 566–582. https://doi.org/10.1177/1073858418803594

Chang, S.-E., Zhu, D. C., Choo, A. L., & Angstadt, M. (2015). White matter neuroanatomical differences in young children who stutter. *Brain*, *138*(3), 694–711. https://doi.org/10.1093/brain/awu400

Chesters, J., Möttönen, R., & Watkins, K. E. (2018). Transcranial direct current stimulation over left inferior frontal cortex improves speech fluency in adults who stutter. *Brain*, *141*(4), 1161–1171. https://doi.org/10.1093/brain/awy011

Craig, A., & Calver, P. (1991). Following up on treated stutterers: Studies of perceptions of fluency and job status. *Journal of Speech and Hearing Research*, *34*(2), 279–284. https://doi.org/10.1044/jshr.3402.279

Craig, A., & Hancock, K. (1995). Self-reported factors related to relapse following treatment for stuttering. *Australian Journal of Human Communication Disorders*, *23*(1), 48–60. https://doi.org/10.3109/asl2.1995.23.issue-1.04

Daliri, A., & Max, L. (2015a). Electrophysiological evidence for a general auditory prediction deficit in adults who stutter. *Brain and Language*, *150*, 37–44. https://doi.org/10.1016/j.bandl.2015.08.008

Daliri, A., & Max, L. (2015b). Modulation of auditory processing during speech movement planning is limited in adults who stutter. *Brain and Language*, *143*, 59–68. https://doi.org/10.1016/j.bandl.2015.03.002

Daliri, A., & Max, L. (2016). Modulation of auditory responses to speech vs. nonspeech stimuli during speech movement planning. *Frontiers in Human Neuroscience*, *10*, 1–9. https://doi.org/10.3389/fnhum.2016.00234

Daliri, A., & Max, L. (2018). Stuttering adults' lack of pre-speech auditory modulation normalizes when speaking with delayed auditory feedback. *Cortex*, *99*, 55–68. https://doi.org/10.1016/j.cortex.2017.10.019

Garnett, E. O., Chow, H. M., Choo, A. L., & Chang, S.-E. (2019). Stuttering severity modulates effects of non-invasive brain stimulation in adults who stutter. *Frontiers in Human Neuroscience*, *13*, 411. https://doi.org/10.3389/fnhum.2019.00411

Garnett, E. O., Chow, H. M., Limb, S., Liu, Y., & Chang, S.-E. (2022). Neural activity during solo and choral reading: A functional magnetic resonance imaging study of overt continuous speech production in adults who stutter. *Frontiers in Human Neuroscience*, *16*, 894676. https://doi.org/10.3389/fnhum.2022.894676

Guitar, B. (2019). *Stuttering: An Integrated Approach to Its Nature and Treatment*. Lippincott Williams & Wilkins.

Kim, K. S., & Max, L. (2014). Estimating feedforward vs. feedback control of speech production through kinematic analyses of unperturbed articulatory movements. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00911

Kim, K. S., Daliri, A., Flanagan, J. R., & Max, L. (2020). Dissociated development of speech and limb sensorimotor learning in stuttering: Speech auditory-motor learning is impaired in both children and adults who stutter. *Neuroscience*, *451*, 1–21. https://doi.org/10.1016/j.neuroscience.2020.10.014

Ito, T., Bai, J., & Ostry, D. J. (2020). Contribution of sensory memory to speech motor learning. *Journal of Neurophysiology*, *124*(4), 1103–1109. https://doi.org/10.1152/JN.00457.2020

Ito, T., Ohashi, H., & Gracco, V. L. (2020). Changes of orofacial somatosensory attenuation during speech production. *Neuroscience Letters*, *730*, 135045. https://doi.org/10.1016/j.neulet.2020.135045

Ito, T., & Ostry, D. J. (2010). Somatosensory contribution to motor learning due to facial skin deformation. *Journal of Neurophysiology*, *104*(3), 1230–1238. https://doi.org/10.1152/jn.00199.2010

Ito, T., Ostry, D. J., & Gracco, V. L. (2015). Somatosensory event-related potentials from orofacial skin stretch stimulation. *Journal of Visualized Experiments*, *2015*(106), 53621. https://doi.org/10.3791/53621

Ito, T., Tiede, M., & Ostry, D. J. (2009). Somatosensory function in speech perception. *Proceedings of the National Academy of Sciences*, *106*(4), 1245–1248. https://doi.org/10.1073/pnas.0810063106

Jones, M., Onslow, M., Packman, A., Williams, S., Ormond, T., Schwarz, I., & Gebski, V. (2005). Randomised controlled trial of the Lidcombe programme of early stuttering intervention. *BMJ*, *331*(7518), 659. https://doi.org/10.1136/BMJ.38520.451840.E0

Laiho, A., Elovaara, H., Kaisamatti, K., Luhtalampi, K., Talaskivi, L., Pohja, S., Routamo-Jaatela, K., & Vuorio, E. (2022). Stuttering interventions for children, adolescents, and adults: A systematic review as a part of clinical guidelines. *Journal of Communication Disorders*, *99*, 106242. https://doi.org/10.1016/j.jcomdis.2022.106242

Max, L. (2004). Stuttering and internal models for sensorimotor control: A theoretical perspective to generate testable hypotheses. In B. Maassen, R. Kent, H. Peters, P. van Lieshout, & W. Hulstijn (Eds.), *Speech Motor Control in Normal and Disordered Speech* (pp. 357–387). Oxford University Press.

Max, L., Caruso, A. J., & Gracco, V. L. (2003). Kinematic analyses of speech, orofacial nonspeech, and finger movements in stuttering and nonstuttering adults. *Journal of Speech, Language, and Hearing Research*, *46*(1), 215–232. https://doi.org/10.1044/1092-4388(2003/017)

Max, L., & Daliri, A. (2019). Limited pre-speech auditory modulation in individuals who stutter: Data and hypotheses. *Journal of Speech, Language, and Hearing Research*, *62*(8S), 3071–3084. https://doi.org/10.1044/2019_JSLHR-S-CSMC7-18-0358

Max, L., Daniels, J. C., Curet, K. M., & Cronin, K. L. (2008). Modulation of auditory and somatosensory processing during the planning of speech movements. *Proceedings of ISSP 2008 - 8th International Seminar on Speech Production*, *2*, 41–44.

Max, L., Guenther, F. H., Gracco, V. L., Ghosh, S. S., & Wallace, M. E. (2004). Unstable or Insufficiently activated internal models and feedback-biased motor control as sources of dysfluency: A theoretical model of stuttering. *Contemporary Issues in Communication Science and Disorders*, *31*, 105–122.

Moein, N., Mohamadi, R., Rostami, R., Nitsche, M., Zomorrodi, R., & Ostadi, A. (2022). Investigation of the effect of delayed auditory feedback and transcranial direct current stimulation (DAF-tDCS) treatment for the enhancement of speech fluency in adults who stutter: A randomized controlled trial. *Journal of Fluency Disorders*, *72*, 105907. https://doi.org/10.1016/J.JFLUDIS.2022.105907

Neef, N. E., Angstadt, M., Koenraads, S. P. C., & Chang, S.-E. (2022). Dissecting structural connectivity of the left and right inferior frontal cortex in children who stutter. *Cerebral Cortex*, bhac328. https://doi.org/10.1093/cercor/bhac328

Neef, N. E., Anwander, A., & Friederici, A. D. (2015). The neurobiological grounding of persistent stuttering: From structure to function. *Current Neurology and Neuroscience Reports*, *15*(9). https://doi.org/10.1007/s11910-015-0579-4

O'Brian, S., Iverach, L., Jones, M., Onslow, M., Packman, A., & Menzies, R. (2013). Effectiveness of the Lidcombe Program for early stuttering in Australian community clinics. *International Journal of Speech-Language Pathology*, *15*(6), 593–603. https://doi.org/10.3109/17549507.2013.783112

Ohashi, H., & Ito, T. (2019). Recalibration of auditory perception of speech due to orofacial somatosensory inputs during speech motor adaptation. *Journal of Neurophysiology*, *122*(5), 2076–2084. https://doi.org/10.1152/jn.00028.2019

Shadmehr, R. (2017). Learning to predict and control the physics of our movements. *Journal of Neuroscience*, *37*(7), 1663–1671. https://doi.org/10.1523/JNEUROSCI.1675-16.2016

Shadmehr, R., Smith, M. A., & Krakauer, J. W. (2010). Error Correction, Sensory Prediction, and Adaptation in Motor Control. *Annual Review of Neuroscience*, *33*(1), 89–108. https://doi.org/10.1146/annurev-neuro-060909-15313

Section III: Pre-speech auditory modulation prior to sentence-level speech (*Study 1*)

(Manuscript formatted for submission to *Neuroscience Letters*)

Elise LeBovidge[a]
eliseal@uw.edu

Ludo Max[a]*
ludomax@uw.edu

[a]Department of Speech and Hearing Sciences, University of Washington,
1417 N.E. 42nd Street, Seattle, Washington, 98105-6246, USA

*corresponding author

1. ABSTRACT

Adult speakers have been shown to modulate their auditory system prior to the production of monosyllabic words. Specifically, when pure tone stimuli are presented just prior to speech onset, the N100 amplitude of the long latency auditory-evoked potentials is reduced as compared to a control condition without speech production. Here, we aimed to further investigate this phenomenon by extending our paradigm to elicit the production of more complex sentence-level speech, as previous research has only reported pre-speech auditory modulation (PSAM) prior to the production of monosyllabic words. Two matched groups of adult speakers were recruited to perform a delayed response reading task in which they produced either monosyllabic words ($N = 15$) or sentences ($N = 15$). Auditory-evoked potentials were recorded in response to probe tones randomly presented 200 ms prior to the 'go' signal to speak and at an equivalent time during a silent reading control condition. We found that the N1 amplitudes recorded prior to speech were significantly smaller than those recorded during the control condition, regardless of utterance length. Notably, the amplitude of pre-speech auditory modulation did not differ between *Sentences* and *Words* groups, suggesting that this paradigm could be used to investigate sensorimotor behavior in the context of more natural, connected speech. Future work investigating how PSAM relates to fluency status in participants who stutter, for example, could provide important insights into the neurodevelopmental speech disorder.

Keywords: auditory modulation; movement planning; sensory suppression; event-related potentials; speech; sentences

2. INTRODUCTION

The central nervous system is able to predict the sensory consequences of planned motor commands and use that information to modulate its processing of self-produced sensory stimulation [1–5]. More recently, it has been revealed that the central nervous system is already modulating the auditory system prior to speech production. Specifically, a series of experiments from our own laboratory have documented a consistent modulation of auditory processing mechanisms during speech movement planning, just prior to speech initiation [6–11]. When pure tone probe stimuli are presented during a short delay period prior to speech onset, the N1 amplitude of the long latency auditory-evoked potentials is reduced compared to those recorded during a silent control condition without speech production.

Initial hypotheses suggested that this pre-speech auditory modulation (PSAM) could be reflective of a general priming of the auditory system just prior to its role in monitoring auditory feedback during speech production [11]. However, recent psychophysical data from our laboratory does not support this hypothesis, as the formant discrimination of participants' own vowel productions was observed to be *worse* at the same timepoint prior to speech production [12]. Thus, the phenomenon of PSAM seems unlikely to reflect a selective tuning of auditory neurons in preparation for their role in feedback monitoring. Instead, PSAM seems to reflect a general *suppression* of auditory processing, as evidenced by participants' reduced capacity to detect small deviations in their own vowel formant frequencies at the same timepoint during speech movement planning.

Our current working hypothesis for the functional relevance of PSAM is based on a new finding that the amount of modulation is negatively associated with inter-trial formant variability as early as vowel onset [13]. Specifically, the more a speaker's cortical auditory response was reduced during speech movement planning, the less variable their produced formants tended to be – even as early as the initial onset of speech (i.e., prior to feedback-based corrections). There was also no clear correlation between the amount of modulation and within-vowel changes, suggesting that these underlying neural processes are not involved in feedback-based control as once hypothesized. It is therefore possible that this pre-speech cortical auditory modulation may instead occur due to bidirectional, *predictive* communication between premotor and sensory areas. In other words, if a subset of auditory cortex is involved in the premovement fine-tuning of motor commands, then fewer neuronal populations would be available to process the probe stimuli presented during movement planning. Kinematic work is ongoing in our laboratory to directly investigate whether PSAM is also associated with feedforward control mechanisms.

Importantly, our previous work has only recorded auditory-evoked potentials (AEPs) to probe tones presented during the planning phase before the production of individual monosyllabic words, which are thought to be predominantly under feedforward control. Therefore, the current study aimed to extend our typical paradigm to allow for the probing of cortical auditory activity just prior to the production of more natural, connected speech in which feedback-based control mechanisms are also recruited. We recruited two matched groups of typically fluent adult speakers to perform the same delayed response reading task, differing only in the length of linguistic stimuli they produced (*Words* vs. *Sentences* groups).

3. METHODOLOGY

3.1. Participants

A total of 30 typical adult speakers of American English were included in data analysis, recruited pseudo-randomly into two groups that differed only in the length of stimuli read during the experiment (*Words* vs. *Sentences* groups; *N* = 15 per group).

Participants were matched pairwise across groups based on self-reported age (± 3 years), handedness, and biological sex. In each group, there were 4 cisgender men. The remaining 11 participants in each group were all assigned female at birth, but two participants in the *Words* group and one participant in the *Sentences* group were nonbinary or genderqueer. All participants reported writing with their right hand.

The mean age of participants in the *Words* group was 26.5 years (*SD* = 6.8 years, ranging from 18 to 39 years). The mean age of participants in the *Sentences* group was 26.3 years (*SD* = 6.1 years, ranging from 19 to 39 years).

Two additional participants completed the experiment but were removed from analysis because clear AEPs could not be obtained from their recorded data. One participant was unable to stay awake and adhere to the task instructions. The other participant had an insufficient number of clean epochs. Two new participants were recruited to replace these participants in the dataset and maintain 15 participants per group.

Participants reported no speech, language, hearing, neurological, or psychological difficulties or diagnoses. No participants reported taking medications with sensorimotor side effects, and all reported minimal facial hair (to allow for orofacial electromyography). At the beginning of the experimental session, all participants underwent a brief hearing screening to verify pure tone air-conducted hearing thresholds of ≤ 25 dB HL for the octave frequencies from 250 to 4000 Hz.

A self-report questionnaire with the demographic categories used by the National Institutes of Health (NIH) showed that most participants were non-Hispanic/Latino and White (N = 14) or non-Hispanic/Latino and Asian (N = 6). Three further non-Hispanic/Latino participants reported being of Mixed ancestry. Six participants reported being Hispanic/Latino (three of which were White, one was Black, one was American Indian/Alaska Native, and the final of unknown race). The final participant only reported being Asian.

Participants provided informed, written consent before participating. The experimental procedures were approved by the University of Washington's Institutional Review Board. All participants were financially compensated $20 per hour for time spent in the laboratory (approx. 3 hours) and any additional street parking costs.

3.2. Experimental paradigm

The paradigm was extended from previously published studies investigating the phenomenon of pre-speech auditory modulation [6–9]. Participants were seated ~1.5 meters away from a 27" monitor (AW2723DF, Alienware; Miami, FL, USA) in a sound-attenuated booth and completed a delayed response reading task designed to elicit the production of either monosyllabic words or entire sentences. Electroencephalographic (EEG) data was simultaneously recorded to investigate the modulation of cortical auditory activity elicited by probe tones presented immediately prior to speaking (*Speaking* condition) compared to a control condition with no speech production (*No-Speaking* condition).

All participants completed three recording blocks of 90 trials each for both the *Speaking* condition and the *No-Speaking* condition. As in previous PSAM studies, we chose to block overt speaking and silent reading separately to minimize additional cognitive effort during the task (e.g., deciding whether to speak aloud on each trial). Block order was randomly assigned for each participant upon study recruitment. Participants thus completed 6 total blocks, with the duration of each block depending on group assignment (approximately 5.5 minutes for the *Words* group and 8.5 minutes for the *Sentences* group).

At the initialization of each block, 90 stimuli were randomly selected from a possible 273 sentences or 180 monosyllabic words. Sentences were adapted from excerpts about animals designed for school-aged children. Calculated Flesch-Kincaid grade levels confirmed that the stimuli were readable for most Americans, as the sentences were, on average, readable by children in grade levels between 3 and 4 (i.e., approximately 10 syllables per sentence).

A schematic representation of one trial is shown in Figure 1. Each trial consisted of the silent or overt reading of one stimulus, depending on the condition (*Speaking* vs. *No-Speaking*), with stimulus length depending on group assignment (*Words* vs. *Sentences*). At the beginning of each trial, for both groups, a target stimulus appeared in white text in the center of a black screen. The font size was identical for both groups, but sentences were evenly divided onto three lines to minimize horizontal ocular movements.

After 600 ms, the target stimulus turned bright green, indicating for *Speaking* blocks that the participant should now produce the stimulus aloud (i.e., the ‘go’ signal to speak). Participants were explicitly instructed to avoid anticipating the ‘go’ signal, waiting to move until they saw the color change. For *No-Speaking* blocks, participants were instructed to refrain from producing speech when the ‘go’ signal appeared (i.e., ignore the color change and only silently read the stimulus to themselves). Participants were also explicitly asked to avoid orofacial movements (i.e., no mumbling or mouthing the stimuli). The green stimulus remained on-screen for a set amount of time (1 second for *Words*, 3 seconds for *Sentences*) prior to the beginning of a random inter-trial interval (1.5, 1.75, 2, 2.25, or 2.5 seconds).

On a subset of trials, the participant was played a brief probe tone (40 ms duration at 1 kHz, with 10 ms rise/fall times). This auditory stimulus was played 200 ms prior to the ‘go’ signal, in an effort to probe auditory processing mechanisms during speech movement planning (Figure 1). The tones were presented randomly on 40% of all trials within each block (*tone* trials vs. *no-tone* trials) binaurally through insert earphones (ER-1, Etymotic Research Inc.; Grove Village, IL, USA) at approximately 75 dB SPL. This probe stimulus was presented 200 ms prior to the ‘go’ signal to minimize visually evoked potentials and movement artifacts [10].

During all blocks, long latency auditory-evoked potentials were captured using a 128-channel nylon electrode cap (ActiveTwo, BioSemi; Amsterdam, The Netherlands), using an extension of the international 10-10 electrode system [14]. Two additional electrodes were placed on the left and right mastoids for offline re-referencing. Ocular artifacts were monitored using two electrooculography (EOG) electrodes placed on the outer canthus of the left eye and below the left eye.

In addition, four surface electromyography (sEMG) electrodes were used to record the timing of orofacial muscle activation onset. These electrodes were placed on the right side of the face over the muscles of the lower and upper lips (orbicularis oris superior and inferior), a jaw depressor (anterior belly of the digastric), and a jaw elevator (masseter). Despite the inclusionary criterion of minimal facial hair, one participant in the *Words* group arrived at the laboratory with facial hair that prevented the successful placement of sEMG electrodes. For this participant, only the lower lip signal could be reliably measured, so the lower lip was the only articulator used to determine orofacial muscle activation.

A microphone (SM58, Shure Incorporated; Niles, IL, USA) was placed in front of the participant (15 cm distance, ~45° angle) and its signal was continuously digitized along with the EEG, sEMG, and EOG signals at a sampling rate of 1024 Hz. These time-locked signals allowed for trial rejection based on premature muscle activation or audible speech output. All data was recorded on a Windows 10 environment using ActiView (version 8.11) software.

Experimental presentation, auditory stimulus delivery, and the delivery of external triggers for the EEG system were controlled using Psychtoolbox extensions [15–17] run on a Linux workstation using the Octave computational environment [18]. The auditory stimulus was first amplified by a headphone amplifier (HeadAmp6 Pro, ART ProAudio; Niagara Falls, NY, USA) prior to delivery to insert earphones. The participant's speech was transduced with an additional wireless lapel microphone (WL185 with transmitter ULX1-M1 and receiver ULXP4, Shure Incorporated; Niles, IL, USA) also positioned approximately 15 cm from the mouth and amplified by a microphone amplifier (DPS II, ART ProAudio; Niagara Falls, NY, USA) to allow for near real-time auditory feedback after amplification by the headphone amplifier. The delay in this auditory feedback setup was inherently ~10 ms.

Despite the high-precision control provided by Psychtoolbox extensions, paradigm latencies were also calibrated prior to study recruitment using a photodiode on the experimental monitor to ensure proper onset of the auditory stimulus relative to the 'go' signal, as well as overall timing accuracy for visual stimuli presentation durations. EEG trigger latencies corresponding to the beginning of each trial, the *tone*/*no-tone* timepoint, and the presentation of the 'go' signal were also verified.

Additionally, prior to each individual experimental session, the equipment was calibrated such that a speech signal (a recording of a cisgender male producing the word “tuck”) played into the lapel microphone at 75 dB SPL from 15 cm away resulted in an approximately 73 dB SPL output in both insert earphones. In addition to this auditory feedback calibration, the probe tone was verified to play at approximately 75 dB SPL in the insert earphones. Calibration was performed with a sound level meter (Type 2250A Handheld Analyzer with Type 4947 1/2” Pressure Field Microphone, Bruel & Kjaer Inc.; Norcross, GA, USA) connected to a 2 cc coupler (Type 4946, Bruel & Kjaer Inc.; Norcross, GA, USA).

3.3. Data analysis: EEG processing

To derive the necessary AEPs to calculate PSAM, the raw EEG data was offline filtered, epoched, and averaged over a frontocentral region of interest (given that modulation is largest in this region [6]), following similar signal processing protocols used in prior studies. All signal processing was executed in MATLAB with the EEGLAB toolbox (v. 2023.1; [19], based on the recommended analysis pipeline [20]).

Recorded blocks were processed one-by-one within participants, with epochs concatenated as a final step. Preprocessing involved first referencing the EEG data to the average of the mastoid signals and then low-pass finite impulse response (FIR) filtering at 50 Hz using the default Hamming windowed sinc filter. The *clean_rawdata* plugin (v. 2.91) applied a default 0.5 Hz high-pass FIR filter and then identified channels for further cleaning. Based on the recommended EEGLAB pipeline [20], we chose the following settings for channel rejection: flatlining for > 5 seconds, line noise greater than the 4 standard deviation threshold, and correlation thresholds greater than 0.8 (a more conservative value than the recommended pipeline in order to minimize “aggressive” data processing). All other *clean_rawdata* rejection methods were disabled.

Preprocessing for the sEMG channels involved filtering (using EEGLAB’s default Hamming windowed sinc filter) from 10-300 Hz prior to rectification. EOG channels were filtered from 1-10 Hz and a bipolar EOG channel was created by summing the two individual channels.

Datasets were then epoched from -600 to 1400 ms relative to the auditory stimulus (or equivalent time for *no-tone* trials). The mean amplitude of the 100 ms period prior to the stimulus onset was subtracted from the entire epoch to remove baseline differences. Epochs were automatically rejected if they contained EEG amplitudes greater than ± 200 μV or EOG amplitudes greater than ± 150 μV in any of the EOG channels. Finally, a random selection of *no-tone* epochs was removed for each participant to yield an equivalent number of *tone* and *no-tone* epochs within each condition (*Speaking* and *No-Speaking*).

Epochs were then manually inspected for further rejection. Based on the microphone signal recorded with each epoch, individual trials were rejected in which a participant failed to produce speech in a *Speaking* block or if a participant produced any speech in a *No-Speaking* block. Individual trials were also manually rejected if participants spoke much later than their other trials (e.g., attention lapses). In addition, sEMG signals were used to manually reject contaminated epochs in which the participant moved their articulators in advance of the ‘go’ signal. Overwhelmingly, the majority of manually rejected epochs were removed due to articulator movement during the planning phase despite participants being instructed to remain still until the ‘go’ signal to move. Across all participants, an additional 5.5% of epochs (8.4% of all *Speaking* epochs and 2.4% of all *No-Speaking* epochs) were manually rejected after automatic preprocessing.

After manual epoch rejection, three participants yielded a substantial difference in the number of epochs between *Speaking* and *No-Speaking* conditions. To minimize any impact of epoch count discrepancies on AEP amplitudes, we further limited the number of epochs for these 3 participants by randomly removing excess epochs (i.e., yielding comparable epoch counts between *Speaking* and *No-Speaking* conditions). Specifically, one male participant in the *Words* group consistently spoke >1 s after the 'go' signal (yielding a low number of usable *Speaking* epochs), so the number of *No-Speaking* epochs was reduced. The remaining two participants had epoch counts reduced due to excess automatic rejections for one condition, but not the other. A female participant in the *Words* group had the number of *Speaking* epochs reduced due to excessive rejections in the *No-Speaking* condition, likely due to ocular artifacts. One female *Sentences* participant had the number of *No-Speaking* epochs reduced, possibly due to large EEG signals around movement in *Speaking* blocks. Notably, the morphology of all evoked responses after such epoch count reductions remained highly similar to both the individual's starting responses and to other participants in the dataset. Thus, these three participants were retained in data analysis.

Then, for both *Speaking* and *No-Speaking* conditions, the average cortical activity recorded during trials without auditory stimulation (*no-tone* trials) was subtracted from the average EEG activity recorded during trials with auditory stimulation (*tone* trials) to remove "drift" from non-auditory activity related to motor, cognitive, and visual processes. An additional lowpass filter was applied at 15 Hz to smooth the final AEPs.

The final datasets for the *Sentences* group had a median of 93 *Speaking* epochs (range: 62 to 102, SD = 10.4) and 98 *No-Speaking* epochs (range: 61 to 103; SD = 10.5) per participant. The final epoch counts for the *Words* group were comparable at a median of 100 *Speaking* epochs (range: 30 to 108; SD = 19.9) and 99.5 *No-Speaking* epochs (range: 39 to 107; SD = 20.1).

The final AEP for the *Speaking* condition was compared with that elicited by the same auditory stimulus during the control condition without speech movement planning (*No-Speaking*). Specifically, the peak amplitude and latency of the N1 components (the largest negative peak between 70 and 130 ms after stimulus onset) were extracted from each condition over a frontocentral region-of-interest (ROI). The frontocentral ROI included the following 17 electrodes using BioSemi's ABC labeling system: A1 (equivalent to Cz), B1, B20, B21, B32, C1, C2, C11, C22, C23, C24, D1, D2, D13, D14, D15, and D18.

Finally, for each participant, a measure of PSAM was calculated by subtracting the peak *No-Speaking* N1 amplitude from the peak N1 amplitude in the *Speaking* condition. As N1 amplitudes are negative values, positive PSAM values correspond to N1 peak amplitudes that are smaller during speech movement planning (i.e., movement-related attenuation).

3.4. Statistical analyses

All statistical analyses were conducted with R in RStudio [21], using an alpha level of 0.05 to determine statistical significance. Visualizations were created using both the *ggplot2* and *cowplot* packages [22,23].

Using the *rstatix* package [24], a repeated-measures ANOVA (rmANOVA) was conducted for each dependent variable (N1 peak amplitude and N1 peak latency), with Group

(*Words, Sentences*) as a between-subjects variable and Condition (*Speaking*, *No-Speaking*) as a within-subjects variable. The Group × Condition interaction term was also included in both rmANOVAs as our primary hypothesis involved testing this interaction. To account for potential violations of the sphericity assumption, degrees of freedom for the within-subjects comparisons were adjusted with Huynh-Feldt corrections [25]. Where appropriate, partial omegas-squared (ω_p^2) are given as effect size estimates [26,27]. Significant effects in rmANOVAs were followed up with *post hoc* pairwise *t*-tests with Tukey corrections for multiple comparisons using the *emmeans* package [28].

4. RESULTS

De-identified outcome measures can be found at https://osf.io/e3j6c/. The grand average AEP waveforms across the entire ROI for both groups (*Words* vs. *Sentences*) and conditions (*Speaking* and *No-Speaking*) are shown in Figure 2.

4.1. N1 peak amplitude

The N1 component peak amplitudes are visualized for both *Words* and *Sentences* groups in Figure 3A. The main effect of Group was not significant in our N1 amplitude rmANOVA ($F[1,28] = 0.323$, $p = .574$, $\omega_p^2 < 0.01$). However, the main effect of Condition was significant ($F[1,28] = 10.427$, $p = .003$, $\omega_p^2 = 0.03$ with smaller N1 amplitudes in the *Speaking* condition than the *No-Speaking* condition ($t[28] = -3.23$, $p = .003$). The Group × Condition interaction was not statistically significant ($F[1,28] = 0.912$, $p = .348$, $\omega_p^2 < 0.01$), although descriptively there is a clearer effect of speech movement planning on N1 amplitude in the grand average AEPs in the *Words* condition (Figure 2).

4.2. N1 peak latency

Figure 3B shows the distribution of N1 peak latency values across groups and conditions. For N1 peak latency, neither the Group ($F[1,28] = 0.799$, $p = .379$, $\omega_p^2 < 0.01$) nor Condition ($F[1,28] = 0.311$, $p = .581$, $\omega_p^2 < 0.01$) main effects were significant, but Group × Condition interaction was significant ($F[1,28] = 4.259$, $p = .048$, $\omega_p^2 = 0.01$). Follow-up pairwise testing of the interaction contrasts did not yield any significant comparisons (p values all > .271).

4.3. Pre-speech auditory modulation

The calculated PSAM values for all participants are shown in Figure 4. The mean amount of PSAM for the *Words* group was 2.08 µV (range: -3.77 to 6.33; SD = 3.31), and the mean amount of PSAM in the *Sentences* group was 1.13 µV (range: -1.40 to 6.24; SD = 1.97).

5. DISCUSSION AND CONCLUSIONS

We replicated, in a new sample of participants, our established finding that typically fluent adult speakers show a modulation of cortical auditory processing during speech movement planning for monosyllabic words (*Words* group). This result is evidenced by smaller N1 amplitudes in response to probe tones presented just prior to speech production, as compared to cortically evoked responses to tones presented during the same timepoint in a silent reading control condition (i.e., with no speech movement planning).

Extending our previous work, we also found a significant amount of pre-speech auditory modulation when participants produced more complex, connected speech in the same experimental paradigm (*Sentences* group). Indeed, the interquartile ranges of both PSAM boxplots are above zero (Figure 4). The lack of a significant interaction effect between Condition and Group in the N1 amplitude rmANOVA suggests that there is little difference between the pre-speech modulation of auditory-evoked responses when comparing the production of words with the production of sentences.

In this study, the *Words* group produced a mean PSAM value of 2.08 μV, which is comparable to our prior work in typically fluent adults [6–9]. Notably, however, there were more participants in the *Words* group with larger PSAM values (greater than 2 μV; Figure 4), and many *Sentences* participants with PSAM values close to 0 μV. Indeed, the group AEPs show a much clearer modulation of N1 amplitudes during speech movement planning for the *Words* group (Figure 2). None of the PSAM values for the *Sentences* group were statistical outliers, but only four participants showed a modulation of more than 2 μV, as compared to more than half of the *Words* group (*N* = 9).

The descriptively smaller PSAM values in the *Sentences* group is particularly interesting in light of our current working hypothesis for this modulatory phenomenon. Specifically, we have recently hypothesized that smaller auditory-evoked responses during movement planning could possibly reflect a reduced processing capacity of cortical auditory neurons due to their ongoing bidirectional communication with premotor and motor areas [13]. If the neural processes reflected in PSAM are indeed involved in the selection of feedforward motor commands, then it is perhaps unsurprising that a greater amount of PSAM is observed for individual monosyllabic words, which are thought to be predominately under feedforward control. On the contrary, the production of longer utterances (i.e., entire sentences) relies more heavily on feedback-based control mechanisms as the duration of the utterance is long enough for the auditory feedback signal to be integrated at the level of cortex. Thus, the pre-speech period prior to the production of sentences may therefore correspond to a reduced reliance on predictive bidirectional communication, as feedback-based control mechanisms will soon also be recruited. As such, cortical auditory neuronal populations will still be readily ‘available’ to respond to the probe tone, yielding a smaller dampening of N1 amplitude prior to sentence production (i.e., less PSAM) as compared to single word utterances.

The distribution of PSAM values across groups also brings up the difference in motor preparation needed for sentences vs. words. We specifically chose to extend our established paradigm by varying *only* the length of the utterances produced. However, it is possible that the 600 ms “delay” period prior to the onset of the ‘go’ signal should have been extended for the production of more complex stimuli. Specifically, it is possible that the descriptively smaller amount of PSAM in the *Sentences* group was due, in part, to participants still *reading* the stimuli at the time of the probe tone’s presentation. As such, participants may not have yet begun preparing for subsequent movement. If we were inadvertently probing the behavior of the auditory system during the process of simply *reading* stimuli (i.e., speech movement planning has yet to occur), the N1 amplitudes would be much more similar between *Speaking* and *No-Speaking* conditions. Furthermore, we chose to divide the sentences onto three lines in order to minimize horizontal ocular artifacts, so participants in the *Sentences* group certainly required more time in order to saccade back to the beginning of the stimulus upon the onset of the ‘go’ signal. Despite these speculations, we did indeed measure a significant main effect of Condition for N1 amplitude, and showed no significant difference between the two groups in the amount of PSAM measured. However, future work could adjust the paradigm timeline, visual layout, or task instructions in an effort to more closely match the two tasks and further describe pre-speech auditory behavior prior to complex utterances.

Relatedly, the underlying nature of speech movement planning is brought to question with our chosen paradigm. Our study design assumed that movement planning is indeed more complex prior to the production of sentences compared to monosyllabic words. However, there is no straightforward way for researchers to directly measure how much movement planning has occurred at a specific timepoint prior to the onset of movement itself (i.e., when the probe tone is played in our PSAM paradigm). It is possible that participants prepare only one word (or a few movement gestures) at a time, regardless of the stimuli length that they are about to produce. Alternatively, at the time of the probe tone, participants in the *Sentences* condition may have already planned several syllables ahead. Clever experimental paradigms could be used in the future to begin to answer such complex questions and inform future studies of pre-speech auditory processing.

Finally, we also found a significant interaction between Condition and Group when analyzing the latency of the N1 component peaks. However, none of the pairwise comparisons were significant in *post-hoc* contrast testing. Descriptively, the *Sentences* group showed, on average, a longer N1 latency in the *Speaking* condition as compared to the *No-Speaking* control condition while the *Words* group showed a shorter latency in the *Speaking* condition (Figure 3B). It is unlikely that this significant interaction term (p = .048) is highly theoretically meaningful (ω_p^2 = 0.01), although future work could further investigate the impact of speech movement planning complexity on the evolution of auditory processing.

In conclusion, we successfully replicated our previous finding of pre-speech auditory modulation in a new sample of adult speakers. We also demonstrated that pre-speech auditory modulation, a sensorimotor phenomenon easily measured in laboratory settings, can indeed be measured prior to sentence-level speech, extending our prior research involving only monosyllabic utterances. Examining cortical auditory responses prior to the production of longer utterances is critical for understanding the functional relevance of this phenomenon, particularly as we further investigate our hypothesis relating PSAM to the fine-tuning of feedforward motor commands for subsequent speech movements. Additionally, these findings may be critical to understanding neurodevelopmental disorders like stuttering in which we have already documented a limited amount of modulation [6–9,11], as stuttering rarely occurs on isolated monosyllabic utterances. For example, future work measuring PSAM prior to stuttered vs. fluent speech would contribute to a deeper understanding of the neurophysiological basis of stuttering – and stuttered speech itself. In the long term, such findings have the potential to inform novel fluency treatment approaches (e.g., noninvasive brain stimulation) for individuals who stutter, directly targeting neural pathways involving pre-speech motor-to-sensory interactions.

6. ARTWORK

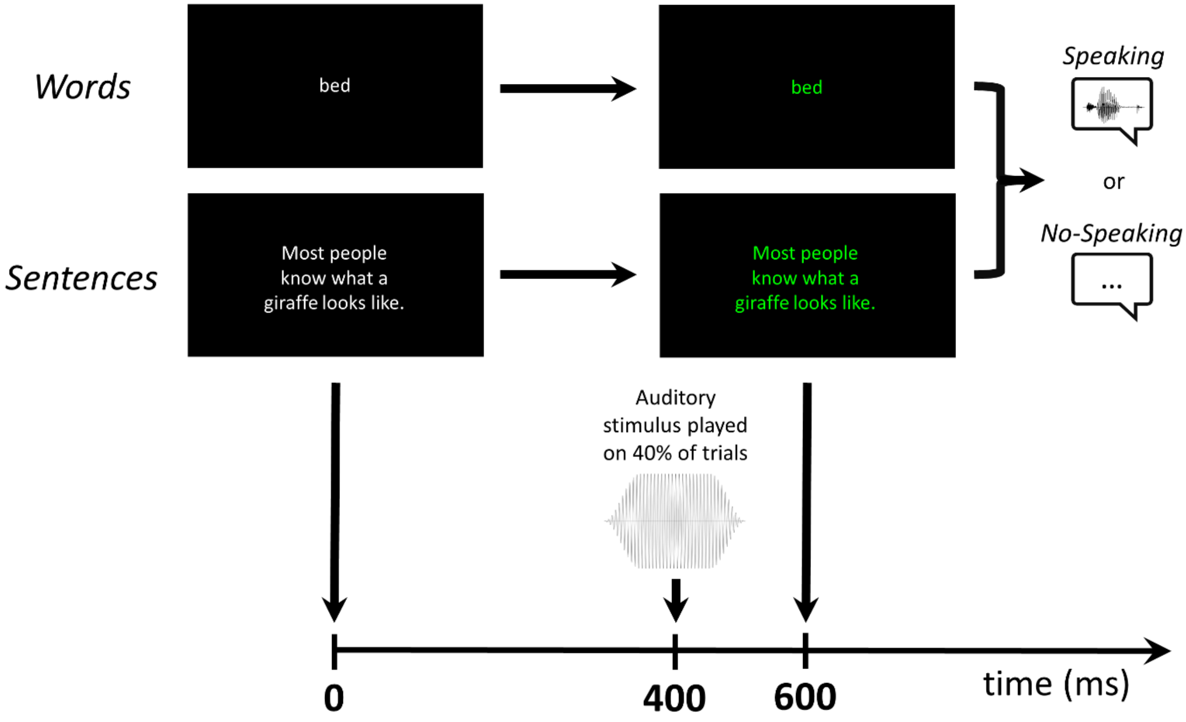

Figure 1

Schematic representation of one trial for both *Words* (top) and *Sentences* (bottom) groups. During *Speaking* blocks, participants produced one stimulus at a time after the ‘go’ signal (i.e., when the text turned green), 600 ms after the initial appearance of the stimulus in white text. During *No-Speaking* blocks, participants refrained from overtly producing the stimulus (i.e., they ignored the ‘go’ signal and silently read the stimulus to themselves). On 40% of all trials, a probe tone was played into the participant’s ears via insert earphones (*tone* trials). This pure tone auditory stimulus was played 200 ms prior to the ‘go’ signal to probe auditory processing mechanisms during speech movement planning.

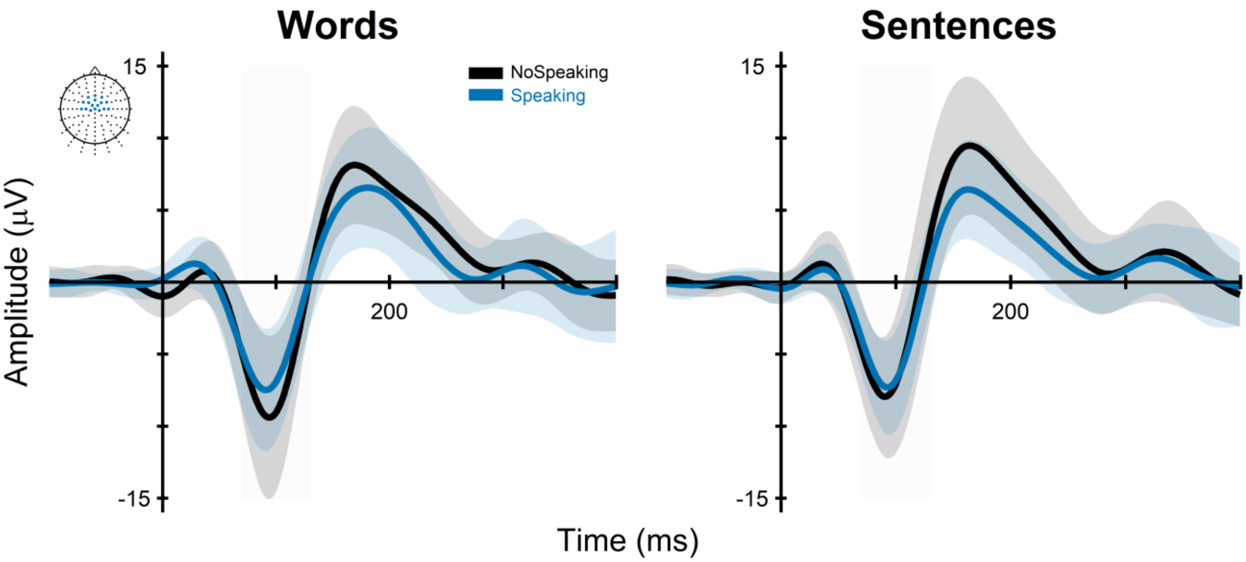

Figure 2

The grand average AEPs elicited by probe tones presented 200 ms in advance of the ‘go’ signal to produce both *Words* (left, blue) and *Sentences* (right, blue), compared to the AEPs measured at the same timepoint during a silent reading control condition (black). The top-down head inset highlights the frontocentral ROI used for analysis.

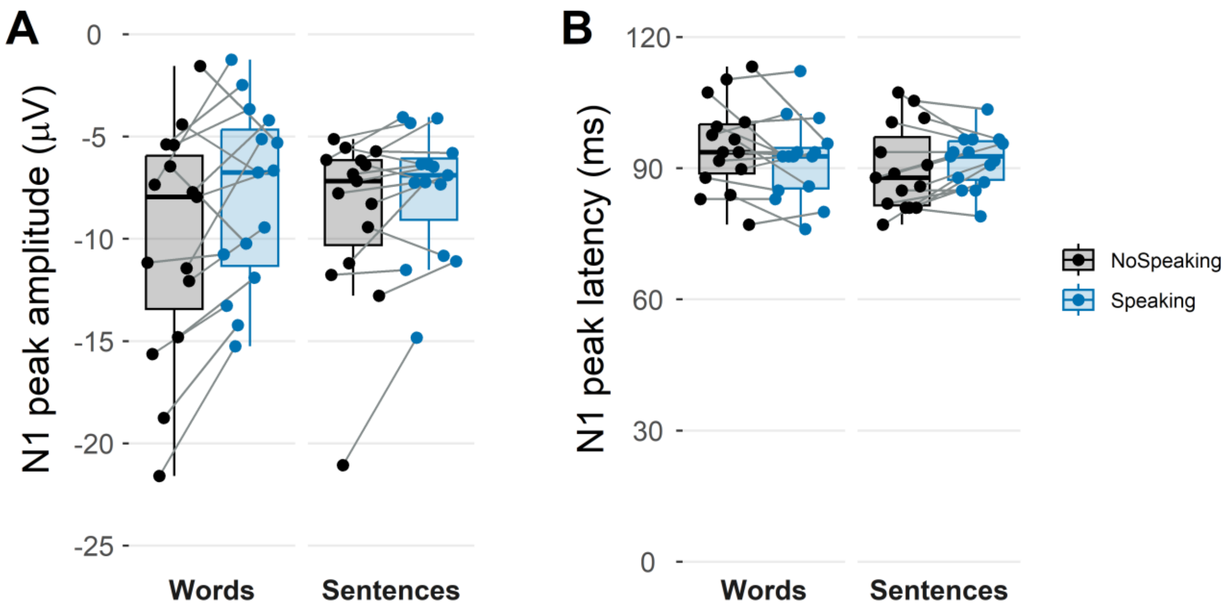

Figure 3

Summary of N1 component parameters for all participants across both the *Words* and *Sentences* groups (*N* = 15 adults each). A) Boxplots for peak N1 amplitudes. Individual colored points correspond to values from single participants, with points connected with a line to show individual performance across conditions (*Speaking* in blue vs. *No-Speaking* in black). For each boxplot, the solid horizontal line represents the group median value, with box edges representing the interquartile range (25th to 75th percentiles) and whiskers extending to the most extreme values within 1.5 times the interquartile range. B) Boxplots for peak N1 latencies. Data representation identical to panel A.

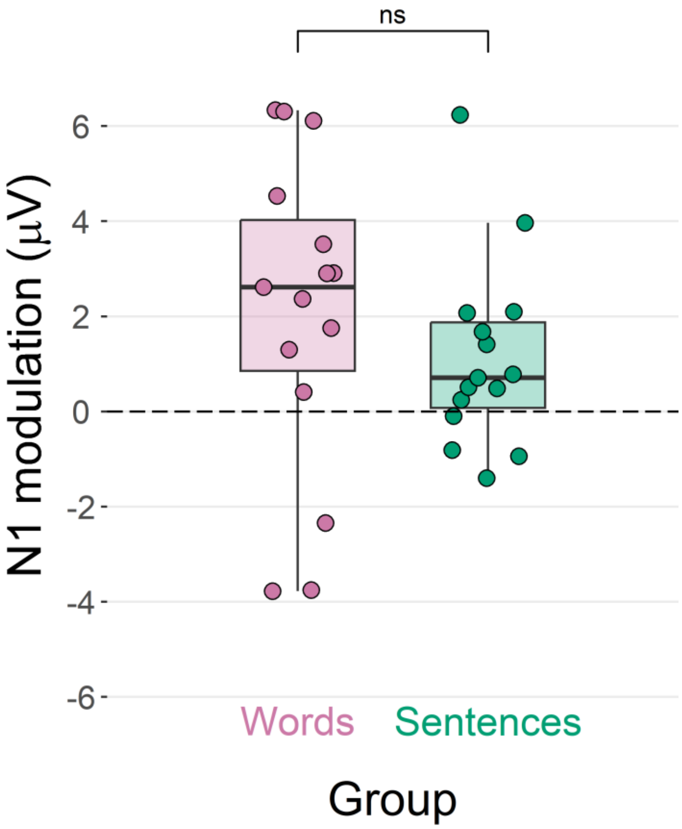

Figure 4

Pre-speech auditory modulation (PSAM) values summarized for all participants. Boxplots are shown for both the *Words* and *Sentences* groups (*N* = 15 adults each). Each colored point corresponds to the amount of modulation (i.e., the difference in peak N1 amplitude between *Speaking* and *No-Speaking* conditions) for one participant (*Words* participants in pink, *Sentences* participants in green). Boxplot visualization details as in Figure 3. Positive amounts of modulation represent an N1 peak that is smaller in amplitude when the participant is preparing to speak.

7. ACKNOWLEDGEMENTS

The authors would like to thank Hannah Goldfarb, Zoie Breland, and Riley Kwinn for assisting with electrode cap preparation. We would also like to thank Priyanka Talur and Zoie Breland for assisting with manual epoch inspections.

8. AUTHOR CONTRIBUTIONS

Elise LeBovidge: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review & editing, Visualization.

Ludo Max: Conceptualization, Methodology, Resources, Writing – review & editing, Supervision, Project administration, Funding acquisition.

9. FUNDING SOURCES

This research was supported by a training grant from the National Center for Advancing Translational Sciences at the National Institutes of Health under award number TL1TR002318, as well as a grant from the National Institute on Deafness and Other Communication Disorders (R01 DC017444; PI: Ludo Max). This content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

10. REFERENCES

[1] P.M. Bays, J.R. Flanagan, D.M. Wolpert, Attenuation of self-generated tactile sensations is predictive, not postdictive, PLoS Biol. 4 (2006) 281–284. https://doi.org/10.1371/journal.pbio.0040028.

[2] S.-J. Blakemore, C.D. Frith, D.M. Wolpert, Spatio-temporal prediction modulates the perception of self-produced stimuli, J. Cogn. Neurosci. 11 (1999) 551–559. https://doi.org/10.1162/089892999563607.

[3] S.-J. Blakemore, D.M. Wolpert, C.D. Frith, Why can't you tickle yourself?, NeuroReport 11 (2000) R11–R16. https://doi.org/10.1038/news981022-7.

[4] R. Shadmehr, Learning to predict and control the physics of our movements, J. Neurosci. 37 (2017) 1663–1671. https://doi.org/10.1523/JNEUROSCI.1675-16.2016.

[5] R. Shadmehr, M.A. Smith, J.W. Krakauer, Error Correction, Sensory Prediction, and Adaptation in Motor Control, Annu. Rev. Neurosci. 33 (2010) 89–108. https://doi.org/10.1146/annurev-neuro-060909-153135.

[6] A. Daliri, L. Max, Electrophysiological evidence for a general auditory prediction deficit in adults who stutter, Brain Lang. 150 (2015) 37–44. https://doi.org/10.1016/j.bandl.2015.08.008.

[7] A. Daliri, L. Max, Modulation of auditory processing during speech movement planning is limited in adults who stutter, Brain Lang. 143 (2015) 59–68. https://doi.org/10.1016/j.bandl.2015.03.002.

[8] A. Daliri, L. Max, Modulation of auditory responses to speech vs. nonspeech stimuli during speech movement planning, Front. Hum. Neurosci. 10 (2016) 1–9. https://doi.org/10.3389/fnhum.2016.00234.

[9] A. Daliri, L. Max, Stuttering adults' lack of pre-speech auditory modulation normalizes when speaking with delayed auditory feedback, Cortex 99 (2018) 55–68. https://doi.org/10.1016/j.cortex.2017.10.019.

[10] L. Max, J.C. Daniels, K.M. Curet, K.L. Cronin, Modulation of auditory and somatosensory processing during the planning of speech movements, Proc. ISSP 2008 - 8th Int. Semin. Speech Prod. 2 (2008) 41–44.

[11] L. Max, A. Daliri, Limited pre-speech auditory modulation in individuals who stutter: Data and hypotheses, J. Speech Lang. Hear. Res. 62 (2019) 3071–3084. https://doi.org/10.1044/2019_JSLHR-S-CSMC7-18-0358.

[12] H. Wang, Y. Ali, L. Max, Perceptual formant discrimination during speech movement planning, PLOS ONE 19 (2024) e0301514. https://doi.org/10.1371/journal.pone.0301514.

[13] J.J. Li, A. Daliri, K.S. Kim, L. Max, Does pre-speech auditory modulation reflect processes related to feedback monitoring or speech movement planning?, Neurosci. Lett. (n.d.) 138025. https://doi.org/10.1101/2024.07.13.603344.

[14] R. Oostenveld, P. Praamstra, The five percent electrode system for high-resolution EEG and ERP measurements, Clin. Neurophysiol. 112 (2001) 713–719. https://doi.org/10.1016/S1388-2457(00)00527-7.

[15] D.H. Brainard, The Psychophysics Toolbox, Spat. Vis. 10 (1997) 433–436.

[16] M. Kleiner, D.H. Brainard, D.G. Pelli, What's new in Psychtoolbox-3?, in: ECVP Abstr. Suppl., 2007.

[17] D.G. Pelli, The VideoToolbox software for visual psychophysics: Transforming numbers into movies, Spat. Vis. 10 (1997) 437–442.

[18] J.W. Eaton, D. Bateman, S. Hauberg, R. Wehbring, GNU Octave version 6.1.0 manual: a high-level interactive language for numerical computations., (2020). https://www.gnu.org/software/octave/doc/v6.3.0/.

[19] A. Delorme, S. Makeig, EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis, J. Neurosci. Methods 134 (2004) 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009.

[20] A. Delorme, EEG is better left alone, Sci. Rep. 13 (2023) 2372. https://doi.org/10.1038/s41598-023-27528-0.

[21] RStudio Team, RStudio: Integrated Development Environment for R, (2020). www.rstudio.com.

[22] H. Wickham, ggplot2: Elegant Graphics for Data Analysis, (2016). https://ggplot2.tidyverse.org.

[23] A. Kassambara, rstatix: Pipe-Friendly Framework for Basic Statistical Tests, (2023). https://CRAN.R-project.org/package=rstatix.

[24] L. Max, P. Onghena, Some issues in the statistical analysis of completely randomized and repeated measures designs for speech, language, and hearing research, J. Speech Lang. Hear. Res. 42 (1999) 261–270. https://doi.org/10.1044/jslhr.4202.261.

[25] R.V. Lenth, emmeans: Estimated Marginal Means, aka Least-Squares Means, (2024). https://rvlenth.github.io/emmeans/.

Section IV: Cortical somatosensory responses evoked by orofacial skin stretch experimentally applied during speech movement planning in stuttering and nonstuttering adults (*Study 2*)

(Manuscript formatted for submission to *Journal of Neurophysiology*)

Elise LeBovidge[1]

Takayuki Ito[2]

Ludo Max[1]

[1]University of Washington, Seattle, Washington, USA

[2]Laboratoire de Recherche Grenoble, Images, Parole, Signal, Automatique (GIPSA-Lab), Grenoble Institute of Technology, Université Grenoble Alpes, Grenoble, France

Correspondence: Ludo Max (LudoMax@uw.edu)

Running head: Somatosensory processing during speech movement planning

1. ABSTRACT

Prior studies have demonstrated that individuals who stutter do not show the typical modulation of auditory processing (reduced N1 amplitude) that is observed during speech movement planning in individuals who do not stutter. We sought to determine whether this lack of sensory modulation during speech movement planning in stuttering adults is specific to the auditory domain. We extended our evoked-potentials paradigm to a somatosensory study with 15 stuttering and 15 nonstuttering adults. Instead of probe tones presented during speech movement planning versus a silent-reading control condition, a robotic device applied rapid orofacial skin stretches to elicit somatosensory evoked potentials in the same conditions and at the same timepoints. We compared N1 amplitudes for *Speaking* and *No-Speaking* conditions to assess a potentially modulating influence of movement planning on somatosensory processing. For these specific stimuli and timepoints, we found no evidence of pre-speech somatosensory modulation in either group of participants. Both groups showed clear N1 peaks in response to the orofacial skin stretch but there was no difference in peak amplitude between the speaking and control conditions. Consequently, there was also no between-group difference. Given that even the typical speakers showed no modulation whereas other studies have yielded evidence for this phenomenon in the context of both orofacial and limb movements, our results suggest that central input from the specific facial mechanoreceptors stimulated here is not subject to modulation during speech planning. Future studies may benefit from direct stimulation of effectors actively involved in performing the relevant articulatory gestures (lip, tongue, jaw).

Keywords: movement planning; sensorimotor integration; somatosensation; speech; stuttering

2. NEW AND NOTEWORTHY

This study is the first demonstration that somatosensory cortical responses can be evoked with experimentally applied orofacial skin stretches during speech movement planning. Unlike the N1 amplitude in auditory evoked potentials in response to probe tones, N1 amplitude in these somatosensory evoked potentials showed no evidence of a modulating influence of speech movement planning in typical speakers.

3. INTRODUCTION

Contemporary computational models of speech production illustrate how the integration of auditory and somatosensory feedback is critical for precise speech motor control (1–3). As is the case for vision and somatosensation in the neural control of limb movements (4), speech articulation (i.e., movements of the lips, tongue, jaw, and velum) requires a continuous weighting of auditory and somatosensory reafferent signals to allow feedback-based online compensatory responses as well as adaptive feedforward-based planning updates for future movements (5, 6). In speech, these articulatory movements shape the vocal tract to generate resonances, bursts, frication noise, and other sounds, depending on the language of the speaker. Given that (a) the effectors move to achieve specific acoustic targets, (b) some speakers have no or limited hearing, and (c) the acoustic signal is readily accessible for the digital implementation of real-time auditory feedback perturbations, the role of auditory feedback has been extensively studied for several decades (7, 8). In contrast, the role of somatosensory feedback (here referring specifically to reafference from proprioceptors and cutaneous mechanoreceptors) in speech articulation has received comparatively much less attention (9–11, but see 12–14).

One consequence of this situation is that efforts to understand the mechanisms underlying *speech disorders* such as stuttering have also been approached with a heavy focus on the auditory system. On the one hand, this has been a logical approach because it was already known as early as the 1950s that experimentally altered auditory feedback is fluency-enhancing for many individuals who stutter (15–22). More recently, various behavioral and neuroimaging paradigms have also provided compelling evidence for functional and anatomical auditory-motor integration differences between stuttering and nonstuttering speakers (23–33). On the other hand, this approach may prevent important insights into the true nature of the mechanisms underlying stuttering if not accompanied by equally strong efforts to study non-auditory systems such as reafferent signals and neural substrates in the *somatosensory* system. In fact, there is some initial evidence suggesting that investigations of the integrity and use of somatosensory feedback in individuals who stutter may indeed be a promising line of research. For example, one study found that, compared with nonstuttering control participants, considerably fewer stuttering participants showed a speech initiation-related reduction in the excitability of lip muscle reflexes evoked by mechanical stretch applied to the corner of the mouth (34). As a second example, recent neuroimaging studies have revealed broad sensorimotor network differences (e.g., decreased functional connectivity between somatosensory regions, reduced interhemispheric connections across somatosensory cortices) between individuals who do and do not stutter (25, 35).

One efficient and fruitful initial approach may be to investigate whether replicated atypical characteristics of auditory-motor interactions in individuals who stutter also extend to somatosensory-motor interactions. Here, we take this approach by modifying a well-established paradigm from our own lab that has produced, across several studies, compelling evidence that typical speakers show a modulation of auditory processing during the movement planning phase immediately prior to speech onset whereas this pre-speech auditory modulation (PSAM) is minimal or absent in speakers who stutter (36–40). Specifically, PSAM is assessed by presenting probe tones during a short delay phase in a delayed response speaking task versus equivalent intervals in no-speaking control conditions and quantifying the observed reduction in amplitude of the N1 component in long-latency auditory-evoked potentials derived from electroencephalographic (EEG) recordings. The neural generators for this N1 component are known to be located in both primary and association auditory cortex (41, 42).

The finding of PSAM in typical speakers is consistent with a large body of work documenting that – across humans and non-human animal models, across data collection techniques (e.g., direct neurophysiological recordings, neuroimaging, and behavioral and psychophysics paradigms), and across effector systems (e.g., oculomotor, limb, orofacial) – there are extensive modulatory influences on sensory processing during movement planning and execution (43–56). Indeed, for limb movements, observations of a suppression of somatosensory-evoked responses even prior to movement onset have been reported for decades (54, 57–59). Thus, our consistent finding of a lack of PSAM in speakers who stutter is an intriguing finding that can be contextualized in an extensive literature on movement-dependent sensory modulation but for which it remains entirely unknown whether the problem is specific to the *auditory* system. Interestingly, even for non-disordered speech, only the influence of producing isolated prolonged vowels or maintaining silent vocal tract postures on cortical *somatosensory*-evoked potentials (SEPs) has been investigated (60). This study analyzed SEPs in response a sudden posterior-directed stretch of the orofacial skin at the corners of the mouth that was experimentally applied with a robotic device while participants sustained (2500 ms) three different vowels and three silent postures corresponding to the same vowels. The amplitude of a negative SEP component around 140 ms after stimulus onset was statistically significantly reduced from the rest-condition baseline for one of the three vowels as well as for the corresponding silent posture. The amplitude of a positive SEP component around 250 ms was statistically significantly reduced from baseline for the vowels but not for the postures. Overall, the study provided initial evidence in support of the idea that, at least during the production of sustained vowel production, speech motor execution exerts a modulatory influence on somatosensory processing.

Therefore, the present study investigated (a) whether, *during movement planning prior to speech onset*, a pre-speech somatosensory modulation (PSSM) can be detected analogous to the well-documented PSAM phenomenon for auditory stimuli, and (b) whether any observations of PSSM differ between speakers who stutter and typically fluent speakers as is also the case for PSAM. Instead of using probe tones to assess cortical auditory processing during speech movement planning, we used a modified version of the above-mentioned method of orofacial skin stretch at the corners of the mouth. The paradigm modification consisted of applying the stretch during *movement planning* for the production of *existing monosyllabic words* (and a no-speaking silent reading control condition with the same word list) and in the upward direction (as all target words involved mouth opening but not necessarily lip spreading). Specifically, the somatosensory stimulation was applied during the delay phase of a delayed-response speaking task and the same time point in a corresponding no-speaking control task identical to those used in our previous PSAM studies. SEPs were recorded by means of scalp electrodes in an EEG cap and the amplitude of the largest SEP component was measured in both conditions.

4. MATERIALS AND METHODS

4.1. Participants

The participants for this study were 15 adults who stutter (mean age = 32 years, range = 19-57 years) and 15 adults who do not stutter (mean age = 32 years, range = 18-56 years). To the extent possible, nonstuttering participants were recruited based on pair-wise matching with a stuttering participant for self-reported age (± 3 years), handedness, and biological sex. An exception to the age-matching criterion was made for the only left-handed stuttering participant due to difficulty finding an age-matched left-handed nonstuttering participant (the included pair of participants differed in age by 6 years).

The stuttering group included 12 cisgender males, 2 cisgender females, and 1 nonbinary participant assigned female at birth. The nonstuttering group included 12 cisgender males and 3 cisgender females. All participants were native speakers of American English (learned prior to age 5 years), according to self-report. Specific exclusion criteria included (a) a history of speech, language, or hearing diagnoses other than stuttering, (b) neurological or psychological diagnoses, (c) medications with potential sensorimotor side effects, and (d) facial hair that would interfere with the experimental procedure.

Participants who stutter self-identified as currently stuttering and reported that the onset of their stuttering occurred during childhood. At our lab, the presence of stuttering was confirmed by the first author (a researcher trained in, and experienced with, clinical measures of stuttering) through administration of the Stuttering Severity Instrument (SSI-4; 61). Participants were included in the study if they received a total SSI score of at least 10 (the clinical threshold for adults) on the day of the experiment or, in rare cases, had met this score during a prior visit to our laboratory and had an extensive clinical history of stuttering. Each participant's stuttering frequency and overall severity, together with other relevant characteristics, are provided in Table 1. With regard to all stuttering measures, it was previously determined that the first author's identification of stuttering moments shows reliable agreement with the senior author (who holds clinical certification from the American Speech-Language-Hearing Association), with an average Cohen's κ statistic of .90 (range .84-.95) across five different SSI samples (62).

According to a self-report demographics questionnaire restricted to the categories used by the National Institutes of Health (NIH), all 15 stuttering participants reported being Not Hispanic/Latino; the majority (*n* = 12) were White whereas the remaining participants were Asian (*n* = 2) or Mixed (*n* = 1). Of the nonstuttering participants, the majority reported being Not Hispanic/Latino and White (*n* = 12); the remaining participants reported being Non-Hispanic/Latino and Asian (*n* = 2) or Hispanic/Latino and White (*n* = 1).

Prior to data collection for the actual experiment, all participants underwent a behavioral pure-tone audiometric screening. Octave frequencies from 250 to 4000 Hz were tested in both ears. Across both groups combined, 27 of 30 participants had typical hearing thresholds ≤ 25 dB HL across all frequencies in both ears. One nonstuttering participant had a threshold of 30 dB HL at 4 kHz in the left ear. One stuttering participant had three frequencies with elevated thresholds: 30 dB HL for the left ear at 1 kHz, 30 dB HL at 4 kHZ in the right ear, and 45 dB HL for the right ear at 500 Hz. A second stuttering participant had left-ear thresholds of 30 dB HL at 2 kHz and 60 dB HL at 4 kHz and a right-ear threshold of 50 dB HL at 4 kHz.

Fourteen of the stuttering participants and 3 of the nonstuttering participants had previously participated in a research study in our laboratory, but none of the participants in either group were included in our published PSAM studies. Thus, participants were generally unfamiliar with the paradigm and tasks. All experimental procedures were approved by the

University of Washington’s Institutional Review Board. Prior to participation, all participants provided informed written consent. Participants were financially compensated for the time spent in the laboratory.

4.2. Experimental paradigm

All participants completed a single experimental session that lasted approximately 3 hours for participants who stutter (including SSI-4 administration) and 2 hours for participants who do not stutter. All participants were allowed breaks in-between separate blocks of test trials, as needed. The paradigm involved a delayed-response speaking task and an equivalent no-speaking control task while cortical evoked potentials were recorded in response to somatosensory stimuli. The stimulation was implemented using an orofacial skin stretch technique developed by Ito and colleagues (12, 60, 63) and described in detail below.

A schematic representation of the setup is provided in Figure 1. Participants were seated in a height-adjustable chair in front of a 27” monitor (U2722D, Dell, Inc.; Round Rock, TX, USA) while wearing a 128-channel nylon electrode cap (ActiveTwo, BioSemi; Amsterdam, The Netherlands). Electrodes were positioned in the cap using an extension of the international 10-10 electrode system (64). Two additional electrodes were placed on the left and right mastoids for offline re-referencing. Ocular artifacts were monitored using electrooculography (EOG) electrodes at the outer canthus and below the left eye.

A detailed description and demonstration of the somatosensory stimulation setup is provided in Ito et al. (2015). The stimulation was controlled by a robotic device (Phantom Premium 1.0A, 3D Systems; Rock Hill, SC, USA) placed behind the participant. A custom-made eye hook was attached at the distal tip of the robot arm. Flexible nylon wire (KastKing

SuperPower Braided Fishing Line, 10-lb. strength; Garden City, NY, USA) and two small plastic tabs (2x3 cm cuts of overhead transparency sheet) were used to couple the robot to the orofacial skin adjacent to both corners of the mouth (left and right). Using double-sided wig tape (Walker Tape 1522 clear medical tape; Saint Paul, MN, USA), the plastic tabs were attached just lateral to the oral angle (or for one pair of matched participants 1 cm more lateral due the stuttering participant's facial hair). The left- and right-side wires were both routed from the robot forward, looping over articulating metal arms (Manfrotto, Videndum plc; Richmond, UK), and then downward to the participant's face (Figure 1).

All participants completed two different conditions, *Speaking* and *No-Speaking*, with trials that were identical in terms of visual appearance and timeline. A schematic representation of one trial is shown in Figure 2. In the *Speaking* condition, participants saw a monosyllabic word appear in white characters on a black background, waited until the characters turned green after 600 ms (the 'go' signal), and then produced the word. In the *No-Speaking* condition, trials contained the same sequence of events on the monitor but participants refrained from producing the word (i.e., they read the word silently and ignored the 'go' signal). The word in green characters remained on the screen for 1 second before a jittered inter-trial interval started (randomly selected from 1000, 1250, 1500, 1750, or 2000 ms). The next trial started immediately after the inter-trial interval. On 40% of all trials, orofacial skin stretch was applied 200 ms prior to the 'go' signal. An upward stretch was applied by moving the robot tip backwards away from the participant. The force trajectory followed a single cycle of a 3 Hz sinusoidal movement, ramping up to 4 N maximum force before returning back to the start position.

Participants completed three blocks of 90 trials for each of the two conditions (*Speaking* and *No-Speaking*), yielding a total of 270 trials per condition. Block order was pseudorandomized for each individual participant. Given that only 40% of the trials involved somatosensory stimulation, a total of 108 *stretch* trials were recorded per condition for each participant. Due to equipment issues, two nonstuttering participants and three stuttering participants had either one shorter or one missing *No-Speaking* block. However, we verified that the final number of available data epochs did not differ significantly between conditions for either the Stuttering or the Nonstuttering group.

All words shown and produced during the tasks were consonant-vowel-consonant (CVC) sequences (e.g., “gag” /gæg/). Given that an upward orofacial skin stretch was applied, we selected words that minimized forward-directed lip movement for the onset consonant (i.e., no initial bilabials with lip protrusion) but maximized downward jaw movement (i.e., all words contained the low vowel /æ/). Words were pseudo-randomly selected at the start of each block of 90 trials, drawn from a total word bank of 49 items: individual words occurred no more than twice in a single block and never back-to-back.

As mentioned by Ito and colleagues (63), applying sudden mechanical stimulation to the participants’ face has the potential to induce motion artifacts that may contaminate the EEG data. For this reason, a familiarization block of 10 *Speaking* stretch trials was presented prior to beginning the first experimental block. The familiarization ensured that (a) the participant understood the task instructions, (b) there was no excessive blinking or movement during the pre-speech time window of interest, and (c) the somatosensory stimulation did not induce motion artifacts. Each participant required only one familiarization block.

Visual presentation of the target words, somatosensory stimulus delivery, and triggers for the EEG recording system were controlled by MATLAB (version 2022a; The MathWorks, Inc.; Natick, MA, USA) using Psychtoolbox extensions (65–67) in a Windows 10 environment. All data were recorded in Biosemi ActiView (version 9.02) software on a Linux workstation. Despite the high-precision control provided by Psychtoolbox extensions, correct timing of all visual stimuli, somatosensory stimuli, and trigger signals was first tested, adjusted, and verified by using a high-speed camera to detect robot movement and detect color changes on the computer monitor in front of the participant.

Participants heard their self-produced speech in free field during data collection (i.e., no headphones or insert earphones were used). Participants spoke at a comfortable, self-selected intensity. A microphone (SM58, Shure; Niles, IL, USA) in front of the participant (~15 cm from the mouth) was used to digitize the acoustic speech signal into the same ActiveTwo system recording the EEG and EOG signals at a sampling rate of 1024 Hz. These time-locked acoustic and physiological signals allowed for trial rejection based on premature or incorrect speech output (see *Data analysis*).

4.3. “No stretch” auditory control group

As a control experiment to verify that the motors and movement of the robot itself did not produce sound that would evoke a cortical *auditory* response, three additional nonstuttering participants (all assigned female at birth and all right-handed) performed the same task and with the robot performing the same movements on a random subset of trials but without a physical connection between the robot and orofacial skin. Thus, no skin stretch was applied to the participant. Following the same automatic preprocessing steps described below, there was no evidence of an auditory evoked potential (i.e., no clear peaks in the grand average signal) for either the *Speaking* or *No-Speaking* condition.

4.4. Data analysis

Data were processed using the EEGLAB (68) toolbox in MATLAB version R2022a (Mathworks, Natick, USA). First, the raw EEG signals were re-referenced to the average signal from the two mastoid electrodes and low-pass filtered (finite impulse response filter, 50 Hz cut-off, Hamming window). Following the recommended EEGLAB pipeline (69), the *clean_rawdata* plugin (version 2.91) was then used to detect bad channels after a default high-pass filter (finite impulse response filter, 0.5 Hz cut-off). Specifically, flatlined electrodes (exceeding 5 seconds), electrodes with line noise (4 standard deviation threshold), and channels exceeding a 0.8 correlation threshold were detected and interpolated. All other *clean_rawdata* rejection methods were disabled. The EOG channels were band-pass filtered from 1-10 Hz and a bipolar EOG channel was created by summing the two individual channels.

The continuous EEG data were epoched from -600 to 1400 ms relative to the onset of the somatosensory stimulus (or equivalent time for *no-stretch* trials). The epochs were baseline-normalized by subtracting the mean amplitude of the 100 ms pre-stimulus period from the entire epoch. Individual epochs with EOG signals over ±150 µV or EEG amplitudes greater than ±200 µV in any channel were automatically rejected. Finally, for each participant, randomly selected *no-stretch* epochs were removed such that the number of retained *no-stretch* epochs matched the number of *stretch* epochs for each condition (as only 40% of trials were *stretch* trials). Epoch sets were then manually inspected for further rejection. Based on the time-locked microphone and EOG signals, epochs were rejected if a participant (a) failed to produce speech in the *Speaking* condition, (b) produced any sound in the *No-Speaking* condition, (c) began speaking prior to the event marker for the ‘go’ signal, or (d) had ocular artifacts around the time of the applied stretch that did not meet the EOG threshold for automatic rejection.

Epochs were also rejected due to unambiguous, overt speech dysfluencies. One participant stuttered on a large number of *Speaking* trials (55.7%), yielding a substantial difference in the number of epochs available for the *Speaking* (48 epochs) versus *No-Speaking* (100 epochs) conditions. To avoid confounding effects of the large difference in number of epochs on SEP amplitude, we further limited the number of *No-Speaking* epochs by randomly removing 50 additional epochs after manual inspection (i.e., yielding 50 *No-Speaking* epochs). The morphology of the evoked responses after this rejection appeared highly similar to other participants in the dataset, so this participant was retained in all subsequent analyses despite the reduced number of trials.

Two nonstuttering participants had a similar discrepancy between the number of epochs across the two conditions, driven by a large number of ocular artifacts during the *No-Speaking* condition. Therefore, a random selection of otherwise acceptable *Speaking* epochs was removed for both of these participants as described above (removing 65 additional epochs for one participant and 40 epochs for another), yielding a similar number of final epochs for both the *Speaking* and *No-Speaking* conditions. One participant's dataset ended with a total of 60 *No-Speaking* epochs and 59 *Speaking* epochs while the other participant ended with the lowest epoch totals across the entire experimental dataset: 35 *No-Speaking* epochs and 35 *Speaking* epochs. Both participants were still included in data analysis after verifying that the morphology of the SEPs after additional epoch rejection remained similar to that of other participants.

After the automatic and manual rejection phases for both *Speaking* and *No-Speaking* conditions, each participant's averaged cortical activity for trials without somatosensory stimulation (*no-stretch* trials) was subtracted from that for trials with such stimulation (*stretch* trials) for the same condition. This subtraction corrected for electrophysiological fluctuation associated with non-somatosensory activity (e.g., motor, cognitive, and visual processes).

Automatic data processing yielded a total of 11,748 clean epochs across all participants. A further 2.0% of epochs were removed during manual epoch inspection prior to the subtraction of *no-stretch* trials (not including the additional epochs removed to balance the number of epochs across conditions for the 3 participants with large discrepancies).

The final datasets for the Stuttering group had a median of 100 *Speaking* epochs (ranging from 45 to 107 with a SD of 14.7) and 99.5 *No-Speaking* epochs (range: 50 to 109; SD = 17.4) per participant. The final epoch counts for the Nonstuttering control group were comparable at a median of 101 *Speaking* epochs (range: 35 to 108; SD = 21.1) and 100 *No-Speaking* epochs (range: 35 to 108; SD = 20.4).

The final SEPs for the *Speaking* condition were then compared with those for the *No-Speaking* condition. Following previous electrophysiological studies involving a similar orofacial skin stretch, our analysis was restricted to the time window 100-200 ms post-stimulus onset (henceforth: N1; 63, 70). And given that the same previous studies with this paradigm have shown the largest responses occur over frontal midline electrodes (63, 70), we focused our analyses on a frontocentral region-of-interest (ROI). For each participant, peak amplitudes and latencies of the N1 component were extracted from the SEP averaged across 9 frontal electrodes (C11, C12, C13, C20, C21 [equivalent to Fz], C22, C24, C25, and C26 in BioSemi's ABC labeling system).

In addition, given the history of laterality in stuttering (71) and our previous finding that pre-speech auditory modulation is larger at central and left hemisphere ROIs (36), we also explored the effect of laterality. Therefore, two additional ROIs were investigated. Specifically, the peak amplitudes and latencies of the N1 components were also obtained from the average SEPs across 9 electrodes comprising a Left ROI (electrodes D10, D11, D12, D19 [equivalent to C3], D20, D21, D26, D27, and D28) and a Right ROI (electrodes B16, B17, B18, B22 [equivalent to C4], B23, B24, B29, B30, and B31).

Lastly, a measure of PSSM was calculated for each participant. Following the comparison of the three ROIs, the ROI with the largest N1 amplitudes was used for these calculations. Specifically, PSSM was calculated for each participant by subtracting the peak N1 amplitude in the *No-Speaking* condition from the peak N1 amplitude in the *Speaking* condition. Given that N1 amplitudes have negative polarity, positive PSSM values represent an N1 that is reduced in amplitude when the participant is preparing to speak.

4.5. Statistical analyses

All statistical analyses were conducted in R using RStudio (72) with the *rstatix* package (73). An alpha level of .05 was used to determine statistical significance. Visualizations were created using the *ggplot2* package (74). Grand average SEPs in the figures are shown with an additional low-pass filter applied at 15 Hz.

Peak N1 amplitude and latency were both used as dependent variables in two separate repeated-measures analyses of variance (rmANOVAs) with Condition (*No-Speaking*, *Speaking*) and ROI (Left, Frontal, Right) as within-subjects variables and Group (Stuttering, Nonstuttering) as a between-subjects variable. Both rmANOVAs tested all main effects and all interactions. The degrees of freedom for within-subjects factors were adjusted using the Huynh-Feldt correction to account for potential violations of the sphericity assumption (75). Partial omega-squared (ω_p^2)

was calculated as a measure of effect size (76, 77). Significant effects were followed up with *post hoc* pairwise Welch's *t*-tests adjusted for multiple comparisons using Holm-Bonferroni corrections (78).

For the Stuttering group, Pearson correlation coefficients were calculated to assess whether participants' stuttering frequency (i.e., mean percent stuttered syllables across the speaking and reading tasks from the SSI-4) or total SSI-4 score were related to their amount of PSSM.

5. RESULTS

Grand average SEPs for the *Speaking* and *No-Speaking* conditions by group and by ROI are shown in Figure 3. Group data as well as individual participant data for the N1 measures (peak amplitude and peak latency) are shown in boxplots in Figure 4.

5.1. N1 peak amplitude

For N1 amplitude, the main effects of Group ($F[1, 28] < 0.001$, $p = .982$, $\omega_p^2 < 0.01$) and Condition ($F[1, 28] < 0.001$, $p = .999$, $\omega_p^2 < 0.01$) were not statistically significant, and the Group × Condition interaction was also not significant ($F[1, 28] = 2.63$, $p = .116$, $\omega_p^2 < 0.01$). Thus, no PSSM was observed, and this was true for both groups.

The main effect of ROI was statistically significant ($F[1.98, 55.49] = 0.807$, $p < .001$, $\omega_p^2 = 0.25$), with N1 amplitudes at the Frontal ROI being larger than those at the Left ($t[29] = 11.5$, $p < .001$) and Right ROIs ($t[29] = -8.96$, $p < .001$). N1 amplitudes at the Right ROI were also significantly larger than those at the Left ROI ($t[29] = 2.24$, $p = .03$). None of the interactions with the ROI factor were statistically significant (all *F* values < 2.57, *p* values > .092, and partial omegas-squared < 0.01).

5.2. N1 peak latency

There were no significant main effects or interactions for N1 latency (all *F* values < 2.786, *p* values > .106, and partial omegas-squared < 0.02).

5.3. Pre-speech somatosensory modulation (PSSM) and stuttering severity

Individual participant PSSM values were calculated for the frontocentral ROI given that N1 peak amplitudes were largest over these midline electrodes. The distribution of individual participant modulation values for each group is shown in Figure 5.

Individual participant data for the correlation between calculated PSSM and the measures of stuttering severity and stuttering frequency are shown in Figure 6. Pearson correlation coefficients revealed no statistically significant relationship between PSSM and either total SSI score (r = -0.337, $\mathrm{CI}_{95\%}$ = [-0.724, 0.212], p = .219) or mean stuttering frequency (r = -0.337, $\mathrm{CI}_{95\%}$ = [-0.724, 0.212], p = .220).

6. DISCUSSION

Several prior studies had shown that typically fluent adult speakers, but not adults who stutter, show a modulation of auditory processing during the movement planning phase immediately prior to speech onset (40). In the current study, we extended the same paradigm to investigate whether this between-group difference in pre-speech sensory modulation is also present in the somatosensory domain. Participants who stutter and matched participants who do not stutter completed a delayed-response speaking task during which a rapid, upward orofacial skin stretch was applied 200 ms before the predictable ‘go’ signal. In a no-speaking control task, the same stimulus was applied at the same timepoint during silent reading of the words. Participants wore an EEG cap and SEPs in response to the skin stretches were recorded during both conditions.

The data revealed, for all participants, large SEPs with a clear N1 component that peaked between 100-200 ms after stimulation onset. The SEPs lacked a clear P2 component, in contrast with those reported for a similar orofacial skin stretch in previous work by Ito and colleagues (60, 63, 70). However, SEP morphology was consistent both within and between groups, suggesting that the absence of a clear P2 component here is not a random artefact but, more likely, a result of our specific implementation of the somatosensory stimuli, the EEG processing pipeline (e.g., we used an average-mastoids reference whereas the aforementioned prior studies use a whole-head average reference).

Most importantly, in the present study, peak amplitudes and latencies of the N1 component evoked by facial stimulation were not statistically significantly different between the *Speaking* and *No-Speaking* conditions and this was true for both the stuttering and the nonstuttering groups. In other words, the data revealed no N1 modulation during speech movement planning as compared with a silent reading control condition. Given that even the nonstuttering participants showed no somatosensory modulation, there was also no difference in modulation between this group and the group of stuttering participants. Moreover, the stuttering and nonstuttering participants did not differ in topography of the SEP N1 amplitudes, which were larger in the frontal region than in the right and left regions (consistent with 63, 70), and larger over the right hemisphere than over the left hemisphere. Lastly, within the group of stuttering participants, there was no statistically significant correlation between individuals’ amount of PSSM (difference in N1 amplitude between the *Speaking* and *No-Speaking* conditions) and clinical measures of either stuttering frequency (in percent stuttered syllables) or overall stuttering severity (SSI-4 score).

The absence of PSSM in the present paradigm may have at least three possible explanations: (a) true differences in the neural processing of somatosensory versus auditory input during speech movement planning, (b) a methodological need to adjust stimulus parameters when using somatosensory stimulation, or (c) task relevance-related specificity in terms of which somatosensory inputs are modulated for speech. First, our results leave open the possibility that there is an actual difference in how somatosensory input is processed during speech movement planning (i.e., no modulation of response magnitude) as compared with auditory input (i.e., modulated response magnitude). That is, there may be no physiological change in the cortical processing of somatosensory input when an individual is preparing motor commands to initiate speech. This seems unlikely, however, in light of ample evidence supporting the notion of somatosensory modulation for both nonspeech and speech movements. For example, somatosensory input is certainly modulated *during* human limb movement execution (46, 54, 57, 59, 79–82). Additionally, research with nonhuman primates has documented somatosensory modulation both during and *prior to* active limb movements (47, 48). Moreover, specifically for speech, it has been shown that lip muscle reflexes evoked by a mechanical posterolateral stretch at the corner of the mouth are suppressed at the time of speech initiation, consistent with the notion that movement preparation suppresses somatosensory input (83). More recently, an EEG study found that SEPs evoked by posterolateral stretch of the skin next to the corners of the mouth were modulated in a lip configuration-specific manner during sustained vowel production (60), and a psychophysics study found that participants perceived electrical stimulation of the lower lip as less intense when it was applied during speech movement planning as compared with not speaking (84).

Second, the absence of PSSM in the present paradigm may be due to the selection of less-than-optimal parameters for stimulus direction, time of application, or duration. With regard to direction, we applied an upward orofacial skin stretch (70, 85) whereas a posterolateral stretch was used in a study that did find lip configuration-specific SEP modulation during sustained vowels (60). Given that our speech task involved real word productions with jaw lowering and sometimes lip spreading movements, it is possible that a posterolateral or downward stretch of the skin at the corners of the mouth would have yielded different results. Similarly, the timing of stimulation onset was adopted from our prior work with an analogous auditory paradigm but this timing, and stimulus duration, may require specific modifications for somatosensory stimuli. For example, the auditory probe tones that we used previously were only 40 ms in duration whereas actual implementation of the present orofacial skin stretches lasted for almost the entire delay period before the ‘go’ signal. In other words, after being stretched upward, the resting mouth position was often not restored until very close to the ‘go’ signal. Furthermore, *if* PSSM is occurring, its time course may be substantially different from that observed for PSAM. In the aforementioned psychophysics study with electrical stimulation of the lip during speech movement planning, the greatest perceptual modulation was observed only 60 ms prior to the onset of speech and no significant modulation occurred until approximately 200 ms before speech onset (84). Consequently, our stimulation starting 200 ms prior to the ‘go’ signal (on average ~500-600 ms prior to the actual onset of speech) may have been too early to capture PSSM.

Third, the fact that no PSSM was detected in the present study may relate to a high degree of specificity in terms of modulating task relevant somatosensory inputs for speech motor control. In fact, this hypothesis is consistent with the findings from all three previous studies that yielded evidence for the occurrence of PSSM: McClean and Clay (83) observed a suppression of mechanically evoked upper and lower *lip* reflexes at the onset of productions of the bilabial consonant /p/; McGuffin et al. (84) found modulated perception of electrical stimulation to the low *lip* before the initiation of words with bilabial consonants /p, b, m/ in word onset position; and Ito et al. (60) found that the modulation of SEPs to posterolateral stretch at the corner of the mouth (i.e., inducing a more “spread” lip position) depended on the specific *lip* configuration for different vowels. We therefore propose that future studies aiming to gain a better understanding of somatosensory integration in speech movement planning and execution may benefit from stimulating somatosensory receptors that will transduce afferent signals generated by the effectors that are directly involved in performing the articulatory gestures for the targeted speech sounds. Thus, directly stimulating structures such as the lips or the tongue during speech movement preparation for bilabial or lingual sounds might be an optimal approach (86–88).

7. CONCLUSIONS AND FUTURE DIRECTIONS

In conclusion, we examined the occurrence of PSSM in both typical speakers and individuals who stutter by extending to the somatosensory domain a well-established paradigm that previously revealed *auditory* modulation during speech planning in typical speakers but not individuals who stutter. A robotic device applied orofacial skin stretches that elicited SEPs during speech movement planning and a no-speaking control condition. For the specific type, duration, and timeline of stimulation applied here, we found no evidence of PSSM in either group of participants. Both groups showed clear N1 peaks in response to the orofacial skin stretch but there was no difference in peak amplitude between the speaking and control conditions for either group. Future studies can be designed to distinguish among the possibilities that the absence of PSSM in the present paradigm resulted from true differences in the neural processing of somatosensory versus auditory input during speech movement planning, the use of suboptimal stimulus parameters, or specificity in the central nervous system's modulation of task relevant somatosensory inputs for speech motor control.

8. DATA AVAILABILITY

Outcome measures used in our analyses (including participant demographics) are openly available via an Open Science Framework (OSF) repository at https://osf.io/wr5hc/. To protect the identity of the participants in this study, no speech signals or raw data are openly shared.

9. ACKNOWLEDGMENTS

The authors would first and foremost like to thank our research participants for contributing to this study and advancing our understanding of stuttering, especially those who shaved their facial hair to participate. Additionally, we would like to thank Hannah Goldfarb, Zoie Breland, Priyanka Talur, Lisa Paroni, and Daria D'Alessandro for assisting with electrode cap preparation and/or assisting with manual inspection of epochs.

10. GRANTS

This research was supported by the National Institute on Deafness and Other Communication Disorders of the National Institutes of Health (R01DC017444 and R01DC020162). This content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or the National Institute on Deafness and Other Communication Disorders.

11. DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the authors.

12. AUTHOR CONTRIBUTIONS

L.M., T.I., and E.L. conceived and designed the research. E.L. performed experiments, analyzed data, prepared figures, and drafted the manuscript. All authors interpreted the results of the experiments, edited and revised the manuscript, and approved the final version of the manuscript.

13. REFERENCES

1. **Guenther FH**, **Ghosh SS**, **Tourville JA**. Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain Lang* 96: 280–301, 2006. doi: 10.1016/J.BANDL.2005.06.001.

2. **Hickok G**, **Houde J**, **Rong F**. Sensorimotor integration in speech processing: Computational basis and neural organization. *Neuron* 69: 407–422, 2011. doi: 10.1016/j.neuron.2011.01.019.

3. **Parrell B**, **Houde JF**. Modeling the role of sensory feedback in speech motor control and learning. *J Speech Lang Hear Res* 62: 2963–2985, 2019. doi: 10.1044/2019_JSLHR-S-CSMC7-18-0127.

4. **Shadmehr R**. Learning to predict and control the physics of our movements. *J Neurosci* 37: 1663–1671, 2017. doi: 10.1523/JNEUROSCI.1675-16.2016.

5. **Feng Y**, **Gracco VL**, **Max L**. Integration of auditory and somatosensory error signals in the neural control of speech movements. *J Neurophysiol* 106: 667–679, 2011. doi: 10.1152/jn.00638.2010.

6. **Lametti DR**, **Nasir SM**, **Ostry DJ**. Sensory preference in speech production revealed by simultaneous alteration of auditory and somatosensory feedback. *J Neurosci* 32: 9351–9358, 2012. doi: https://doi.org/10.1523/JNEUROSCI.0404-12.2012.

7. **Caudrelier T**, **Rochet-Capellan A**. Changes in speech production in response to formant perturbations: An overview of two decades of research. In: *Speech Production and Perception: Learning and memory*, edited by Fuchs S, Cleland J, Rochet-Capellan A. Peter Lang, 2019, p. 15–75.

8. **Perkell JS**. Five decades of research in speech motor control: What have we learned, and where should we go from here? *J Speech Lang Hear Res* 56: 1857–1874, 2013. doi: 10.1044/1092-4388(2013/12-0382).

9. **Bradshaw AR**, **Lametti DR**, **McGettigan C**. The role of sensory feedback in developmental stuttering: a review. *Neurobiol Lang* 2: 308–334, 2021. doi: 10.1162/nol_a_00036.

10. **Kent RD**, **Martin RE**, **Sufit RL**. Oral sensation: a review and clinical prospective. In: *Human Communication and its Disorders*. 1990, p. 135–191.

11. **Kent RD**. The feel of speech: Multisystem and polymodal somatosensation in speech production. *J Speech Lang Hear Res* 67: 1424–1460, 2024. doi: 10.1044/2024_JSLHR-23-00575.

12. **Ito T**, **Ostry DJ**. Somatosensory contribution to motor learning due to facial skin deformation. *J Neurophysiol* 104: 1230–1238, 2010. doi: 10.1152/jn.00199.2010.

13. **Nasir SM**, **Ostry DJ**. Somatosensory precision in speech production. *Curr Biol* 16: 1918–1923, 2006. doi: 10.1016/j.cub.2006.07.069.

14. **Tremblay S**, **Shiller DM**, **Ostry DJ**. Somatosensory basis of speech production. *Nature* 423: 866–869, 2003. doi: 10.1038/nature01710.

15. **Andrews G**, **Howie PM**, **Dozsa M**, **Guitar BE**. Stuttering: Speech pattern characteristics under fluency-inducing conditions. *J Speech Hear Res* 25: 208–216, 1982. doi: 10.1044/jshr.2502.208.

16. **Lotzmann G**. Zur anwendung variierter verzögerungszeiten bei balbuties. *Folia Phoniatr* 13: 276–312, 1961.

17. **Soderberg GA**. Delayed auditory feedback and the speech of stutterers: A review of studies. *J Speech Hear Disord* 34: 20–29, 1969. doi: 10.1044/jshd.3401.20.

18. **Webster RL**, **Schumacher SJ**, **Lubker BB**. Changes in stuttering frequency as a function of various intervals of delayed auditory feedback. *J Abnorm Psychol* 75: 45–49, 1970.

19. **Cherry C**, **Sayers BMcA**. Experiments upon the total inhibition of stammering by external control, and some clinical results. *J Psychosom Res* 1: 233–246, 1956.

20. **Sutton S**, **Chase RA**. White noise and stuttering. *J Speech Hear Res* 4: 72–72, 1961. doi: 10.1044/jshr.0401.72.

21. **Howell P**, **El-Yaniv N**, **Powell DJ**. Factors affecting fluency in stutterers when speaking under altered auditory feedback. In: *Speech Motor Dynamics in Stuttering*, edited by Peters HFM, Hulstijn W. 1987, p. 361–369.

22. **Martin RR**, **Siegel GM**, **Johnson LJ**, **Haroldson SK**. Sidetone amplification, noise, and stuttering. *J Speech Lang Hear Res* 27: 518–527, 1984.

23. **Fox PT**, **Ingham RJ**, **Ingham JC**, **Hirsch TB**, **Downs JH**, **Martin C**, **Jerabek P**, **Glass T**, **Lancaster JL**. A PET study of the neural systems of stuttering. *Nature* 382: 158–162, 1996. doi: 10.1038/382158a0.

24. **Belyk M**, **Kraft SJ**, **Brown S**. Stuttering as a trait or a state revisited: motor system involvement in persistent developmental stuttering. *Eur J Neurosci* 45: 622–624, 2017. doi: 10.1111/EJN.13512.

25. **Neef NE**, **Anwander A**, **Friederici AD**. The neurobiological grounding of persistent stuttering: From structure to function. *Curr Neurol Neurosci Rep* 15, 2015. doi: 10.1007/s11910-015-0579-4.

26. **Budde KS**, **Barron DS**, **Fox PT**. Stuttering, induced fluency, and natural fluency: A hierarchical series of activation likelihood estimation meta-analyses. *Brain Lang* 139: 99–107, 2014. doi: 10.1016/j.bandl.2014.10.002.

27. **Cai S**, **Beal DS**, **Ghosh SS**, **Tiede MK**, **Guenther FH**, **Perkell JS**. Weak responses to auditory feedback perturbation during articulation in persons who stutter: Evidence for abnormal auditory-motor transformation. *PLoS ONE* 7: 1–13, 2012. doi: 10.1371/journal.pone.0041830.

28. **Daliri A**, **Wieland EA**, **Cai S**, **Guenther FH**, **Chang S-E**. Auditory-motor adaptation is reduced in adults who stutter but not in children who stutter. *Dev Sci* 21: e12521, 2018. doi: 10.1111/desc.12521.

29. **Kim KS**, **Max L**. Speech auditory-motor adaptation to formant-shifted feedback lacks an explicit component: Reduced adaptation in adults who stutter reflects limitations in implicit sensorimotor learning. *Eur J Neurosci* 53: 3093–3108, 2021. doi: 10.1111/ejn.15175.

30. **Kim KS**, **Daliri A**, **Flanagan JR**, **Max L**. Dissociated development of speech and limb sensorimotor learning in stuttering: Speech auditory-motor learning is impaired in both children and adults who stutter. *Neuroscience* 451: 1–21, 2020. doi: 10.1016/j.neuroscience.2020.10.014.

31. **Chang S-E**, **Erickson KI**, **Ambrose NG**, **Hasegawa-Johnson MA**, **Ludlow CL**. Brain anatomy differences in childhood stuttering. *NeuroImage* 39: 1333–1344, 2008. doi: 10.1016/j.neuroimage.2007.09.067.

32. **Koenraads SPC**, **El Marroun H**, **Muetzel RL**, **Chang S-E**, **Vernooij MW**, **Baatenburg de Jong RJ**, **White T**, **Franken MC**, **van der Schroeff MP**. Stuttering and gray matter morphometry: A population-based neuroimaging study in young children. *Brain Lang* 194: 121–131, 2019. doi: 10.1016/J.BANDL.2019.04.008.

33. **Neef NE**, **Chang S-E**. Knowns and unknowns about the neurobiology of stuttering [Online]. https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.3002492.

34. **McClean MD**. Lip-muscle reflexes during speech movement preparation in stutterers. *J Fluen Disord* 21: 49–60, 1996. doi: 10.1016/0094-730X(95)00058-F.

35. **Gracco VL**, **Sares AG**, **Koirala N**. Structural brain network topological alterations in stuttering adults. *Brain Commun* 4: fcac058, 2022. doi: 10.1093/BRAINCOMMS/FCAC058.

36. **Daliri A**, **Max L**. Modulation of auditory processing during speech movement planning is limited in adults who stutter. *Brain Lang* 143: 59–68, 2015. doi: 10.1016/j.bandl.2015.03.002.

37. **Daliri A**, **Max L**. Modulation of auditory responses to speech vs. nonspeech stimuli during speech movement planning. *Front Hum Neurosci* 10: 1–9, 2016. doi: 10.3389/fnhum.2016.00234.

38. **Daliri A**, **Max L**. Electrophysiological evidence for a general auditory prediction deficit in adults who stutter. *Brain Lang* 150: 37–44, 2015. doi: 10.1016/j.bandl.2015.08.008.

39. **Daliri A**, **Max L**. Stuttering adults’ lack of pre-speech auditory modulation normalizes when speaking with delayed auditory feedback. *Cortex* 99: 55–68, 2018. doi: 10.1016/j.cortex.2017.10.019.

40. **Max L**, **Daliri A**. Limited pre-speech auditory modulation in individuals who stutter: Data and hypotheses. *J Speech Lang Hear Res* 62: 3071–3084, 2019. doi: 10.1044/2019_JSLHR-S-CSMC7-18-0358.

41. **Godey B**, **Schwartz D**, **de Graaf JB**, **Chauvel P**, **Liégeois-Chauvel C**. Neuromagnetic source localization of auditory evoked fields and intracerebral evoked potentials: a comparison of data in the same patients. *Clin Neurophysiol* 112: 1850–1859, 2001. doi: 10.1016/S1388-2457(01)00636-8.

42. **Näätänen R**, **Picton T**. The N1 wave of the human electric and magnetic response to sound: A review and an analysis of the component structure. *Psychophysiology* 24: 375–425, 1987. doi: 10.1111/j.1469-8986.1987.tb00311.x.

43. **Hazemann P**, **Audin G**, **Lille F**. Effect of voluntary self-paced movements upon auditory and somatosensory evoked potentials in man. *Electroencephalogr Clin Neurophysiol* 39: 247–254, 1975. doi: 10.1016/0013-4694(75)90146-7.

44. **Numminen J**, **Salmelin R**, **Hari R**. Subject’s own speech reduces reactivity of the human auditory cortex. *Neurosci Lett* 265: 119–122, 1999. doi: 10.1016/S0304-3940(99)00218-9.

45. **Houde JF**, **Nagarajan SS**, **Sekihara K**, **Merzenich MM**. Modulation of the auditory cortex during speech: An MEG study. *J Cogn Neurosci* 14: 1125–1138, 2002. doi: 10.1162/089892902760807140.

46. **Blakemore S-J**, **Wolpert DM**, **Frith CD**. Central cancellation of self-produced tickle sensation. *Nat Neurosci 1998 17* 1: 635–640, 1998. doi: 10.1038/2870.

47. **Seki K**, **Fetz EE**. Gating of sensory input at spinal and cortical levels during preparation and execution of voluntary movement. *J Neurosci* 32: 890–902, 2012. doi: 10.1523/JNEUROSCI.4958-11.2012.

48. **Seki K**, **Perlmutter SI**, **Fetz EE**. Sensory input to primate spinal cord is presynaptically inhibited during voluntary movement. *Nat Neurosci* 6: 1309–1316, 2003. doi: 10.1038/nn1154.

49. **Chang EF**, **Niziolek CA**, **Knight RT**, **Nagarajan SS**, **Houde JF**. Human cortical sensorimotor network underlying feedback control of vocal pitch. *Proc Natl Acad Sci U S A* 110: 2653–2658, 2013. doi: 10.1073/PNAS.1216827110/-/DCSUPPLEMENTAL.

50. **Numminen J**, **Curio G**. Differential effects of overt, covert and replayed speech on vowel-evoked responses of the human auditory cortex. *Neurosci Lett* 272: 29–32, 1999.

51. **Eliades SJ**, **Wang X**. Sensory-motor interaction in the primate auditory cortex during self-initiated vocalizations. *J Neurophysiol* 89: 2194–2207, 2003. doi: 10.1152/jn.00627.2002.

52. **Eliades SJ**, **Wang X**. Corollary discharge mechanisms during vocal production in marmoset monkeys. *Biol Psychiatry Cogn Neurosci Neuroimaging* 4: 805–812, 2019. doi: 10.1016/J.BPSC.2019.06.008.

53. **Suga N**, **Schlegel P**. Neural attenuation of responses to emitted sounds in echolocating bats. *Science* 177: 82–84, 1972.

54. **Starr A**, **Cohen LG**. ‘Gating’ of somatosensory evoked potentials begins before the onset of voluntary movement in man. *Brain Res* 348: 183–186, 1985. doi: 10.1016/0006-8993(85)90377-4.

55. **Singla S**, **Dempsey C**, **Warren R**, **Enikolopov AG**, **Sawtell NB**. A cerebellum-like circuit in the auditory system cancels responses to self-generated sounds. *Nat Neurosci 2017 207* 20: 943–950, 2017. doi: 10.1038/nn.4567.

56. **Nelson A**, **Schneider DM**, **Takatoh J**, **Sakurai K**, **Wang F**, **Mooney R**. A Circuit for Motor Cortical Modulation of Auditory Cortical Activity. *J Neurosci* 33: 14342–14353, 2013. doi: 10.1523/JNEUROSCI.2275-13.2013.

57. **Cohen LG**, **Starr A**. Localization, timing and specificity of gating of somatosensory evoked potentials during active movement in man. *Brain* 110: 451–467, 1987.

58. **Shimazu H**, **Kaji R**, **Murase N**, **Kohara N**, **Ikeda A**, **Shibasaki H**, **Kimura J**, **Rothwell JC**. Pre-movement gating of short-latency somatosensory evoked potentials. *NeuroReport* 10: 2457–2460, 1999.

59. **Wasaka T**, **Hoshiyama M**, **Nakata H**, **Nishihira Y**, **Kakigi R**. Gating of somatosensory evoked magnetic fields during the preparatory period of self-initiated finger movement. *NeuroImage* 20: 1830–1838, 2003. doi: 10.1016/S1053-8119(03)00442-7.

60. **Ito T**, **Ohashi H**, **Gracco VL**. Changes of orofacial somatosensory attenuation during speech production. *Neurosci Lett* 730: 135045, 2020. doi: 10.1016/j.neulet.2020.135045.

61. **Riley G**. SSI-4: Stuttering severity instrument. Pro-Ed: 2009.

62. **Cohen J**. A coefficient of agreement for nominal scales. *Educ Psychol Meas* 20: 37–46, 1960. doi: 10.1177/001316446002000104.

63. **Ito T**, **Ostry DJ**, **Gracco VL**. Somatosensory event-related potentials from orofacial skin stretch stimulation. *J Vis Exp* 2015: 53621, 2015. doi: 10.3791/53621.

64. **Oostenveld R**, **Praamstra P**. The five percent electrode system for high-resolution EEG and ERP measurements. *Clin Neurophysiol* 112: 713–719, 2001. doi: 10.1016/S1388-2457(00)00527-7.

65. **Brainard DH**. The Psychophysics Toolbox. *Spat Vis* 10: 433–436, 1997.

66. **Kleiner M**, **Brainard DH**, **Pelli DG**. What’s new in Psychtoolbox-3? In: *ECVP Abstract Supplement*. Perception. 2007.

67. **Pelli DG**. The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spat Vis* 10: 437–442, 1997.

68. **Delorme A**, **Makeig S**. EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J Neurosci Methods* 134: 9–21, 2004. doi: 10.1016/j.jneumeth.2003.10.009.

69. **Delorme A**. EEG is better left alone. *Sci Rep* 13: 2372, 2023. doi: 10.1038/s41598-023-27528-0.

70. **Ito T**, **Coppola JH**, **Ostry DJ**. Speech motor learning changes the neural response to both auditory and somatosensory signals. *Sci Rep* 6: 25926, 2016. doi: 10.1038/srep25926.

71. **Travis LE**. The cerebral dominance theory of stuttering: 1931–1978. *J Speech Hear Disord* 43: 278–281, 1978. doi: 10.1044/jshd.4303.278.

72. **RStudio Team**. RStudio: Integrated Development Environment for R [Online]. RStudio, PBC.: 2020. www.rstudio.com.

73. **Kassambara A**. rstatix: Pipe-friendly framework for basic statistical tests [Online]. 2023. https://CRAN.R-project.org/package=rstatix.

74. **Wickham H**. Ggplot2: elegant graphics for data analysis [Online]. Springer-Verlag New York: 2016. https://ggplot2.tidyverse.org.

75. **Max L**, **Onghena P**. Some issues in the statistical analysis of completely randomized and repeated measures designs for speech, language, and hearing research. *J Speech Lang Hear Res* 42: 261–270, 1999. doi: 10.1044/jslhr.4202.261.

76. **Okada K**. Is omega squared less biased? A comparison of three major effect size indices in one-way anova. *Behaviormetrika* 40: 129–147, 2013. doi: 10.2333/bhmk.40.129.

77. **Lakens D**. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Front Psychol* 0: 863, 2013. doi: 10.3389/FPSYG.2013.00863.

78. **Holm S**. A simple sequentially rejective multiple test procedure. *Scand J Stat* 6: 65–70, 1979.

79. **Bays PM**, **Flanagan JR**, **Wolpert DM**. Attenuation of self-generated tactile sensations is predictive, not postdictive. *PLoS Biol* 4: 281–284, 2006. doi: 10.1371/journal.pbio.0040028.

80. **Haggard P**, **Whitford B**. Supplementary motor area provides an efferent signal for sensory suppression. *Cogn Brain Res* 19: 52–58, 2004. doi: 10.1016/J.COGBRAINRES.2003.10.018.

81. **Shergill SS**, **Bays PM**, **Frith CD**, **Wotpert DM**. Two eyes for an eye: The neuroscience of force escalation. *Science* 301: 187, 2003. doi: 10.1126/science.1085327.

82. **Voss M**, **Ingram JN**, **Haggard P**, **Wolpert DM**. Sensorimotor attenuation by central motor command signals in the absence of movement. *Nat Neurosci* 9: 26–27, 2006. doi: 10.1038/nn1592.

83. **McClean MD**, **Clay JL**. Evidence for suppression of lip muscle reflexes prior to speech. *Exp Brain Res* 97: 541–544, 1994. doi: 10.1007/BF00241547.

84. **McGuffin BJ**, **Liss JM**, **Daliri A**. The orofacial somatosensory system is modulated during speech planning and production. *J Speech Lang Hear Res* 63: 2637–2648, 2020. doi: 10.1044/2020_JSLHR-19-00318.

85. **Ito T**, **Tiede M**, **Ostry DJ**. Somatosensory function in speech perception. *Proc Natl Acad Sci* 106: 1245–1248, 2009. doi: 10.1073/pnas.0810063106.

86. **Ito T**, **Bouguerra M**, **Bourhis M**, **Perrier P**. Tongue reflex for speech posture control. *Sci Rep* 14: 6386, 2024.

87. **Bourhis M**, **Perrier P**, **Savariaux C**, **Ito T**. Quick speech motor correction in the absence of auditory feedback. *Front Hum Neurosci* 18, 2024. doi: 10.3389/fnhum.2024.1399316.

88. **Ito T**, **Szabados A**, **Caillet J-L**, **Perrier P**. Quick compensatory mechanisms for tongue posture stabilization during speech production. *J Neurophysiol* 123: 2491–2503, 2020. doi: 10.1152/jn.00756.2019.

14. FIGURE LEGENDS

Figure 1

Schematic representation of the general skin stretch setup for recording somatosensory-evoked potentials during speech movement planning.

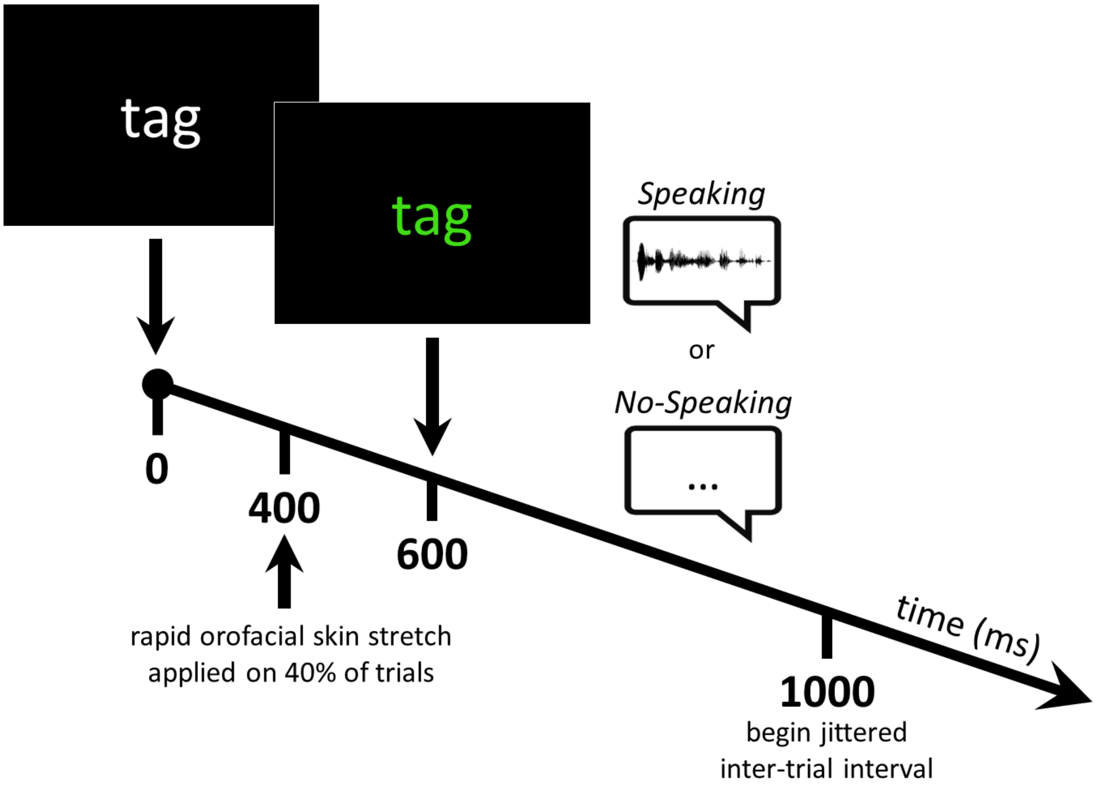

Figure 2

Schematic representation of one trial. In the *Speaking* condition, participants produced a monosyllabic word after the ‘go’ signal (i.e., when the text turned green), 600 ms after the initial appearance of the stimulus in white text. In the *No-Speaking* condition, participants refrained from overtly producing the stimulus (i.e., they ignored the ‘go’ signal). On 40% of the trials in all conditions, a rapid orofacial skin stretch was applied to the corners of the participant’s mouth (*stretch* trials). This somatosensory stimulus was presented 200 ms prior to the ‘go’ signal to probe somatosensory processing mechanisms during speech movement planning. Four hundred ms after the ‘go’ signal, a short inter-trial interval began.

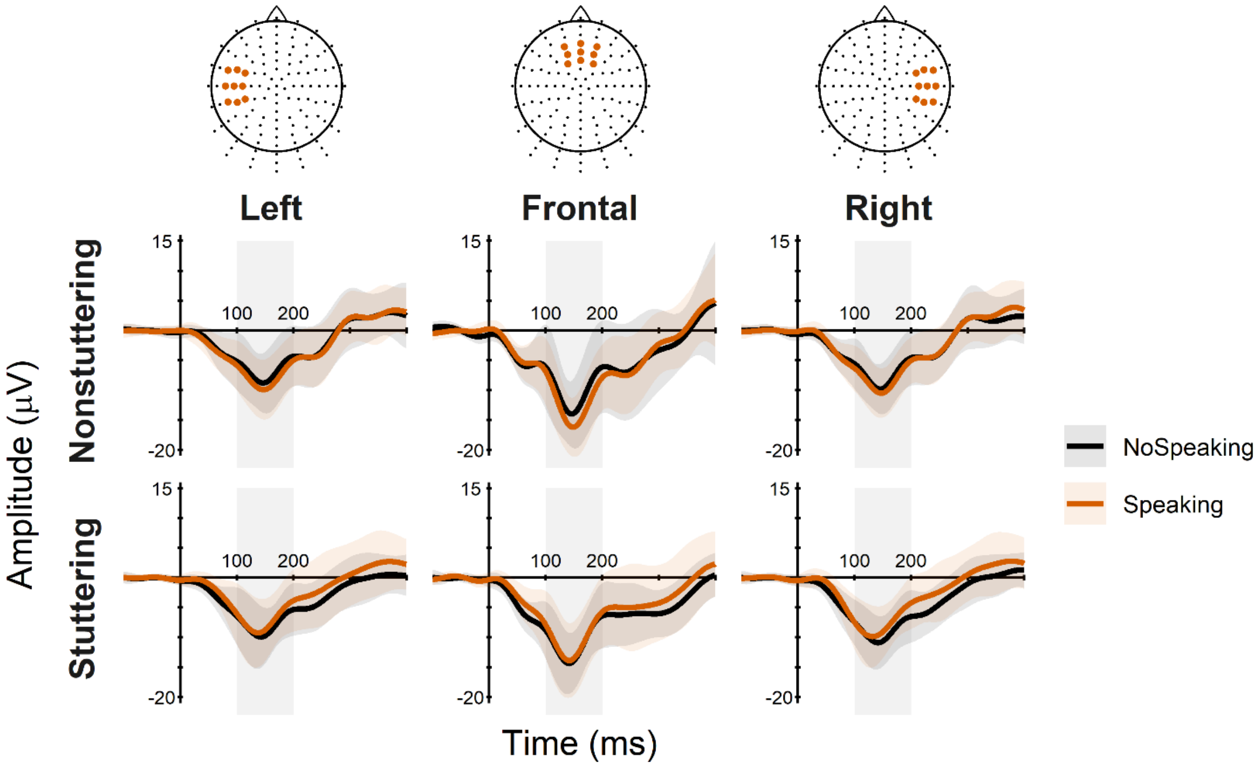

Figure 3

Grand average SEPs for both the Nonstuttering (*N* = 15 adults; top row) and Stuttering (*N* = 15 adults; bottom row) groups for each ROI: Left (leftmost column), Frontal (middle column), and Right (rightmost column). The mean evoked responses elicited by the applied orofacial skin stretch during speech movement planning are shown in orange (*Speaking* condition), while the control condition with no movement planning is shown in black (*No-Speaking* condition). The shaded ribbons indicate one standard deviation around the mean. The light gray shaded box represents the N1 analysis window from 100 to 200 ms post-stimulus onset. The electrode caps above each column highlight the nine electrodes used for each ROI analysis.

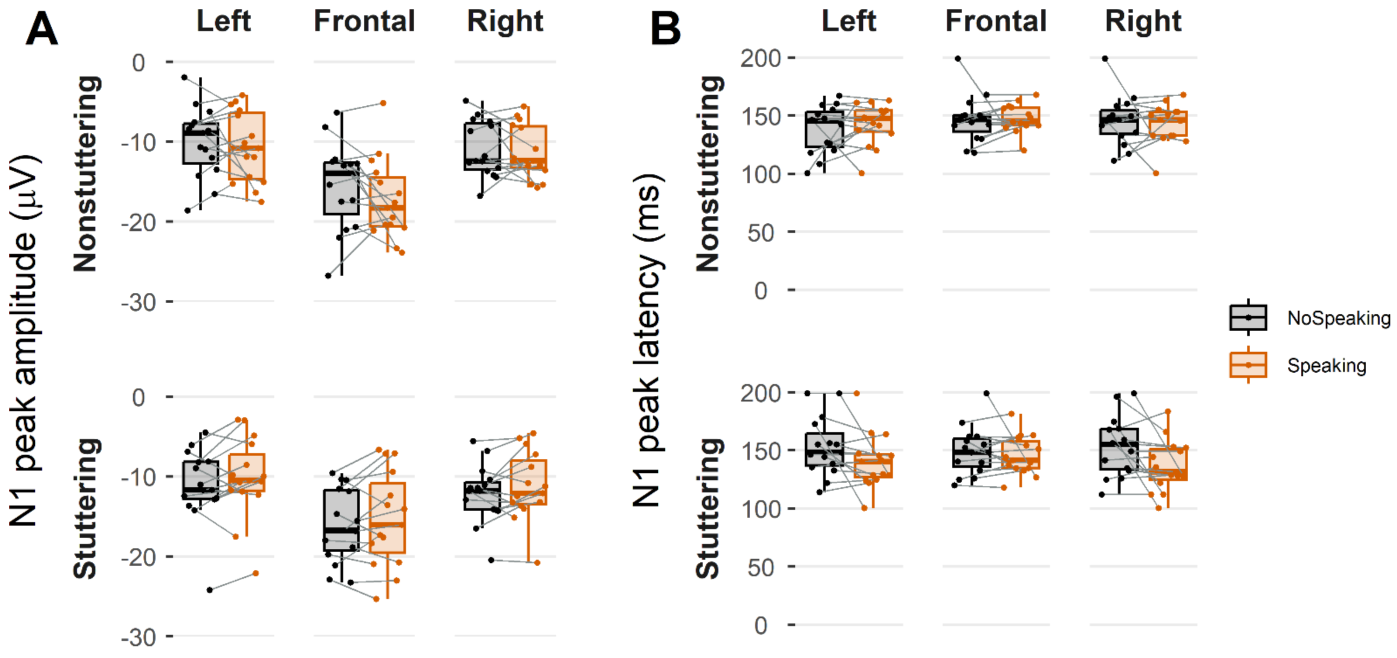

Figure 4

Summary of N1 component parameters for all participants across the three ROIs. Data are shown for both the Nonstuttering group (*N* = 15 adults; top row) and the Stuttering group (*N* = 15 adults; bottom row). Boxplots are shown for both peak amplitude data (A, left) and peak latency data (B, right). Individual points correspond to values from single participants, with points connected with a line to show individual performance across conditions (*Speaking* vs. *No-Speaking*). For each boxplot, the solid horizontal line represents the group median value, with box edges representing the interquartile range (25th to 75th percentiles) and whiskers extending to the most extreme values within 1.5 times the interquartile range.

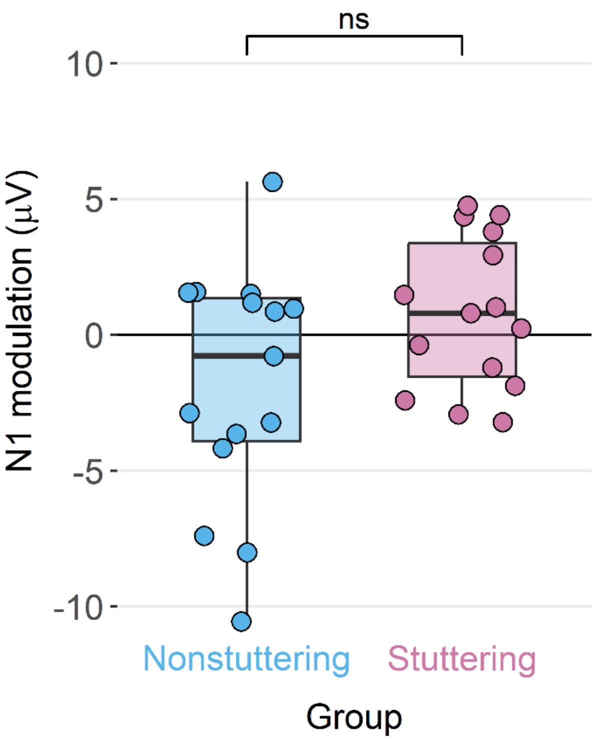

Figure 5

Pre-speech somatosensory modulation values summarized for all participants. Boxplots are shown for both the Nonstuttering (*N* = 15 adults; left) and Stuttering (*N* = 15 adults; right) groups. Each colored dot corresponds to the amount of modulation for one participant. Boxplot details as in Figure 4. Positive amounts of modulation represent an N1 peak that is smaller in amplitude when the participant is preparing to speak.

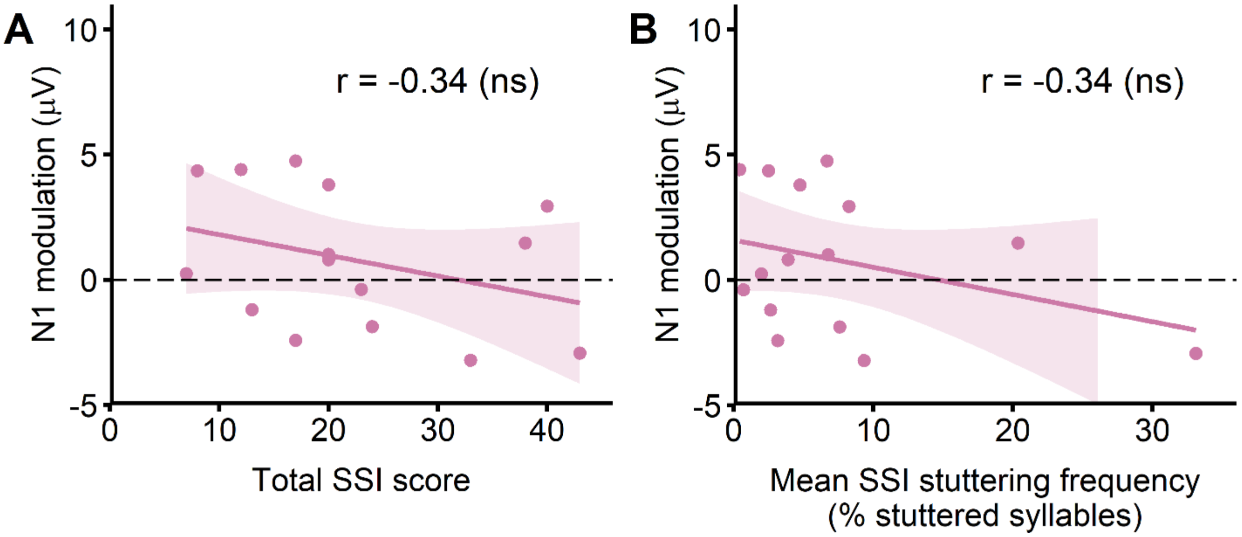

Figure 6

Correlation plots between two stuttering severity proxies and the individual amounts of pre-speech somatosensory modulation in the Stuttering group (*N* = 15 adults). Neither the correlation between N1 modulation and the total SSI score (A; *p* = .219) nor the correlation between N1 modulation and the mean percentage of stuttered syllables across the SSI-4 tasks (B; *p* = .220) were statistically significant. For both A and B, lines and shaded regions represent linear regression fits with 95% confidence intervals. Each data point corresponds to one individual participant.

15. TABLES

Table 1. *Demographic information for all stuttering participants, sorted by their same-day mean stuttering frequency across both SSI tasks.*

Participant ID	Sex assigned at birth	Gender	Age (years)	Handedness	Reported age of onset (years)	Self-identifies as an IWS	SSI score	SSI severity	Mean stuttering frequency (%SS)
1	Male	Male	25	Right	3 to 4	Yes	12	very mild	0.4
2	Female	Female	25	Right	"always"	Yes	23	mild	0.7
3	Male	Male	34	Right	4	Yes	7	below cutoff	2.0
4	Male	Male	40	Right	3	Yes	8	below cutoff	2.5
5	Female	Nonbinary	38	Right	2	Yes	13	very mild	2.6
6	Female	Female	19	Right	"when I began talking"	Yes	17	very mild	3.1
7	Male	Male	28	Right	8	Yes	20	mild	3.9
8	Male	Male	26	Right	"extremely young"	Yes	20	mild	4.8
9	Male	Male	37	Left	4	Yes	17	very mild	6.7
10	Male	Male	43	Right	3	Yes	20	mild	6.8
11	Male	Male	57	Right	"before preschool"	Yes	24	mild	7.6
12	Male	Male	42	Right	6 to 8	Yes	40	very severe	8.3
13	Male	Male	20	Right	"young enough I can't remember"	Yes	33	severe	9.3
14	Male	Male	23	Right	6	Yes	38	very severe	20.4
15	Male	Male	23	Right	5	Yes	43	very severe	33.1

IWS = Individual Who Stutters. SSI = Stuttering Severity Instrument, 4th edition. %SS = percent stuttered syllables.

Section V: Preliminary insights into the functional relevance of pre-speech auditory modulation: Correlations with kinematic estimates of feedback-based control in adults who stutter (*Study 3*)

(Manuscript in preparation)

Elise LeBovidge[1]

Kwang S. Kim[2]

Ludo Max[1]

[1]Department of Speech and Hearing Sciences, University of Washington,
Seattle, Washington, USA

[2]Department of Speech, Language, and Hearing Sciences, Purdue University,
West Lafayette, Indiana, USA

1. ABSTRACT

There is ample evidence that auditory cortical responses are modulated even prior to speech onset. However, the functional relevance of such pre-speech auditory modulation (PSAM) is still unclear. Recent evidence that PSAM correlates with reduced acoustic variability at the very onset of speech production suggests that this phenomenon may reflect ongoing feedforward mechanisms (e.g., fine-tuning the outgoing motor command) rather than the previously hypothesized general preparation in advance of speech feedback monitoring. This study is the first investigation to directly correlate, across individuals, a speaker's extent of PSAM with estimates of the same speaker's reliance on feedforward vs. feedback-based control. Two matched groups of participants (6 adults who stutter and 6 adults who do not stutter) completed two tasks: 1) an electroencephalographic (EEG) task in which probe tones were presented just prior to the onset of monosyllabic words and 2) an electromagnetic articulography task to record speech kinematics during the production of the same words. The EEG task allowed us to obtain individual PSAM measurements by comparing auditory evoked potentials recorded during speech movement planning with those recorded during a silent reading control condition. The kinematic task allowed us to estimate, for each speaker, the contribution of feedforward vs. feedback control systems using a previously established statistical model. Preliminary results suggest that the extent of PSAM is positively correlated with individual estimates of feedforward control and negatively correlated with estimates of feedback control. This finding directly supports our emerging hypothesis that pre-speech auditory modulation may reflect ongoing feedforward mechanisms. Consistent prior findings of reduced PSAM in stuttering speakers may therefore reflect an inherent limitation in feedforward control, supporting previous research hypothesizing an overreliance on sensory feedback in stuttering.

2. INTRODUCTION

It is well known that movement induces a generalized central modulation of sensory evoked potentials – regardless of effector system (Coquery et al., 1972; Hazemann et al., 1975; Horváth, 2015). Indeed, both speech and nonspeech effector systems seem to show similar motor-to-sensory influences in humans. In light of the extensive evidence for *attenuated* neurophysiological responses during movement and movement planning, it is often assumed that the function of this motor-to-sensory signaling is to globally suppress the sensory processing of predicted input (e.g., reafference). Such global suppression may allow for more relevant sensory processing to take precedence (Creutzfeldt et al., 1989; Davidson & Wolpert, 2005). Other theories of sensory suppression suggest that individuals may rely on such a mechanism to better detect self-produced errors or distinguish between self-generated sensations and externally generated events (Blakemore et al., 2000; Chang et al., 2013).

Consistent with nonhuman primate data showing that both suppressed and excited auditory cortical neurons are sensitive to different acoustic properties (e.g., Eliades & Wang, 2003, 2008, 2017), our laboratory has previously hypothesized that the auditory modulation observed prior to speech onset reflects an optimization process for processing speech-relevant acoustic features, whereby some neuronal populations are downregulated while others are upregulated (Max & Daliri, 2019). In other words, this pre-speech auditory modulation (PSAM) may represent a fine-tuning process of relevant sensory systems in order to better monitor upcoming sensory feedback.

However, recent evidence from our laboratory is not in line with this hypothesis. Specifically, at the same time point at which we measure PSAM prior to speech onset, we have now measured *worse* formant discrimination rather than *enhanced* discrimination (Wang et al., 2024). We also discovered a significant negative correlation between an individual's extent of modulation and their acoustic variability in produced vowels (Li et al., 2024). This correlation with variability is observable at the very onset of speech, at latencies that are well before the time of feedback-based corrections. These newer data therefore suggest that PSAM may relate to a process that is (a) already ongoing prior to movement onset and (b) contributes to motor command preparation during the planning phase. In other words, a more complex interaction with sensory systems may be occurring. There may be active bidirectional communication between sensory (e.g., auditory) and motor/premotor areas in order to, for example, determine the most appropriate motor command to send out.

Thus, this study aims to further characterize the functional relevance of pre-speech auditory modulation. If PSAM reflects a predictive correction by the control system that arises early in the movement, it may also be related to reliance on feedback-based control during the movement (i.e., online correction or online compensation). When the central nervous system is not able to correct a movement based on discrepancies detected according to internal predictions, it may further adjust the movement using another correction strategy that relies on online sensory feedback. Therefore, the documented absence of PSAM in individuals who stutter (see Max & Daliri, 2019, for a review) may be associated with not only less predictive correction, but a larger amount of online feedback-based control.

A large amount of reliance (i.e., overreliance) on online feedback-based control in individuals who stutter (IWS) has been previously hypothesized to contribute to unstable control and, subsequently, dysfluencies (Max, 2004; Max et al., 2004). Researchers have computationally simulated overreliance on online feedback control using the DIVA (Directions Into Velocities of Articulators) model (Tourville & Guenther, 2011). This "neurally altered" version of DIVA was implemented with a 'reset' function that would be triggered by an excessive amount of feedback control. When a bias toward feedback control was introduced, this reset signal was triggered repetitively, leading to stuttering behavior (e.g., syllable repetition) in DIVA (Civier et al., 2010). Additional support for a larger amount of feedback control in IWS compared to typical speakers comes from our own study (Kim, Gracco, & Max, in preparation). We used kinematic landmarks to estimate feedforward vs. feedback control (Kim & Max, 2014) and found that earlier kinematics in IWS could not predict movement outcomes as well as they did in typical speakers. This means that the movements were less planned prior to movement onset and adjusted later during the latter phase of the movements. In addition, adding a duration variable to the statistical model could explain movement outcome for IWS significantly more than it did in typical speakers, suggesting that movements made by individuals who stutter were adjusted more by shortening or delaying the movement duration (i.e., online correction).

Nevertheless, there are several studies that have argued instead for *decreased* reliance on online feedback-based control in IWS. These studies applied various kinds of perturbation in random trials and found smaller compensatory responses in IWS (formant shifts in Cai et al., 2012; pitch shifts in Loucks et al., 2012; tendon vibration in Loucks & De Nil, 2006; both proprioceptive and auditory perturbation in Namasivayam et al., 2009). However, these results may not necessarily suggest decreased reliance on online feedback control in IWS. First, the unexpected perturbations applied in these studies are not congruent with speaking situations in which IWS typically experience dysfluencies. Unlike these approaches, Kim et al. (in preparation) explored movements in an unperturbed environment (i.e., a speaking environment in which IWS typically experience dysfluencies) and found *more* reliance on online feedback-based control in IWS. In addition, the findings of Cai et al. (2012) were not replicated in a follow-up study conducted by the same authors (Cai et al., 2014), in which IWS and typical speakers were found to not differ in their response to the unexpected perturbations. Moreover, it also has been suggested that some of these results can be interpreted with an alternative explanation: namely, that IWS may rely too much on a different feedback modality (e.g., somatosensory vs. auditory), depending on the context (Civier et al., 2010).

In sum, exploring the functional relevance of pre-speech auditory modulation is crucial for advancing our understanding of stuttering and its underlying mechanisms. It is possible that PSAM may reflect a predictive correction mechanism and the lack of modulation in IWS may result in large movement variabilities in their speech movements that remain uncorrected at the onset of movement. As a result, IWS may also have to rely *more* on online auditory feedback during speech production to further adjust their speech movements. Hence, this study's aim was to examine whether the amount of pre-speech auditory modulation is associated with the degree of online feedback-based control. To meet this aim, we recorded kinematic data from unperturbed articulatory movements and used a statistical approach that we have previously validated for speech movements in order to estimate the relative contributions of feedforward and feedback mechanisms (Kim & Max, 2014). The same participants also participated in an electrophysiological session to allow for a deeper investigation into the functional relevance of pre-speech auditory modulation.

Based on current theories of stuttering and previous results showing an overreliance on online feedback control in IWS, we hypothesized that individuals who stutter would rely more on online sensory feedback-based control than typical speakers, replicating our preliminary findings (Kim, Gracco, & Max, in preparation). Further, we expected to find a negative correlation between an individual's reliance on feedback control and the observed amount of

PSAM. In other words, the lack of modulation found among IWS may reflect a deficient predictive correction mechanism, leading to a greater need to adjust movements on feedback-based control throughout the utterance.

3. METHODOLOGY

3.1. Participants

A total of 16 native speakers of American English were recruited for this experiment. Eight adults who stutter were matched pairwise with 8 adults who do not stutter (i.e., typical speakers) based on self-reported handedness, biological sex, and age (± 3 years). All participants reported (i) having no speech, language, or hearing difficulties/diagnoses (other than stuttering in the Stuttering group), (ii) having no neurological or psychological diagnoses, (iii) not to be taking medications with sensorimotor side effects, and (iv) having minimal facial hair to allow for the adhesion of facial sensors. Participants also did not report any dental modifications (e.g., false teeth, permanent retainers) that could have been damaged by the dental adhesive. All participants passed a pure tone audiometric screening on their first visit to the laboratory, verifying typical hearing (≤25 dB HL at all octave frequencies from 250 to 8000 Hz).

We sought to recruit an even distribution of male and female participants, as all biological sexes were eligible to participate in the study. Ultimately, the final sex distribution of IWS in the analyzed sample was more balanced than that of the general stuttering population (in adults, approximately 3-4 times more males than females; Bloodstein & Bernstein Ratner, 2008).

Out of the 16 participants who completed the experiment, two stuttering participants (1 female and 1 male, both cisgender) were highly dysfluent throughout the *Kinematic* task. As dysfluencies are known to affect kinematic measurements, these 2 stuttering participants and their 2 matched nonstuttering counterparts were removed from data analysis.

Thus, the final Stuttering group comprised 6 participants (4 males and 2 females, all cisgender and all right-handed). The mean age was 25.2 years, ranging from 19.0 to 34.2 years. Table 1 provides a detailed summary of the sample's demographics. According to a self-report questionnaire restricted to the categories used by the National Institutes of Health (NIH), all 6 participants were not Hispanic/Latino. Most were White (*N* = 5), and one was Asian.

The Nonstuttering group (*N* = 6) had similar demographics. One of the participants assigned female at birth was genderqueer, but all other participants were cisgender. The control participants had a mean age of 25.0 years, ranging from 18.8 to 32.1 years. Most participants in this group were White and not Hispanic/Latino (*N* = 4), one participant was Asian and not Hispanic/Latino, and one female reported Hispanic/Latino ancestry of unknown race.

All stuttering participants self-identified as an individual who stutters (IWS), responding affirmatively to the question "Do you stutter?" in our eligibility questionnaire. The self-reported onset of stuttering must have occurred during childhood, prior to age 8. One participant with a self-reported onset at 9-10 years of age had a parent who reported the onset of stuttering to be around 4-5 years of age, so we opted to include their data in the analysis.

At the beginning of the first experimental session, all participants in the Stuttering group were videotaped while engaging in a short (5-10 minute) conversation with the experimenter. They then read a short passage aloud. To confirm the presence of stuttering, both of these same-day speech samples were scored offline according to the 4th Edition of the Stuttering Severity

Instrument (SSI-4; Riley, 2009) by an ASHA-certified speech-language therapist with extensive stuttering experience.

Out of the 12 participants included in data analysis, a total of 8 had previously participated in one of our prior research studies (3 IWS and 5 IWNS). Participants were therefore relatively comfortable in our laboratory environment and familiar with the experimenter(s).

The University of Washington's Institutional Review Board approved all experimental procedures and participants provided informed written consent prior to participating in all tasks. Participants were monetarily compensated for their time ($15/hr prior to 2023 and $20/hr subsequently).

3.2. Overall experimental paradigm

Participants came to the laboratory on two separate occasions to complete two tasks: the *Auditory Modulation* task and the *Kinematic* task. Upon study recruitment, participants were assigned a counterbalanced order to complete the two sessions, yet due to participant cancellations and exclusion from data analysis, only 2 of the 6 pairs of participants presented here completed the *Kinematic* task first. Despite this, matched pair of participants completed the two sessions in the same order (i.e., Nonstuttering participants completed the task in the same order as their matched Stuttering participant).

Across all participants, there was a mean delay of 12.8 days between the two sessions (range: 4-22 days). There was no significant difference between the two groups with respect to the number of days between the two sessions, as determined by a two-sample *t*-test ($t[10] = -0.772$, $p = .458$).

3.3. *Auditory Modulation* task

The *Auditory Modulation* task was based on our previously published studies investigating the same neurophysiological phenomenon (Daliri & Max, 2015a, 2015b, 2016, 2018). Participants were seated in a sound-attenuated chamber while wearing a 128-channel active-electrode nylon cap (ActiveTwo, BioSemi; Amsterdam, The Netherlands). They faced a monitor with a microphone (SM58, Shure Incorporated; Niles, IL, USA) placed 15 cm in front of their mouth.

The basic pre-speech auditory modulation paradigm is schematized in Figure 1. As in our previous studies, participants were asked to produce monosyllabic words upon the presentation of a ‘go’ signal. This delayed response reading task allowed for the probing of auditory processing just prior to speech production. Specifically, participants performed two different conditions: *Speaking* and *No-Speaking*. Stimuli consisted of a set of three monosyllabic words (*tick*, *tech*, and *tack*), chosen so that they share the same CV syllables as the *Kinematic* task.

In the *Speaking* condition, participants produced one word at a time after the ‘go’ signal (i.e., when the text turned green), 600 ms after the initial appearance of the stimulus in white text (Figure 1). In the *No-Speaking* condition, participants refrained from overtly producing the stimulus (i.e., they ignored the ‘go’ signal) and simply read the stimulus silently to themselves upon initial presentation. These conditions were designed to allow for the comparison of auditory processing during speech movement planning in the *Speaking* condition with a control condition without speech movement planning (retaining similar linguistic, motor, and cognitive processes occurring due to silent reading).

Two hundred milliseconds prior to the ‘go’ signal, on 40% of all trials, a brief probe tone (40 ms duration at 1 kHz, with 10 ms rise/fall times) was played through insert earphones (ER-3A, Etymotic Research Inc.; Grove Village, IL, USA) at an intensity of approximately 75 dB

SPL. The trial ended upon the disappearance of the green word from the screen, approximately 1 second after the ‘go’ signal. The temporal interval between trials was randomly determined from a set of possible inter-trial intervals between 1.5 and 2.5 seconds.

High-density electroencephalographic (EEG) signals were recorded using an extension of the international 10-10 electrode system (Oostenveld & Praamstra, 2001). Two additional electrodes were placed on the left and right mastoids for offline re-referencing. We also recorded electrooculograms (EOGs) to verify and reject ocular artifacts using electrodes placed on the outer canthus of the left eye and below the left eye.

Orofacial electromyograms (sEMGs) were recorded to detect and reject any active muscle movements prior to the ‘go’ signal in the *Speaking* condition or during the *No-Speaking* trials. Surface electrodes were placed on the right side of the face over the muscles of the lower and upper lips (orbicularis oris superior and inferior), a jaw depressor (anterior belly of the digastric), and a jaw elevator (masseter). All three signals (EEG, sEMG, and EOG) were continuously digitized along with the microphone signal at a sampling rate of 1024 Hz. All data was recorded on a Windows environment using ActiView software.

Participants completed a total of 6 blocks (3 *Speaking* blocks and 3 *No-Speaking* blocks), where each block consisted of 90 trials. The order of these blocks was randomized for each participant pair (i.e., each Stuttering participant had a Nonstuttering match who completed the task with the same block order), except for one pair who completed the blocks in a different order due to experimenter error. As a result, we recorded 108 *tone* trials for each condition (90 trials in each block x 40% *tone* trials x 3 blocks for each condition) that could be analyzed for N1 modulation.

The pure tone auditory stimulus was first amplified by a headphone amplifier (HeadAmp6 Pro, ART ProAudio; Niagara Falls, NY, USA) prior to delivery to insert earphones. Participant’s speech was transduced with an additional wireless lavalier microphone (WL185 with transmitter ULX1-M1 and receiver ULXP4, Shure Incorporated; Niles, IL, USA) positioned approximately 15 cm from the mouth and amplified by a microphone amplifier (DPS II, ART ProAudio; Niagara Falls, NY, USA) to allow for near real-time auditory feedback. The measured delay in this auditory feedback setup was ~10 ms.

The experimental presentation, auditory stimulus delivery, and the delivery of external triggers for the EEG system was controlled by either PsyScope X B53 (Cohen et al., 1993) on a Mac environment or via Psychtoolbox extensions (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997) on a Linux workstation using the Octave computational environment (Eaton et al., 2020). All experimental hardware was identical for these participants (the insert earphones, amplifiers, etc.) minus the experiment presentation computer (Mac vs. Linux) and the monitor the participant was reading from (23” vs. 27”). Both setups underwent a verification of paradigm latencies prior to participant recruitment. Specifically, the timing of the onset of the auditory stimuli, the event markers, and the trial phase durations were fine-tuned using photodiodes for visual stimuli and microphones for auditory stimuli.

Prior to each individual recording session, the equipment was calibrated such that a speech signal (a recording of a typical cisgender male producing the word “tuck”) played into the lavalier microphone at 75 dB SPL from 15 cm away results in an approximately 73 dB SPL output in both insert earphones. In addition to this auditory feedback calibration, the equipment was also calibrated to ensure that the probe tone played at approximately 75 dB SPL in the insert earphones. Calibration was performed with a sound level meter (Type 2250A Hand Held

Analyzer with Type 4947 1/2” Pressure Field Microphone, Bruel & Kjaer Inc.; Norcross, GA, USA) connected to a 2 cc coupler (Type 4946, Bruel & Kjaer Inc.; Norcross, GA, USA).

3.4. *Kinematic* task

Electromagnetic articulography (EMA; AG500, Carstens Medizinelektronik GmbH; Bovenden, Germany) was used to transduce articulatory movements throughout the *Kinematic task* (Figure 2). Using a high viscosity cyanoacrylate oral adhesive (PeriAcryl 90, GlueStitch, Inc.; British Columbia, Canada) and adhesive stickers, EMA sensors were attached along the midline to the tongue, jaw, and lips, with reference sensors attached to the upper gums (above the upper central incisors) and each mastoid, closely following recommendations for best practices (Rebernik et al., 2021).

Three different tongue sensors were attached: T1 (positioned ~1 cm from the tongue tip), T2, and T3 (positioned ~1.5 and 3 cm more posterior to T1, respectively). The tongue sensors were placed at their respective, approximate locations while participants protruded their tongue. The jaw sensor was attached to the mandibular gums below the lower central incisors and the lip sensors were attached along the mid-sagittal plane at both the upper and lower vermillion borders. The tongue, jaw, and upper gum reference sensors were coated in latex (EnvironMolds; Summit, NJ, USA) prior to attaching to the participant. The lip sensors were also coated in latex to be used as backup sensors if any intraoral articulator sensors needed replacement during data collection, which only occurred once (one Nonstuttering participant; the task was restarted after sensor adjustment and only the second recording was analyzed, see *Data analysis: Kinematic task* below).

Each participant produced 30 repetitions of the phrase “I say *X* a lot” for each of the three target words: *tick, tech,* and *tack* (e.g., “I say tick a lot”). In sum, participants produced a total of 90 sentences, randomly presented within blocks of 3 in large text on a 27” monitor (U2722D, Dell, Inc.; Round Rock, TX, USA) approximately 2 meters away from the participant. The three target syllables comprised a voiceless stop consonant (/t/) with high, mid, and low front vowels (/i, ɛ, æ/) in order to elicit movements of increasing amplitude. All target syllables were followed by a velar consonant (/k/) to allow for freer positioning of the tongue tip during production of the target vowels. Participants were encouraged to read the sentences naturally, without overemphasizing the stimulus words.

Custom MATLAB (version 2023a; The MathWorks, Inc.; Natick, MA, USA) software (*Marta*, written by Mark Tiede at Haskins Laboratories; New Haven, CT) was used for stimulus presentation and to manage data recording. Kinematic data were recorded on the AG500 control laptop at a sampling rate of 200 Hz, time-locked with audio data from a clip-on EMA-compatible microphone recorded at 16 kHz. An additional microphone was placed near the experimental presentation monitor (Figure 2) to allow for *Marta* experimental control.

Individual occlusal planes were determined using bite plate sweeps recorded at either the beginning or end of the session. Specifically, participants gently held an acrylic bite plate (Carstens Medizinelektronik GmbH; Bovenden, Germany) between their teeth for 3-4 recording sweeps of approximately 5 seconds. The participant’s incisors rested directly behind a small stop in the center of the bite plate. Three EMA sensors were attached to the bite plate in a fixed configuration outside of the mouth (left, right, and central), with the central sensor directly in front of the stop where the incisors were aligned. As mentioned above, these bite plate sensors were used to redefine the data to each individual participant's occlusal plane, making the origin (0,0) at the facial tip of the upper central incisors (Figure 7B).

3.5. Data analysis: *Auditory Modulation* task

The amount of pre-speech auditory modulation for each participant was determined according to our previously developed procedures using the EEGLAB toolbox in MATLAB (Delorme & Makeig, 2004). To obtain the necessary AEPs to calculate PSAM, the raw EEG data was offline filtered, epoched, and averaged over a frontocentral region of interest (given that modulation is largest in this region; Daliri & Max, 2015b).

Specifically, all raw EEG data was first re-referenced to the average of the mastoid signals. The subsequent processing steps depended minimally on the dataset. The datasets recorded with the PsyScope stimulus presentation ($N = 8$) were analyzed with an older version of EEGLAB (v. 14.1.2b). For these datasets, bad channels were identified and interpolated using the EEGLAB functions *pop_rejchan* (with a threshold of 3) and *pop_interp*. The EEG channels were then finite impulse response (FIR) filtered from 2 to 50 Hz using the default Hamming windowed sinc filters.

The remaining datasets ($N = 4$) were analyzed with a newer version of EEGLAB (v. 2023.1) based on the recommended analysis pipeline (Delorme, 2023). Data were first lowpass filtered at 50 Hz. The *clean_rawdata* plugin (v. 2.91) then applied an additional 0.5 Hz highpass FIR filter prior to marking "bad" channels to be interpolated. Flatlined electrodes (exceeding 5 seconds), electrodes with line noise (threshold set at 4 SD), and channels exceeding a 0.8 correlation threshold were marked for interpolation. Importantly, despite minor highpass filter differences, all data were cleaned to a similar extent using conservative channel interpolation settings.

All datasets were then epoched from -600 to 1400 ms relative to the onset of the probe tone (or equivalent time for *no-tone* trials). The mean amplitude of the 100 ms period prior to the stimulus onset was subtracted from the entire epoch to remove baseline differences. Preprocessing for the sEMG channels involved filtering from 10-300 Hz prior to rectification. Finally, the EOG channels were filtered from 1-10 Hz and a bipolar EOG channel was created by summing the two individual channels.

Epochs were automatically rejected from analysis if they contained EEG amplitudes greater than ±200 µV or EOG amplitudes greater than ±150 µV. Epochs were then manually inspected for further rejection. Based on the microphone signal recorded with each epoch, individual trials were rejected if a participant failed to adhere to the task instructions (e.g., they failed to produce speech in the *Speaking* condition or they produced any speech in the *No-Speaking* condition). In addition, the sEMG signals were used to reject contaminated epochs in which the participant moved their articulators in advance of the ‘go’ signal for all conditions. For the Stuttering group, if a trial was perceptibly dysfluent, that epoch was also manually rejected.

After automatic and manual epoch rejections, three participants (2 IWS and 1 IWNS) ended with a substantial difference in the number of epochs between *Speaking* and *No-Speaking* conditions (fewer *Speaking* epochs). We further limited the number of *No-Speaking* epochs for these participants in order to minimize the effect of a large difference in the number of epochs on the comparison of AEP amplitudes. Specifically, for each participant, a random selection of *No-Speaking* epochs was removed in order to approximate the final number of clean *Speaking* epochs. Despite the reduced number of epochs, the morphology of the average evoked responses after this additional rejection step appeared highly similar to the other participants in the dataset, so they were retained in subsequent analyses.

Within each condition (*Speaking* and *No-Speaking*), the average cortical activity recorded during trials without auditory stimulation (*no-tone* trials) were then subtracted from the average EEG activity recorded during trials with auditory stimulation (*tone* trials) to remove “drift” from non-auditory activity related to motor, cognitive, and visual processes.

Automatic preprocessing of the EEG data yielded a total of 4,330 epochs. Across all participants, 7.6% of epochs (12% of *Speaking* epochs and 4% of *No-Speaking* epochs) were further rejected during a manual inspection of the data, prior to the subtraction of *no-tone* trials (not including the epochs removed to balance the number of epochs across conditions).

Participants in the Stuttering group had final datasets with a median of 87.5 *No-Speaking* epochs (ranging from 55 to 99; SD = 16.7) and a median of 70.5 *Speaking* epochs (range: 54 to 87; SD = 11.2). Participants in the Nonstuttering group had a final median of 90.5 *No-Speaking* epochs (ranging from 33 to 105; SD = 24.2) and a median of 84.5 *Speaking* epochs (range: 31 to 101; SD = 23.8). There was no significant difference between groups for either *Speaking* (t[15.7] = 1.110, p = .284) or *No-Speaking* epochs (t[19.5] = 0.511, p = .615), as determined by two-sample t-tests on individual participants’ final epoch counts.

The final AEPs for each condition were smoothed by applying an additional lowpass filter at 15 Hz. Our analysis involved comparing the AEPs measured during the *Speaking* condition with those elicited by the same auditory stimulus during the control condition without speech movement planning (*No-Speaking*). Specifically, the peak amplitudes and latencies of the N1 component (the largest negative peak between 70 and 130 ms after stimulus onset) were extracted over a frontocentral region-of-interest (ROI; based on Daliri & Max, 2015b). The frontocentral ROI included the following 17 electrodes using BioSemi's ABC labeling system: A1 (equivalent to Cz), B1, B20, B21, B32, C1, C2, C11, C22, C23, C24, D1, D2, D13, D14, D15, and D18.

Finally, pre-speech auditory modulation values for each participant were calculated by taking the difference in peak N1 amplitude between the *Speaking* and *No-Speaking* conditions. Positive values of PSAM thus represent an N1 amplitude that is smaller when the participant is about to produce speech (i.e., during speech movement planning).

3.6. Data analysis: *Kinematic* task

Immediately after data collection, kinematic data were lowpass filtered with a cut-off frequency of 15 Hz for moving articulator sensors and 5 Hz for stationary reference sensors, corrected for head movement using the mastoid sensors and the upper incisor sensor, and re-expressed relative to an anatomically defined coordinate system based on each individual participant's occlusal plane (as described above). These processing steps were applied using the default Cartsens procedures on the AG500 control laptop.

Offline kinematic measurements were made using a custom MATLAB interface that allowed for the visual and auditory inspection of each trial and isolation of the opening movement for the /tV/ syllables. We limited our kinematic investigation to the tongue tip (T1) and the jaw (lower incisor) sensors as we were interested in the opening movements associated with the alveolar consonant /t/, matching the stimuli of the *Auditory Modulation* task. For each trial and each effector, the audio signal was visualized with the *x* and *y* position traces, the tangential velocity trace, and the tangential acceleration trace.

Specifically, for each utterance and each effector, we extracted three key kinematic variables for the CV opening movement: peak tangential acceleration, peak tangential velocity, and total movement extent – as well as the effectors' spatial coordinates at the time of each landmark.

Data extraction involved manually inspecting each of the 90 utterances for each participant and restricting the analysis window to only the CV movements under investigation (i.e., isolating just the /tV/ opening movements in the stimulus words). The audio signals were used to determine the general time window of interest for each opening movement. For each effector, movement onset and offset timepoints were defined as the local minima immediately preceding and following the tangential velocity peak corresponding to the opening movement. Peak tangential acceleration was defined as the maximum tangential acceleration value between the time of movement onset and the tangential velocity peak. Total movement extent for each effector was defined as the Euclidean distance between the onset and offset positions in (x, y) space. Total movement duration was defined as the total time between movement onset and offset. For each participant, movement paths within each vowel category were re-expressed to start at the median starting (x, y) position for that vowel (Figure 7A).

Any utterances with perceptible stuttering moments or speech errors were excluded from data analysis. Stuttering judgments were made by a certified speech-language pathologist and a doctoral student with appropriate training and experience evaluating stuttered speech. Movements were also excluded from analysis due to noise in the kinematic data or ambiguous kinematic parameters. Occasionally, individual movements had multiple tangential velocity peaks, as in our prior work (Kim & Max, 2014). Movements with multiple peaks in the tangential velocity profile were excluded from analysis. See *Results* for a detailed breakdown of the rejected movements.

At an individual level, we estimated the strength of feedback-based corrections using a statistical technique based on multiple correlation, following our previously described methods validating this procedure in typical speech (Kim & Max, 2014). In brief, we operationalized a speaker's reliance on feedback control as the additional variance in movement extent that could be explained by adding movement duration (a variable assumed to be affected by feedback-based adjustments) to a multiple correlation model with peak kinematic measures (both peak tangential acceleration and peak tangential velocity) as additional predictor variables. Crucially, these correlations were compared to simpler bivariate correlations where the same peak kinematic measures *alone* predicted movement extent (i.e., the feedforward estimates).

Such feedback-based contributions can be inferred as long as 1) peak tangential acceleration and peak tangential velocity scale with movement extent, 2) for movements toward the same goal (e.g., the same vowel identity), movement duration correlates negatively with such peak kinematic measures, and 3) the inclusion of movement duration in the multiple correlation models significantly increases the overall explained variance in movement extent. We have previously validated this procedure for the speech movements of nonstuttering speakers (Kim & Max, 2014).

Here, we defined an individual's reliance on feedforward control as the percentage of variance explained in movement extent by just peak acceleration/velocity values. Feedback-based control estimates, on the other hand, were defined as the additional variance explained by *adding movement duration* to the same models. Thus, for each participant, we extracted 8 control estimates: feedforward and feedback reliance (2) across both peak acceleration and peak velocity measures (x2) for both effectors (jaw and tongue movements separately; x2).

3.7. Statistical analyses

All statistical analyses were conducted in R with RStudio (RStudio Team, 2020) using the *rstatix* package (Kassambara, 2023), unless otherwise noted, using an alpha level of 0.05 to determine statistical significance. Visualizations were created using the *ggplot2* (Wickham, 2016) and *cowplot* (Wilke, 2024) packages.

For the *Auditory Modulation* task, a separate repeated-measures ANOVA was conducted for both N1 peak amplitude and N1 peak latency, with Group (*Stuttering*, *Nonstuttering*) as a between-subjects variable and Condition (*Speaking*, *No-Speaking*) as a within-subjects variable. To account for potential violations of the sphericity assumption, degrees of freedom for the within-subjects comparisons were adjusted with Huynh-Feldt corrections (Max & Onghena, 1999). Partial omegas-squared are given as effect size estimates where appropriate (Lakens, 2013; Okada, 2013).

Pearson correlation coefficients were calculated to investigate whether same-day measures of stuttering severity were related to PSAM values. Specifically, two proxies of stuttering severity were correlated with our neurophysiological measure: (a) the mean percent of syllables stuttered during the clinical evaluation, averaged across both SSI-4 tasks, as well as (b) the total SSI-4 score (as often reported in stuttering literature).

For the *Kinematic* task, movement parameters and correlation coefficients were compared between groups or across vowel categories with two-sample *t*-tests or analyses of variances (ANOVAs), as appropriate.

Estimates of feedforward vs. feedback control were calculated in MATLAB (v. 2022b). Pearson correlation coefficients were first calculated to determine the relationships between peak acceleration/velocity values (separately) and movement extent. Bivariate correlations were conducted separately for the Stuttering and Nonstuttering groups and separately for both effectors (jaw vs. tongue tip movements). These correlations were also run both across all vowels (i.e., when movements were planned to achieve distinct distances) and separately within each vowel category (i.e., where movements were planned to achieve similar extents). The coefficients of determination for these bivariate correlations were considered to be estimates of feedforward contributions *alone*.

Next, the strength of feedback-based contributions to these opening movements was estimated by adding movement duration as an additional predictor to the above-described correlation models. Specifically, the coefficients of multiple determination for these correlations were considered to quantify the contribution of both feedforward *and* feedback control in determining overall movement extent. Thus, feedback-based control estimates *alone* were calculated by subtracting the coefficients of determination from both modeling procedures.

F statistics were calculated to determine whether the total variance explained by including movement duration in the models was significantly greater than the total variance explained by the bivariate models alone, tested at an alpha level of 0.05. See Kim & Max (2014) for further mathematical details. Due to the large number of correlation tests run for this exploratory kinematic analysis, the *p*-values presented are uncorrected.

For both feedforward and feedback control estimates, a three-factor ANOVA was conducted across models, with Group (*Stuttering*, *Nonstuttering*), Effector (*Tongue Tip*, *Jaw*), and Kinematic Parameter (*Peak Tangential Acceleration*, *Peak Tangential Velocity*) as main effects of interest as well as all of the respective interaction terms. Partial omegas-squared are given as effect size estimates where appropriate. Significant effects were followed up with *post hoc* pairwise comparisons using the Tukey Honest Significant Difference test.

To test our primary hypothesis of interest, the feedforward and feedback control estimates were entered into separate Pearson correlation analyses to determine their relationships with the amount of PSAM measured in the *Auditory Modulation* task across all individuals. As this investigation was exploratory in nature, we opted to run correlations for all 8 models separately (feedforward vs. feedback estimates × 2 effectors × 2 kinematic parameters). As above, all *p*-values presented for these correlations are uncorrected.

4. RESULTS

4.1. *Auditory Modulation* task

Grand average auditory-evoked responses for both *Speaking* and *No-Speaking* conditions for both groups are shown in Figure 3. Group data for the distribution of PSAM values are summarized in Figure 4.

4.1.1. N1 peak amplitude

The main effects of Group ($F[1,10] = 0.014$, $p = .910$, $\omega_p^2 < 0.01$) and Condition ($F[1,10] = 2.655$, $p = .134$, $\omega_p^2 = 0.03$) were not significant. The Group × Condition interaction was also not significant ($F[1,10] = 0.815$, $p = .388$, $\omega_p^2 < 0.01$).

4.1.2. N1 peak latency

There were no significant main effects or interactions for N1 latencies (in all cases, $F[1,10] < 0.784$, p values > .397, and partial omegas-squared < 0.01).

4.1.3. Pre-speech auditory modulation (PSAM) and stuttering severity

Correlations between PSAM values and stuttering severity proxies revealed no significant relationships. Pearson correlation coefficients were not significant for either mean stuttering frequency ($r = 0.212$, $CI_{95\%} = [-0.72, 0.87]$, $p = .686$) or total SSI score ($r = 0.430$, $CI_{95\%} = [-0.59, 0.92]$, $p = .395$).

4.2. *Kinematic* task

All participants were able to complete at least one full recording of 90 sentences (30 productions of each target syllable) except for one nonstuttering dataset with only 50 sentences due to a recording issue. Most participants also repeated the task a second time to record a back-up dataset, even though only one complete dataset of 90 sentences was analyzed for each participant. The first set was always analyzed unless there was a) a recording issue (e.g., noisy sensor) that prompted a second recording after adjusting the setup ($N = 1$ Nonstuttering participant) or b) there were more fluent trials in the second set ($N = 2$ Stuttering participants).

As previously mentioned, a total of 2 Stuttering participants and their Nonstuttering counterparts were entirely removed from data analysis due to consistent dysfluencies throughout the *Kinematic* task. Beyond these excluded datasets, the primary reason that individual movements were rejected in the Stuttering group was due to perceptible dysfluencies.

In both groups, participants also occasionally made movements with multiple peaks in the tangential velocity profile, a common finding for movements with more curvilinear paths (Abend et al., 1982). In the Stuttering group, a total of 8 jaw movements and 38 tongue movements were excluded for this reason. A total of 19 jaw and 24 tongue movements with multiple peaks in the tangential velocity profiles were excluded in the Nonstuttering group. Figure 5 shows the group-level distributions of multiple peak movements per participant. There was no significant difference between the two groups in the percentage of movements that were rejected due to multiple tangential velocity peaks ($M_{Nonstuttering}$ = 4.5% of movements, $M_{Stuttering}$ = 4.39% of movements; *t*[9.71] = -0.047, *p* = .963).

Overall, out of 515 analyzable utterances collected from the Stuttering group, a total of 172 jaw movements (33.4%) and 183 tongue movements (35.5%) were excluded from analysis. Out of 500 analyzable utterances collected from the Nonstuttering group, a total of 93 jaw movements (18.6%) and 94 tongue movements (18.8%) were excluded. There was no significant difference between the two groups with respect to the total number of tongue (*t*[8.24] = 1.550, *p* = .159) or jaw (*t*[7.29] = 1.148, *p* = .287) movement rejections per participant.

Thus, the final Stuttering dataset included 343 jaw movements and 332 tongue movements, while the Nonstuttering dataset included 407 jaw movements and 406 tongue movements. The manufacturer's recommended threshold for acceptable sensor calibration RMS values (i.e., the average RMS calculated during system calibration across all positions for each calibrated sensor) is 20, shown with subsequent accuracy assessments to minimize position errors (Yunusova et al., 2009). Notably, all recording sessions in the current study obtained calibration RMS values for all sensors that were significantly smaller, with values less than 2. The same threshold of 20 is often used to determine unacceptable RMS values recorded *during each utterance* (i.e., a threshold for discarding trials). Here, however, we opted to retain movements with slightly elevated RMS values as long as the kinematic recordings met all other criteria for analysis. For the final movements included in data analysis, the distribution of peak RMS error values for the sensors on each effector is shown in Figure 6. These peak RMS values were extracted within the analyzed window of the /tV/ opening movements for all retained movements (i.e., excluding movements that were rejected). There was no difference between the sensors with respect to peak RMS values, as determined by a two-sample *t*-test (M_{jaw} = 9.52, M_{tongue} = 9.35; *t*[1351] = 0.923, *p* = .356).

Position data from a representative participant, both before and after movement traces were shifted to the mean starting points, is shown in Figure 7. Jaw opening movements from a representative participant are shown in Figure 8. Across participants, opening movements yielded tangential velocity profiles that were generally bell-shaped with a single peak (Figure 8C), as expected based on prior literature (Abend et al., 1982; Morasso, 1981). Tangential acceleration profiles had typical acceleration and deceleration phases. Tongue tip opening movements appeared highly similar, although tongue movements were more curved on average.

Notably, the peak kinematic variables generally scaled with vowel height, a key prerequisite for the feedforward vs. feedback control estimation procedure (Kim & Max, 2014). Specifically, our task successfully obtained movements of varying extent across vowel categories, with extent increasing as vowel height decreased (Figure 8D; *F*[2,1484] = 189.8, $p < .001$, $\omega_p^2 = 0.20$).

Across all movements (i.e., regardless of vowel identity), movement duration was significantly longer for the Stuttering participants than the Nonstuttering participants (*t*[1388.6] = -12.5, $p < .001$). As expected, movement duration also varied significantly with vowel category (*F*[2,1484] = 44.4, $p < .001$, $\omega_p^2 = 0.06$).

4.2.1. Correlations between peak kinematic parameters and movement extent

Bivariate correlations showed that, generally, both peak tangential acceleration and peak tangential velocity scaled with movement extent. Three example participants with increasing correlation coefficients are shown in Figure 9, with the lowest correlation coefficients corresponding to data from a Stuttering participant (Figure 9A; pink regression line). The overall distribution of coefficients extracted from these bivariate correlations are visualized in Figure 10, with quantitative summaries tabulated in Table 2.

Descriptively, correlation coefficients were higher when including movements across all vowels together (Figure 10A) as compared to correlations within each individual vowel category, where many coefficients were, in fact, negative (Figure 10B). Indeed, all bivariate correlations *across* vowels were statistically significant except for 17% of the correlations between peak tangential velocity and extent for the tongue tip sensors in the Stuttering group (Table 2). Correlations *within* each vowel were less significant overall, particularly for the tongue tip movements in the peak acceleration models.

Regardless of effector and group, correlation coefficients were higher for the velocity models than the acceleration models both across vowels ($t[43.12] = -4.00$, $p < .001$) and within vowels ($t[135.6] = -6.53$, $p < .001$).

Estimates of individual participants' reliance on feedforward control are represented with the light bars in Figure 12. Across participants and models (e.g., calculated using tongue tip sensor measurements vs. jaw sensor measurements), the proportion of variance explained by these bivariate correlations varied considerably. There was a significant main effect of Kinematic Parameter ($F[1,40] = 17.612$, $p < .001$, $\omega_p^2 = 0.26$), but neither the main effect of Group ($F[1,40] = 2.802$, $p = .102$, $\omega_p^2 = 0.04$) or Effector ($F[1,40] = 1.360$, $p = .250$, $\omega_p^2 < 0.01$) were significant in this sample. None of the interactions terms were significant either (all F values < 0.325, all p values > .572, and all $\omega_p^2 < 0.01$). *Post hoc* comparisons showed that feedforward control estimates were higher for velocity models than acceleration models (M_{diff} = 27.78, $CI_{95\%}$ = [12.4, 41.2], $p < .001$). Differences in the group means showed that participants who stutter had a smaller reliance on feedforward control than the control participants, yet this comparison did not reach statistical significance (M_{diff} = -11.08, $CI_{95\%}$ = [-24.46, 2.30], $p = .102$).

4.2.2. Feedforward vs. feedback control estimates

Next, we considered the addition of movement duration as an additional predictor variable for movement extent based on prior results suggesting that shortening or delaying movement duration may reflect online adjustments to better accomplish movement goals. Prior research showed that movement duration correlated positively with peak kinematic measures across movements to different targets, but *negatively* when calculating within movements of the same target (within vowel categories; Messier & Kalaska, 1999).

In this dataset, for jaw movements across vowels, the correlation between peak tangential acceleration and movement duration was positive in 6/12 cases (6 participants × 2 groups), but negative within vowels in an overwhelming 31/36 cases (6 participants × 2 groups × 3 vowels). For tongue movements, the proportion of positive correlations between peak acceleration and movement duration across vowels was much lower (3/12 cases), but there was a strong negative trend for within-vowel correlations (32/36 cases). Correlations between peak velocity and movement duration showed similar results. Jaw movements showed positive correlations between peak velocity and movement duration across vowels in 7/12 cases, but negatively within-vowels in 27/36 cases. For tongue movements, 9/12 cases across vowels were positive while 22/36 cases were negative when calculated within vowels.

Figure 11 shows two example participants (one Stuttering and one Nonstuttering) in which there was a positive correlation between peak tangential velocity and movement duration when calculating across vowels, but negative correlations when calculating within vowel categories. This negative correlation across movements toward the same target suggests that participants are compensating for variability in the achieved initial kinematics by shortening or delaying movement duration (Gordon & Ghez, 1987; Messier & Kalaska, 1999).

For all participants and models (i.e., regardless of effector and peak kinematic parameter), adding movement duration as a predictor variable significantly improved the prediction of movement extent when compared to the bivariate correlation models (all p values < 0.05).

As shown in Figures 12 and 13, the additional percent of variance explained after adding movement duration as a predictor varied across participants and models. There was a significant main effect of Group ($F[1,40] = 5.475$, $p = .024$, $\omega_p^2 = 0.09$) and Kinematic Parameter ($F[1,40] = 6.842$, $p = .013$, $\omega_p^2 = 0.11$) on feedback control estimates, but no significant main effect of Effector ($F[1,40] = 0.049$, $p = .826$, $\omega_p^2 < 0.01$). None of the interaction terms were significant (all F values < 0.763, all p values $> .388$, and all $\omega_p^2 < 0.01$). *Post hoc* testing showed a significantly lower reliance on feedback control for models calculated with peak tangential velocity compared to those calculated with peak tangential acceleration ($M_{diff} = -12.15$, $CI_{95\%} = [-21.5, -2.76]$, $p_{adjusted} = .012$). Furthermore, the Stuttering group relied significantly more on feedback control than the Nonstuttering group ($M_{diff} = 10.87$, $CI_{95\%} = [1.48, 20.25]$, $p_{adjusted} = .024$).

4.2.3. Correlating PSAM with feedback vs. feedforward control estimates

All four correlations between individual estimates of feedforward reliance and amounts of pre-speech auditory modulation were positive (Figure 14). The strongest correlations were for the acceleration models, with both jaw (r = 0.65, *p* = .021) and tongue (r = 0.72, *p* = .008) movements reaching significance (uncorrected *p* values). Neither of the correlations with feedforward estimates taken from velocity models reached significance (r = 0.37, *p* = .233 for jaw movements and r = 0.52, *p* = .084 for tongue movements).

Furthermore, all correlations between estimates of feedback reliance and extent of PSAM were negative (Figure 15). None of the correlations with feedback estimates taken from acceleration models (r = -0.40, *p* = .196 for jaw movements and r = -0.55, *p* = .066 for tongue movements) or velocity models (r = -0.30, *p* = .337 for jaw movements and r = -0.46, *p* = .133 for tongue movements) were significant.

5. DISCUSSION

This study investigated the relationship between pre-speech auditory modulation and an individual's reliance on online feedback-based control. To do so, we recruited participants who stutter and participants who do not stutter (*N* = 6 adults per group) to complete two tasks in our laboratory. One session involved measuring an individual's extent of cortical auditory modulation (i.e., the difference in N1 component amplitude in response to probe tones presented during speech movement planning and a control silent reading condition without motor planning). The other session involved recording kinematic data from unperturbed articulatory movements, allowing us to estimate, for each participant, the relative contributions of feedforward vs. feedback mechanisms.

Our sample of participants did not show significant amounts of pre-speech auditory modulation for either group, failing to replicate our previous finding that typical speakers show a significant modulation of cortical auditory processing due to speech movement planning. By extension, we did not replicate our prior findings of a between-group difference in the extent of PSAM. We believe that this failure to replicate prior results is driven mainly by the small sample size in this study (*N* = 6 adults per group), as both 1) a significant extent of PSAM for typical speakers and 2) a difference in modulation between groups have already been replicated in different samples, with large effect sizes, in multiple other studies (Daliri & Max, 2016, 2018; Max & Daliri, 2019). Notably, the small sample in the current study includes one Stuttering participant with a large amount of PSAM as well as several Nonstuttering participants with limited extents of modulation (Figure 4). A larger sample size would thus allow for a more rigorous investigation of pre-speech auditory behavior.

Our kinematic analyses showed that /tV/ opening movements were, on average, produced more slowly for speakers who stutter, corroborating previous literature showing longer movement durations in stuttering (Caruso et al., 1988; Zimmermann, 1980; but see Max et al., 2003 in which opening movement durations did not differ between groups).

We obtained generally high, positive correlations between peak initial kinematic parameters and movement extent. When calculating correlations between peak kinematic measures and movement extent for movements to different target vowels (across vowel analyses), the overwhelming majority of correlations were significant (47/48; Table 2, left side). When calculating correlations for movements to the same target vowel (within vowel analyses), the proportion of coefficients that were statistically significant dropped considerably across groups, kinematic parameters, and effectors (Table 2, right side). Specifically, the acceleration models had fewer significant correlations than those using peak velocity (ranging from 33-56% significant compared to velocity models ranging from 61-83% significant). Furthermore, within-vowel correlations based on tongue tip measurements had a lower proportion of individual models reaching statistical significance compared to jaw measurements (tongue models ranged from 33-67% significant, jaw models ranged from 56-83% significant), which could be explained by the tongue's additional degrees of freedom compared to the jaw. Overall, this finding of a lower proportion of correlations reaching statistical significance when analyzing movements aimed at each vowel target separately corresponds well with prior work investigating arm movements (Messier & Kalaska, 1999) as well as our own prior work with speech movements (Kim & Max, 2014). We obtained correlation coefficients for both across- and within-vowel analyses that closely resemble those found in the limb literature (Messier & Kalaska, 1999).

In general, the bivariate correlation coefficients between peak tangential velocity and movement extent were higher than their corresponding acceleration models (Figure 10). Across groups and vowels, mean correlation coefficients between peak acceleration and extent ranged from 0.59-0.72 while mean coefficients between peak velocity and extent ranged from 0.74-0.89. This finding was unsurprising given previous work, and the fact that peak tangential acceleration occurs very early in the movement (Kim & Max, 2014; Messier & Kalaska, 1999).

Our applied statistical procedure for estimating reliance on feedforward vs. feedback control replicates previous results suggesting that speech articulation is primarily under feedforward control (Kim & Max, 2014). Mean bivariate correlations between initial kinematic values and movement extent ranged between 0.62 to 0.89 for Nonstuttering participants and 0.59 to 0.85 for Stuttering speakers. In this relatively small sample, we did not obtain a significant difference in feedforward control estimates between groups, although the participants who stutter did, on average, show a descriptively lower reliance on feedforward control than those who do not (with an 11% lower mean percent variance explained by the bivariate models).

The additional inclusion of movement duration as a predictor in multiple correlation models significantly improved the prediction of movement extent in all cases, across all participants, corroborating prior work in which movement duration was established as a strong contender for estimating feedback-based control in such analyses (Gordon & Ghez, 1987; Kim & Max, 2014; Messier & Kalaska, 1999). We obtained a significant difference in feedback control estimates between groups, with Stuttering speakers relying significantly more on feedback control than the Nonstuttering speakers. This result replicates an earlier study using the same statistical analysis with speech articulatory movements that also found a larger extent of feedback-based control in a group of participants who stutter (Kim, Gracco, & Max, in preparation). This finding of a greater extent of online corrections during articulatory movements in adults who stutter is particularly relevant to our primary hypothesis, as previous research has often shown conflicting evidence regarding stuttering speakers' reliance on feedback-based control. Indeed, several experiments measuring online compensation to applied, unexpected perturbations have instead shown *smaller* compensatory responses in individuals who stutter compared to those who do not (formant shifts in Cai et al., 2012; pitch shifts in Loucks et al., 2012; tendon vibration in Loucks & De Nil, 2006; both proprioceptive and auditory perturbation in Namasivayam et al., 2009). The findings of the current study, in addition to Kim et al. (in preparation), instead suggest that speakers who stutter may indeed rely more on online feedback-based control compared to those who do not.

Our primary hypothesis involved testing the relationship between pre-speech auditory modulation and feedback-based control measures. We performed several correlational analyses to begin exploring such relationships and found, for all kinematic parameters and effectors tested, a *positive* correlation between PSAM and feed*forward* control estimates (Figure 14). Somewhat analogously, we also found, for all models tested, a *negative* correlation between PSAM and feed*back* control estimates (Figure 15). Specifically, we observed statistically significant positive correlations between PSAM and feedforward control estimates calculated with peak acceleration measures (i.e., regardless of effector; Figure 14, top). Despite moderate correlation coefficients, none of the correlations between PSAM and feedback control estimates were significant (Figure 15).

Notably, the feedforward and feedback control estimates are intimately connected within participants. For example, if a large proportion of variance in movement extent was already explained by the simple bivariate models, then there would be a smaller proportion of additional variance that could theoretically be explained by adding movement duration to the correlation models. As such, it is perhaps unsurprising that we obtained opposing correlation trends with PSAM and feedforward vs. feedback control estimates. The strength of correlations is, however, of note, particularly in light of the large variation across participants (Figure 12) and the small sample size in this preliminary exploration.

These correlational results provide the first direct evidence that PSAM could be capturing ongoing feedforward processes in advance of movement, in contrast to our earlier hypothesis that PSAM may reflect a selective tuning of auditory neurons in preparation for their upcoming role in speech feedback monitoring (Max & Daliri, 2019). Specifically, pre-speech auditory modulation may instead occur due to bidirectional, *predictive* communication between (pre)motor and sensory cortices. For example, a subset of auditory cortical neurons may be involved in the premovement fine-tuning of the outgoing motor command (i.e., feedforward planning) and thus be unavailable to respond to the incoming probe stimulus presented during just prior to speech onset in such paradigms. As such, we observe a reduced amplitude in the N1 component of auditory evoked responses to probe tones when presented during speech movement planning.

This emerging hypothesis follows recent work in our laboratory that showed that the amount of modulation was negatively correlated with intertrial acoustic variability as early as vowel onset (i.e., prior to the onset of auditory feedback mechanisms; Li et al., 2024). Recent psychophysical data recorded on an identical timeline is also in line with this hypothesis. Specifically, we measured *worse* formant discrimination of participants' own vowel productions at the same timepoint prior to speech production (Wang et al., 2024).

There are several limitations to this experiment and its findings. The most notable limitation is the small sample size (*N* = 6 adults per group). Despite the challenges presented by the nature of kinematic recordings of speech movements, the recruitment of participants from a clinical population (and their matched control participants), and the fact that participants had to be willing to return to the laboratory on two separate occasions, we acknowledge that our conclusions would certainly be stronger with a larger number of participants in the sample. It is particularly difficult to reconcile the common expectation that articulography studies often have fewer than 5 participants (indeed, almost half of those reviewed by Rebernik et al., 2021) with the limitations of statistical power in the *Auditory Modulation* task. We have previously observed large between-group differences in PSAM, yet the few PSAM outliers (in both groups) likely drove the lack of significant modulation observed in this sample.

In addition to the small sample size, our methodology involved correlating neurophysiological measurements with kinematic estimates recorded on separate occasions (sometimes as long as 3 weeks apart). Both tasks were inherently time-consuming for both the experimenters and the participants and involved extensive calibration and preparation in advance of data collection. Certainly, fatigue would be an additional issue for recording such datasets on the same day. We do not, however, have evidence to date regarding the replicability of PSAM measurements *within* individuals over time. The analysis presented here presupposes that PSAM measurements would remain somewhat consistent within individuals (e.g., a participant always modulates, regardless of the extent, or always does not modulate), as an index of some intraindividual pre-speech mechanism, although that remains to be investigated. Future work documenting the test-retest reliability of this neurophysiological index would therefore be informative for future investigations on PSAM.

Another limitation involves methodological choices in our kinematic data analysis, such as the choice to exclude movements with multiple peaks in the tangential velocity traces, following our prior methods (Kim & Max, 2014; Kim, Gracco, & Max, in preparation). Previous datasets had to discard approximately 10% of recorded movements across both tongue and jaw sensors, yet the present study rejected, on average, fewer than 5% of movements per participant (Figure 5). Here, we only rejected movements with unambiguous multiple peaks in the tangential velocity profiles (i.e., an additional inflection point must be present), leading to the inclusion of several “near” double peaks in which there was no additional peak in the velocity signal despite evidence to suggest online adjustments were being made (e.g., additional acceleration peaks). Notably, a descriptively greater proportion of tongue movements were discarded due to multiple tangential velocity peaks than jaw movements, which is not surprising given the tongue’s multiple degrees of freedom compared to the jaw's rotational hinge. However, the decision to *discard* movements that are likely adjusted *online* to achieve movement targets may have further reduced our estimations of reliance on feedback-based control. Regardless of the decision to include these multiple peak movements, there would be clear evidence that within-movement adjustments are being made by both groups of participants.

6. CONCLUSION

In sum, we successfully replicated prior findings that showed an increased reliance on feedback-based control in speakers who stutter compared to those who do not stutter. We did not, however, replicate prior results of pre-speech speech auditory modulation, as neither the Stuttering nor Nonstuttering groups showed a significant extent of cortical auditory modulation due to speech movement planning. Finally, we showed a direct, positive relationship between PSAM and reliance on feedforward control during unperturbed, fluent speech movements. In line with other recent work from our laboratory, we therefore suggest that PSAM reflects active, bidirectional communication between sensory and (pre)motor areas due to the fine-tuning of motor commands for subsequent speech movements. Prior studies have repeatedly observed atypical motor-to-auditory influences during speech movement planning in individuals who stutter (for a review, see Max & Daliri, 2019), which may reflect a fundamental limitation in their ability to rely on feedforward control mechanisms for speech production.

7. ACKNOWLEDGEMENTS

The authors would like to thank Allie Stewart, CCC-SLP for scoring the clinical stuttering evaluations for this experiment, as well as Churan Li and Lisa Paroni for assisting with experimental data collection. The authors are also particularly grateful for the technical assistance and AG500 experimental management software (*Marta*) provided by Mark Tiede. This research was supported by a grant from the National Institute on Deafness and Other Communication Disorders (R01 DC017444; PI: Ludo Max). This content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

8. REFERENCES

Abend, W., Bizzi, E., & Morasso, P. (1982). Human arm trajectory. *Brain*, *105*, 331–348.

Blakemore, S.-J., Wolpert, D. M., & Frith, C. D. (2000). Why can't you tickle yourself? *NeuroReport*, *11*(11), R11–R16. https://doi.org/10.1038/news981022-7

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433–436.

Cai, S., Beal, D. S., Ghosh, S. S., Guenther, F. H., & Perkell, J. S. (2014). Impaired timing adjustments in response to time-varying auditory perturbation during connected speech production in persons who stutter. *Brain and Language*, *129*(1), 24–29. https://doi.org/10.1016/j.bandl.2014.01.002

Cai, S., Beal, D. S., Ghosh, S. S., Tiede, M. K., Guenther, F. H., & Perkell, J. S. (2012). Weak responses to auditory feedback perturbation during articulation in persons who stutter: Evidence for abnormal auditory-motor transformation. *PLoS ONE*, *7*(7), 1–13. https://doi.org/10.1371/journal.pone.0041830

Caruso, A. J., Abbs, J. H., & Gracco, V. L. (1988). Kinematic analysis of multiple movement coordination during speech in stutterers. *Brain*, *111*, 439–455.

Chang, E. F., Niziolek, C. A., Knight, R. T., Nagarajan, S. S., & Houde, J. F. (2013). Human cortical sensorimotor network underlying feedback control of vocal pitch. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(7), 2653–2658. https://doi.org/10.1073/PNAS.1216827110/-/DCSUPPLEMENTAL

Civier, O., Tasko, S. M., & Guenther, F. H. (2010). Overreliance on auditory feedback may lead to sound/syllable repetitions: Simulations of stuttering and fluency-inducing conditions with a neural model of speech production. *Journal of Fluency Disorders*, *35*(3), 246–279. https://doi.org/10.1016/j.jfludis.2010.05.002

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers*, *25*(2), 257–271.

Coquery, J. M., Coulmance, M., & Leron, M. C. (1972). Modifications des potentiels évoqués corticaux somesthésiques durant des mouvements actifs et passifs chez l'homme. *Electroencephalography and Clinical Neurophysiology*, *33*(3), 269–276. https://doi.org/10.1016/0013-4694(72)90153-8

Creutzfeldt, O., Ojemann, G., & Lettich, E. (1989). Neuronal activity in the human lateral temporal lobe—II. Responses to the subjects own voice. *Experimental Brain Research*, *77*(3), 476–489. https://doi.org/10.1007/BF00249601

Daliri, A., & Max, L. (2015a). Electrophysiological evidence for a general auditory prediction deficit in adults who stutter. *Brain and Language*, *150*, 37–44. https://doi.org/10.1016/j.bandl.2015.08.008

Daliri, A., & Max, L. (2015b). Modulation of auditory processing during speech movement planning is limited in adults who stutter. *Brain and Language*, *143*, 59–68. https://doi.org/10.1016/j.bandl.2015.03.002

Daliri, A., & Max, L. (2016). Modulation of auditory responses to speech vs. Nonspeech stimuli during speech movement planning. *Frontiers in Human Neuroscience*, *10*, 1–9. https://doi.org/10.3389/fnhum.2016.00234

Daliri, A., & Max, L. (2018). Stuttering adults' lack of pre-speech auditory modulation normalizes when speaking with delayed auditory feedback. *Cortex*, *99*, 55–68. https://doi.org/10.1016/j.cortex.2017.10.019

Davidson, P. R., & Wolpert, D. M. (2005). Widespread access to predictive models in the motor system: A short review. *Journal of Neural Engineering*, *2*(3), S313–S319. https://doi.org/10.1088/1741-2560/2/3/S11

Delorme, A. (2023). EEG is better left alone. *Scientific Reports*, *13*(1), Article 1. https://doi.org/10.1038/s41598-023-27528-0

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Eaton, J. W., Bateman, D., Hauberg, S., & Wehbring, R. (2020). *GNU Octave version 6.1.0 manual: A high-level interactive language for numerical computations.* https://www.gnu.org/software/octave/doc/v6.3.0/

Eliades, S. J., & Wang, X. (2003). Sensory-motor interaction in the primate auditory cortex during self-initiated vocalizations. *Journal of Neurophysiology*, *89*(4), 2194–2207. https://doi.org/10.1152/jn.00627.2002

Eliades, S. J., & Wang, X. (2008). Neural substrates of vocalization feedback monitoring in primate auditory cortex. *Nature*, *453*(7198), 1102–1106. https://doi.org/10.1038/nature06910

Eliades, S. J., & Wang, X. (2017). Contributions of sensory tuning to auditory-vocal interactions in marmoset auditory cortex. *Hearing Research*, *348*, 98–111. https://doi.org/10.1016/j.heares.2017.03.001

Gordon, J., & Ghez, C. (1987). Trajectory control in targeted force impulses: III. Compensatory adjustments for initial errors. *Experimental Brain Research*, *67*(2), 253–269. https://doi.org/10.1007/BF00248547

Hazemann, P., Audin, G., & Lille, F. (1975). Effect of voluntary self-paced movements upon auditory and somatosensory evoked potentials in man. *Electroencephalography and Clinical Neurophysiology*, *39*(3), 247–254. https://doi.org/10.1016/0013-4694(75)90146-7

Horváth, J. (2015). Action-related auditory ERP attenuation: Paradigms and hypotheses. *Brain Research*, *1626*, 54–65. https://doi.org/10.1016/J.BRAINRES.2015.03.038

Kassambara, A. (2023). *rstatix: Pipe-friendly framework for basic statistical tests* (Version R package version 0.7.2) [Computer software]. https://CRAN.R-project.org/package=rstatix

Kim, K.S., Gracco, V., & Max, L. (in preparation). Feedforward vs. feedback control of unperturbed speech movements in stuttering and nonstuttering adults.

Kim, K. S., & Max, L. (2014). Estimating feedforward vs. Feedback control of speech production through kinematic analyses of unperturbed articulatory movements. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00911

Kleiner, M., Brainard, D. H., & Pelli, D. G. (2007). What's new in Psychtoolbox-3? *ECVP Abstract Supplement*, *36*.

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, *0*(NOV), 863. https://doi.org/10.3389/FPSYG.2013.00863

Li, J. J., Daliri, A., Kim, K. S., & Max, L. (2024). Does pre-speech auditory modulation reflect processes related to feedback monitoring or speech movement planning? *Neuroscience Letters*, *843*, 138025. https://doi.org/10.1016/j.neulet.2024.138025

Loucks, T., Chon, H., & Han, W. (2012). Audiovocal integration in adults who stutter. *International Journal of Language & Communication Disorders*, *47*(4), 451–456. https://doi.org/10.1111/j.1460-6984.2011.00111.x

Loucks, T. M., & De Nil, L. F. (2006). Oral kinesthetic deficit in adults who stutter: A target-accuracy study. *Journal of Motor Behavior, 38*(3), 238–247. https://doi.org/10.3200/JMBR.38.3.238-247

Max, L. (2004). Stuttering and internal models for sensorimotor control: A theoretical perspective to generate testable hypotheses. In B. Maassen, R. Kent, H. Peters, P. van Lieshout, & W. Hulstijn (Eds.), *Speech Motor Control in Normal and Disordered Speech* (pp. 357–387). Oxford University Press.

Max, L., Caruso, A. J., & Gracco, V. L. (2003). Kinematic analyses of speech, orofacial nonspeech, and finger movements in stuttering and nonstuttering adults. *Journal of Speech, Language, and Hearing Research*, *46*(1), 215–232. https://doi.org/10.1044/1092-4388(2003/017)

Max, L., & Daliri, A. (2019). Limited pre-speech auditory modulation in individuals who stutter: Data and hypotheses. *Journal of Speech, Language, and Hearing Research*, *62*(8S), 3071–3084. https://doi.org/10.1044/2019_JSLHR-S-CSMC7-18-0358

Max, L., Guenther, F. H., Gracco, V. L., Ghosh, S. S., & Wallace, M. E. (2004). Unstable or insufficiently activated internal models and feedback-biased motor control as sources of dysfluency: A theoretical model of stuttering. *Contemporary Issues in Communication Science and Disorders*, *31*, 105–122.

Max, L., & Onghena, P. (1999). Some issues in the statistical analysis of completely randomized and repeated measures designs for speech, language, and hearing research. *Journal of Speech, Language, and Hearing Research*, *42*(2), 261–270. https://doi.org/10.1044/jslhr.4202.261

Messier, J., & Kalaska, J. F. (1999). Comparison of variability of initial kinematics and endpoints of reaching movements. *Experimental Brain Research*, *125*, 139–152.

Morasso, P. (1981). Spatial control of arm movements. *Experimental Brain Research*, *42*(2), 223–227. https://doi.org/10.1007/BF00236911

Namasivayam, A. K., van Lieshout, P., McIlroy, W. E., & De Nil, L. (2009). Sensory feedback dependence hypothesis in persons who stutter. *Human Movement Science*, *28*(6), 688–707. https://doi.org/10.1016/j.humov.2009.04.004

Okada, K. (2013). Is omega squared less biased? A comparison of three major effect size indices in one-way anova. *Behaviormetrika*, *40*(2), 129–147. https://doi.org/10.2333/bhmk.40.129

Oostenveld, R., & Praamstra, P. (2001). The five percent electrode system for high-resolution EEG and ERP measurements. *Clinical Neurophysiology*, *112*(4), 713–719. https://doi.org/10.1016/S1388-2457(00)00527-7

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437–442.

Rebernik, T., Jacobi, J., Jonkers, R., Noiray, A., & Wieling, M. (2021). A review of data collection practices using electromagnetic articulography. *Laboratory Phonology*, *12*(1), Article 1. https://doi.org/10.5334/labphon.237

Riley, G. (2009). *SSI-4: Stuttering severity instrument* [Computer software]. Pro-Ed.

RStudio Team. (2020). *RStudio: Integrated Development Environment for R* [Computer software]. RStudio, PBC. www.rstudio.com

Tourville, J. A., & Guenther, F. H. (2011). The DIVA model: A neural theory of speech acquisition and production. *Language and Cognitive Processes*, *26*(7), 952–981. https://doi.org/10.1080/01690960903498424

Wang, H., Ali, Y., & Max, L. (2024). Perceptual formant discrimination during speech movement planning. *PLOS ONE*, *19*(4), e0301514. https://doi.org/10.1371/journal.pone.0301514

Wickham, H. (2016). *Ggplot2: Elegant graphics for data analysis* [Computer software]. Springer-Verlag New York. https://ggplot2.tidyverse.org

Wilke, C. O. (2024). *cowplot: Streamlined plot theme and plot annotations for "ggplot2"* (Version R package version 1.1.3) [Computer software]. https://wilkelab.org/cowplot/

Yunusova, Y., Green, J. R., & Mefferd, A. (2009). Accuracy Assessment for AG500, Electromagnetic Articulograph. *Journal of Speech, Language, and Hearing Research*, *52*(2), 547–555. https://doi.org/10.1044/1092-4388(2008/07-0218)

Zimmermann, G. (1980). Stuttering: A Disorder of Movement. *Journal of Speech and Hearing Research*, *23*(1), 122–136. https://doi.org/10.1044/JSHR.2301.122

9. FIGURES

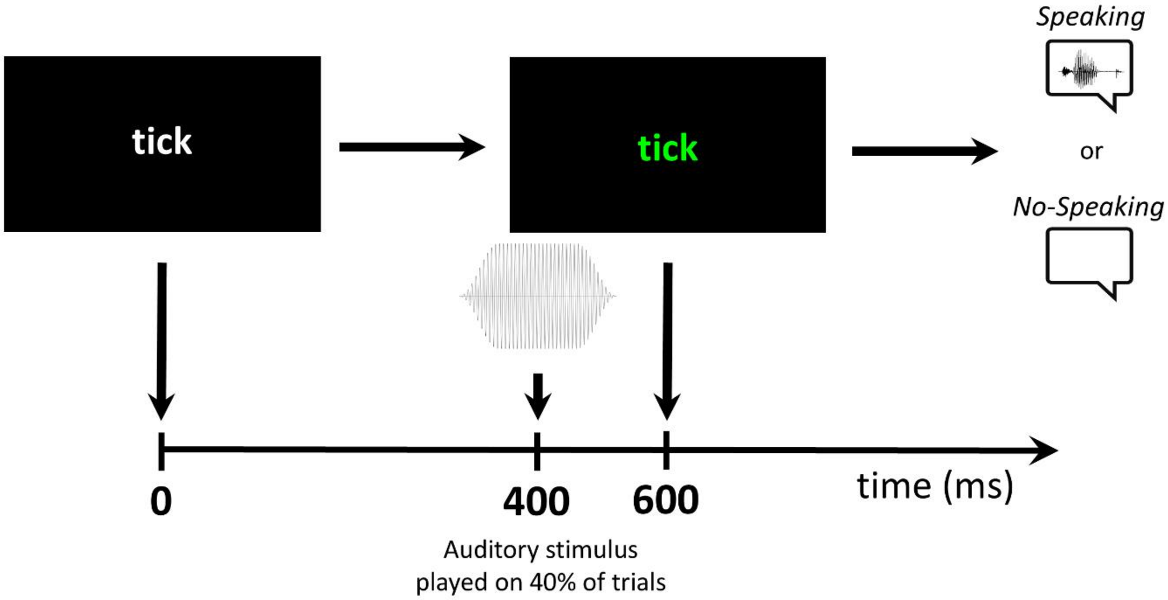

Figure 1

Schematic representation of the basic pre-speech auditory modulation paradigm for the *Auditory Modulation* task. In the *Speaking* condition, participants produced one word at a time after the 'go' signal (i.e., when the text turned green), 600 ms after the initial appearance of the stimulus in white text. In the *No-Speaking* condition, participants simply read the words silently to themselves (i.e., ignoring the 'go' signal). On 40% of the trials in all conditions, a probe tone was played into the participant's ears via insert earphones (*tone* trials). This pure tone auditory stimulus was played 200 ms prior to the 'go' signal to probe auditory processing mechanisms during speech movement planning.

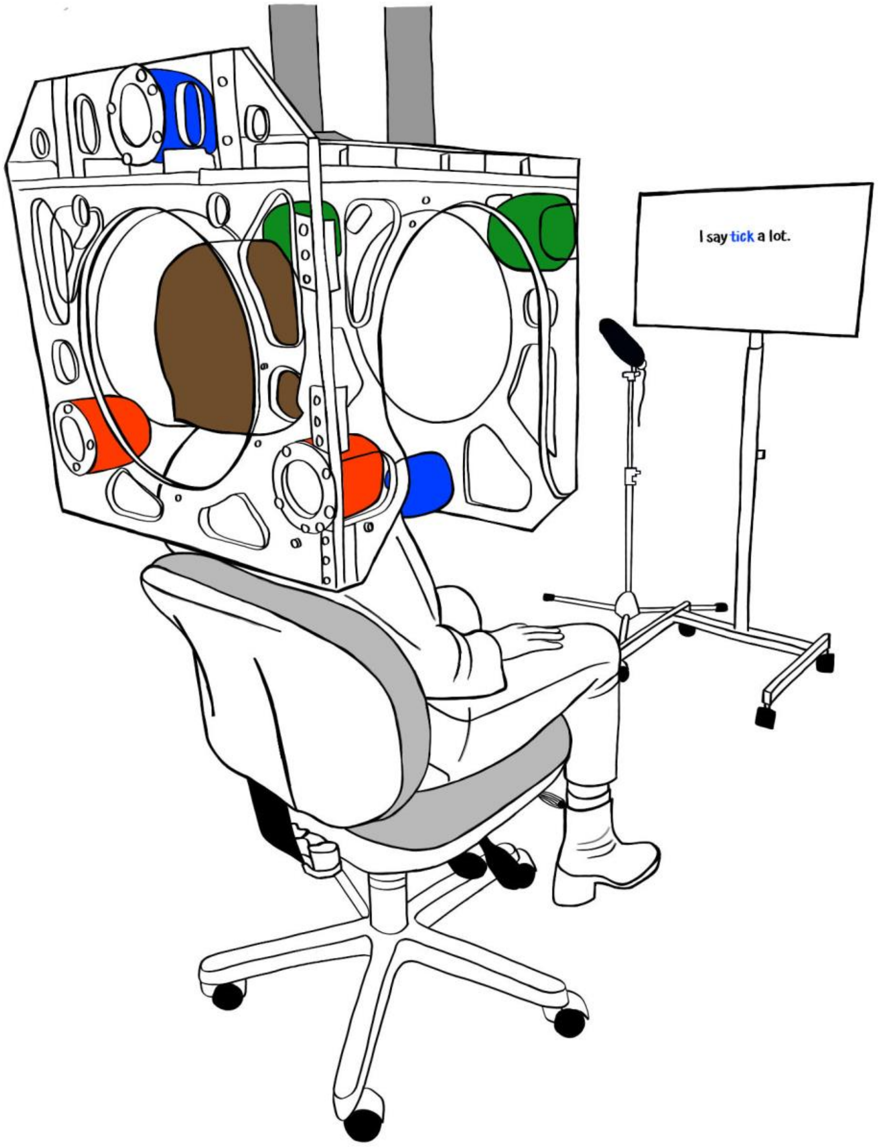

Figure 2

Schematic representation of the general setup for the *Kinematic task* with the AG500 articulograph. Participants sat inside the acrylic cube while reading sentences off a monitor.

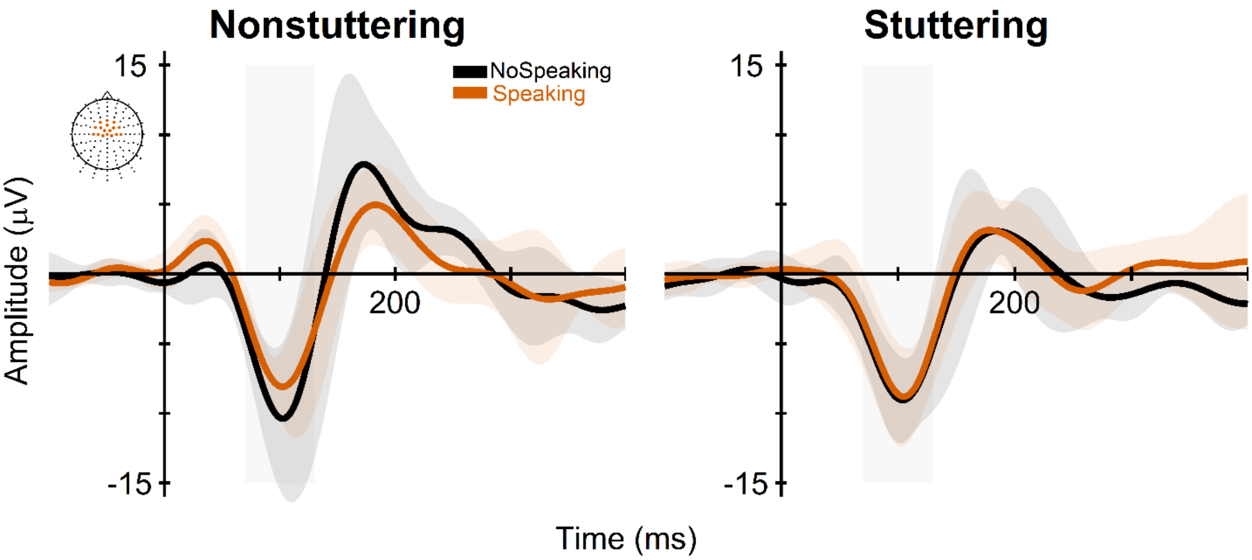

Figure 3

Grand average AEPs for both Nonstuttering (left) and Stuttering (right) groups. AEPs in the *Speaking* condition (orange) were elicited by a 1000 Hz probe tone presented 200 ms prior to the ‘go’ signal to begin producing a stimulus word (*tick*, *tech*, or *talk*). AEPs in the *No-Speaking* condition (black) were elicited by the same probe tone presented at an identical timepoint during a silent reading control condition (i.e., without speech movement planning). The shaded time window marks the N1 component analysis window from 70 to 130 ms post-stimulus onset. The top-down head inset shows the electrodes used for analysis.

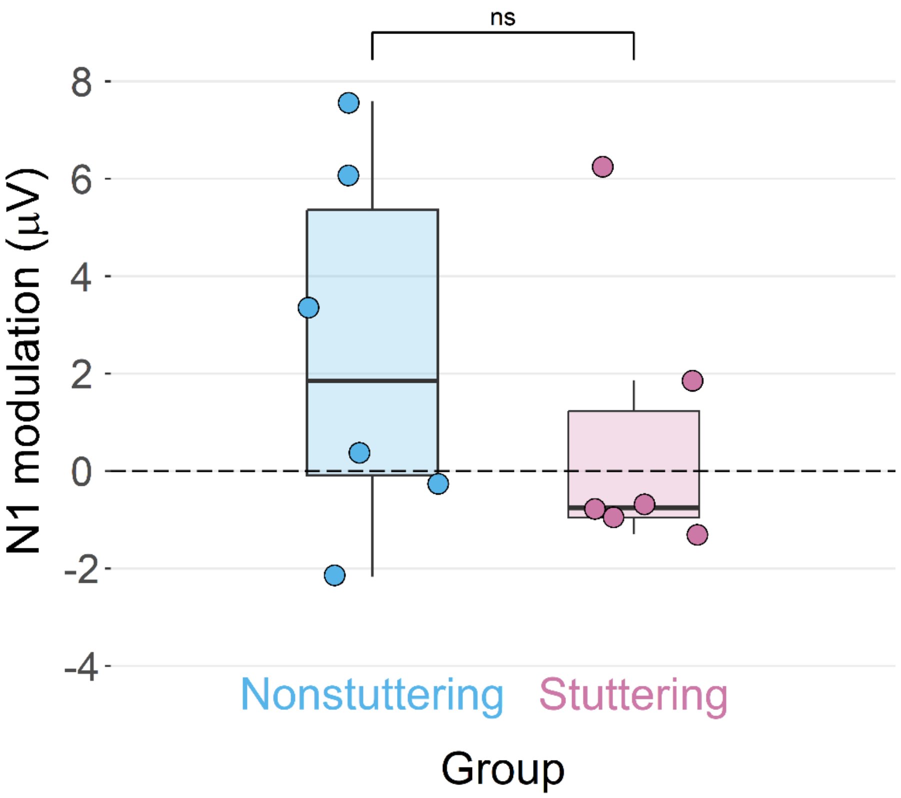

Figure 4

Distributions of PSAM values for both Nonstuttering (left, blue) and Stuttering (right, pink) groups (*N* = 5 adults each). Each colored point corresponds to one individual participant's amount of N1 modulation, calculated as the difference in N1 peak amplitude between *Speaking* and *No-Speaking* conditions. Positive N1 modulation values represent an N1 peak that is smaller in amplitude in the *Speaking* condition (i.e., evoked during speech movement planning). For each boxplot, the solid horizontal line marks the group median value, with box extents corresponding to the interquartile range (25th to 75th percentiles) and whiskers extending to the most extreme non-outlier values (within 1.5 times the interquartile range).

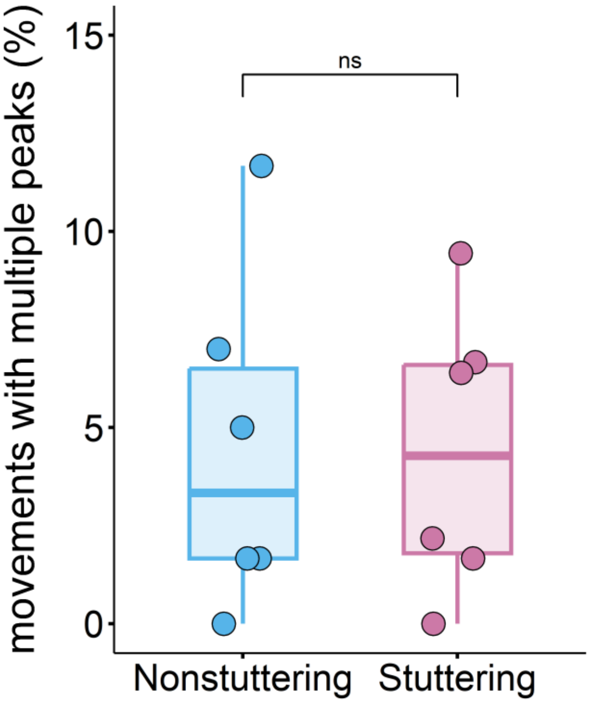

Figure 5

Group-level distributions of the percentage of movements that were rejected due to multiple peaks in the tangential velocity traces. Each individual datapoint corresponds to one participant, with percentages including both jaw and tongue tip movements. Boxplot visualization details as in Figure 4.

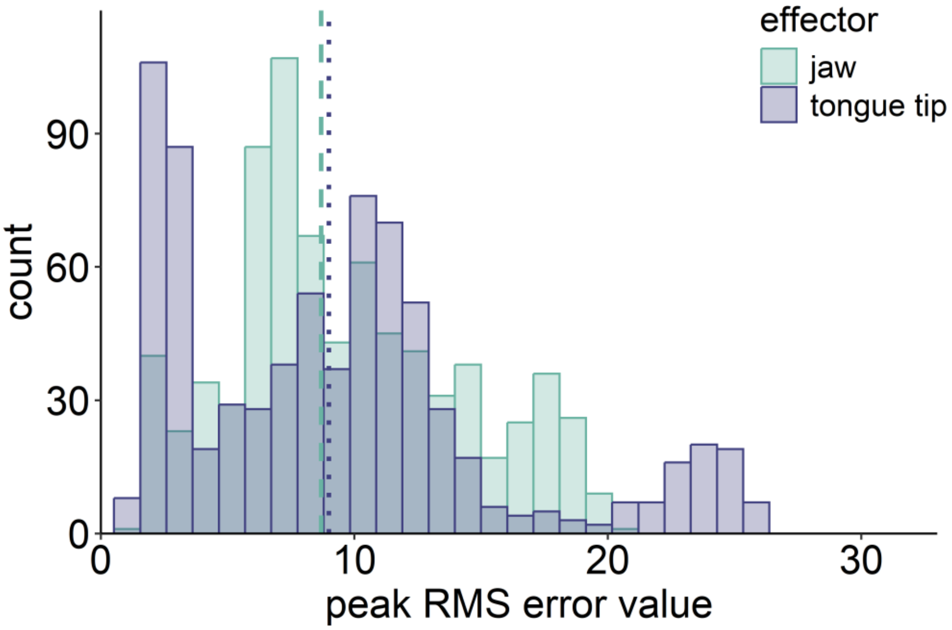

Figure 6

Histogram showing the distribution of peak RMS error values for all kinematic trials across all participants, for both the jaw (green) and tongue tip (purple) sensors. The median peak RMS value for each effector is also marked for both jaw (dashed line) and tongue tip (dotted line).

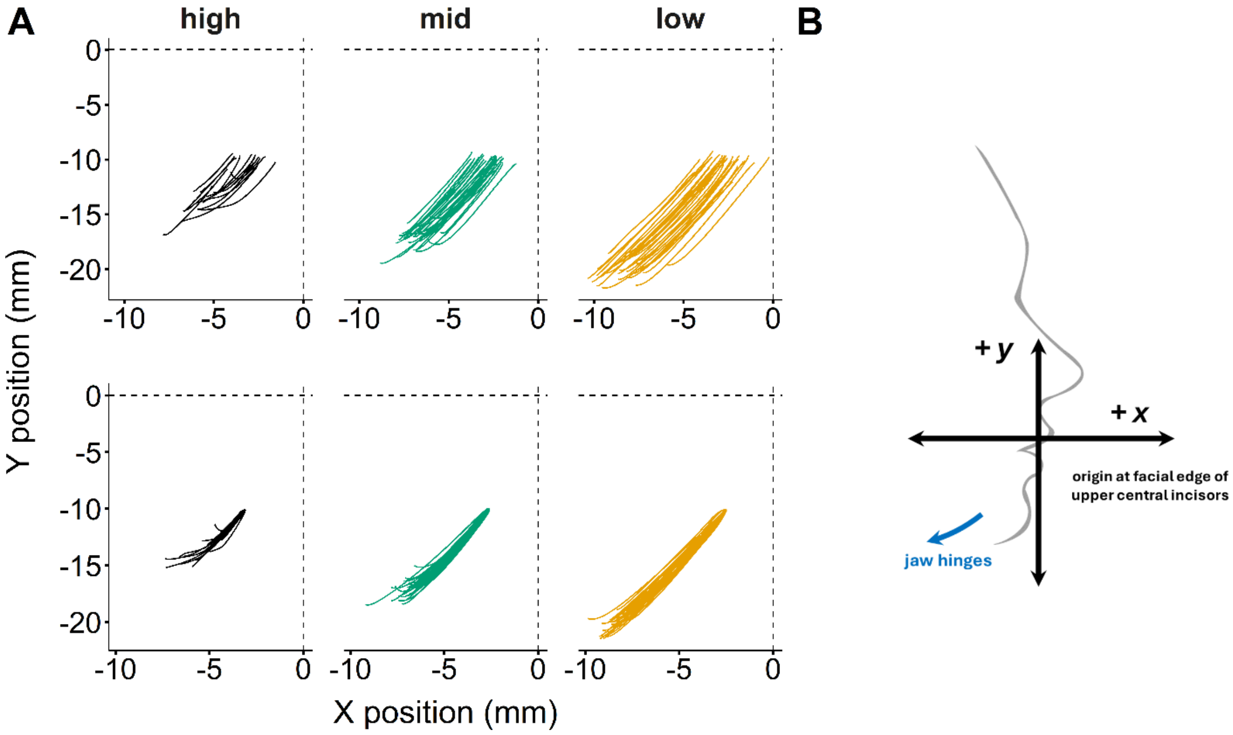

Figure 7

A) Example position data across jaw opening movements from a single Nonstuttering participant. Movement traces are shown for all utterances both in their original positions (top row) and after shifting all traces to start at the same mean starting point (bottom row). Data is divided by vowel identity to separate high (black), mid (green), and low (yellow) vowels (left to right columns, respectively). B) Schematic representation of the Cartesian plane used for analysis. Increasingly negative *x* and *y* values correspond to jaw opening movements as the jaw hinges.

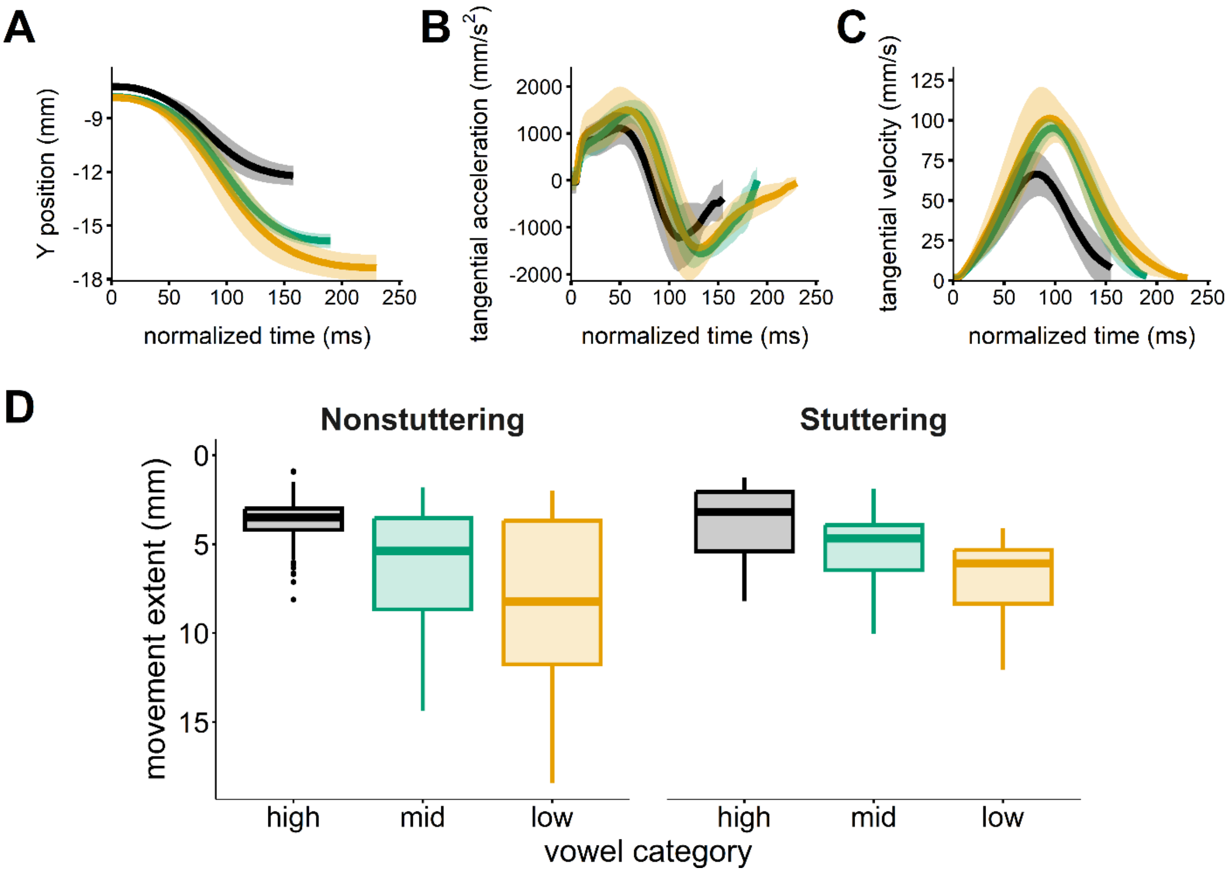

Figure 8

Kinematic variables scale with vowel height: example of a single Stuttering participant's opening movements for the jaw sensor only. Panels A-C show the *y* position, the tangential acceleration, and the tangential velocity as a function of normalized time, respectively. Solid lines correspond to the mean values across trials, with the shaded areas denoting one standard deviation around the mean. Panel D shows group-level boxplots for all jaw movement extents, divided by vowel height. Vowel category colors are as in Figure 6. Boxplot visualization details as in Figure 4.

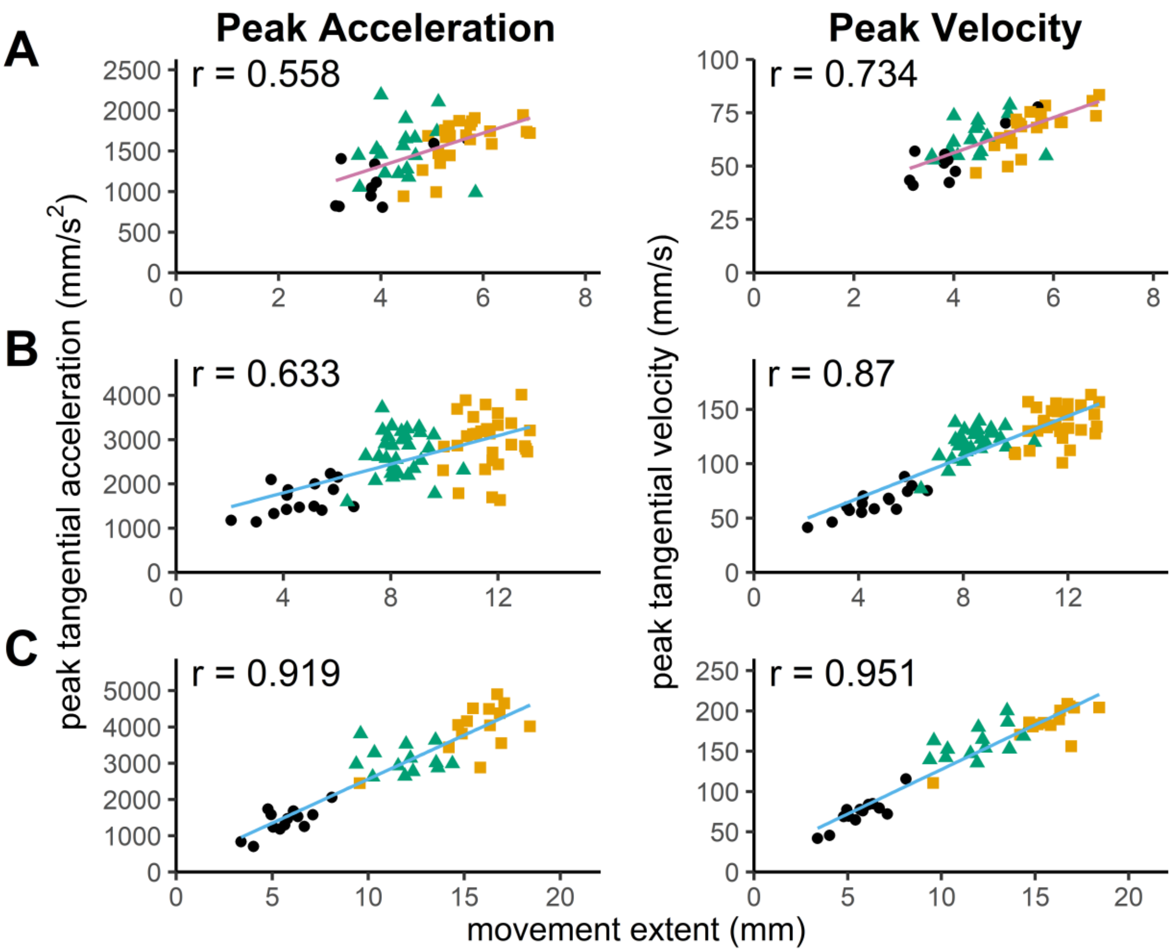

Figure 9

Example correlations between movement extent and both peak tangential acceleration (left column) and peak tangential velocity (right column) for jaw opening movements from three participants. Each row corresponds to data from one participant, with correlation values spanning from low to high (rows A-C, respectively). Row A is data from a Stuttering participant (pink trend line), rows B and C are data from Nonstuttering participants (blue trend lines). Individual datapoints represent individual utterances, divided by vowel: high (black dots), mid (green triangles), and low (yellow squares). Pearson correlation coefficients are given for each relationship.

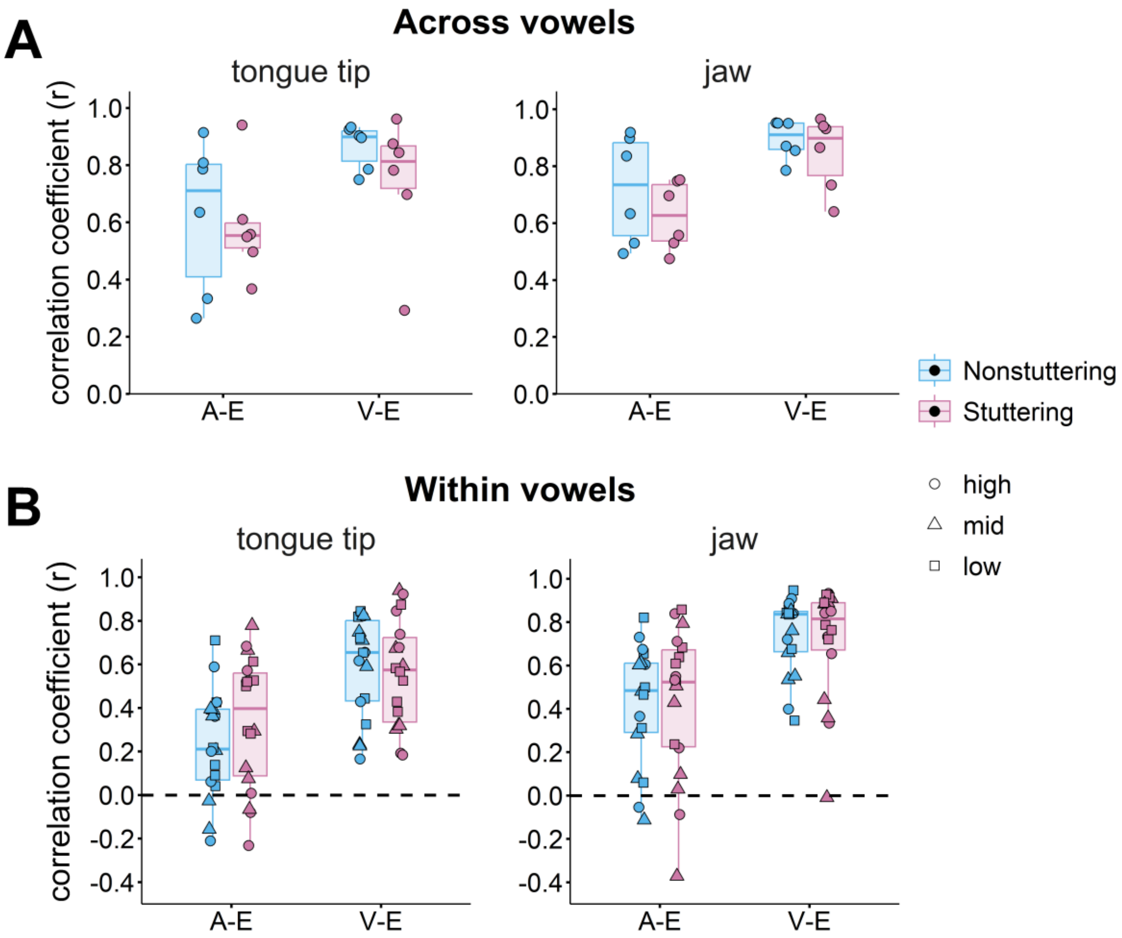

Figure 10

Group-level distributions of correlation coefficients between both peak tangential velocity and movement extent (V-E) and peak tangential acceleration and movement extent (A-E). Data are shown for both Nonstuttering (blue) and Stuttering (pink) groups as well as by effector (tongue tip on left, jaw on right). Panel A shows correlation coefficients collapsed across all vowels (top row), with each point corresponding to an individual participant. Panel B shows correlation coefficients within individual vowel categories, with each point representing a vowel group (high vowels as circles, mid vowels as triangles, and low vowels as squares) for a specific participant. Boxplot visualization details as in prior figures. Quantitative data summarized in Table 2.

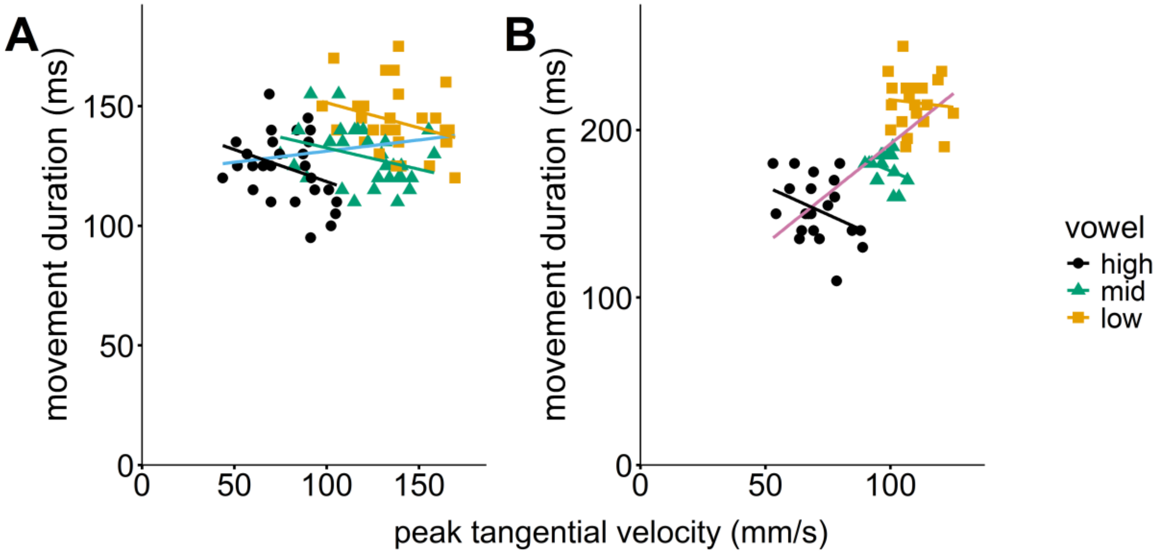

Figure 11

Examples of a positive correlation between peak tangential velocity and movement duration when calculating across vowels (correlation trendlines in pink for Stuttering, blue for Nonstuttering), but negative correlation values when calculating within individual vowel categories (high vowels in black, mid vowels in green, and low vowels in yellow). Example data are shown for both a Nonstuttering participant (A, tongue tip data only) and a Stuttering participant (B, jaw data only).

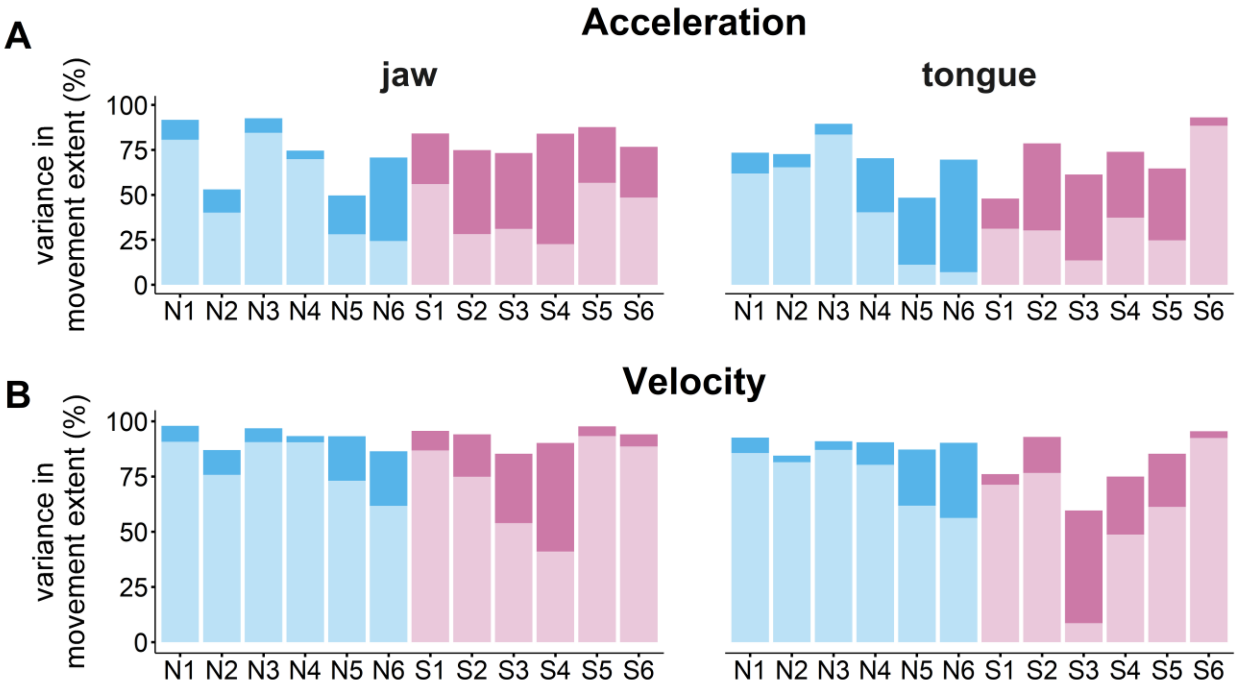

Figure 12

Individual barplots depicting the total percent variance explained in movement extent for all correlation models across participants. Results are divided by effector (jaw on left, tongue tip on right) and by kinematic variable (peak tangential acceleration in A, top row and peak tangential velocity in B, bottom row). N1-N6 and S1-S6 correspond to individual Nonstuttering (blue) or Stuttering (pink) participants, respectively. Light bars represent the variance explained in movement extent by the bivariate correlations (i.e., just the peak kinematic variables), our estimate of feedforward contributions. The darker bars highlight the *additional* variance explained by adding movement duration as a predictor to the same correlation models. The dark bars alone correspond to estimates of feedback-based contributions for each participant.

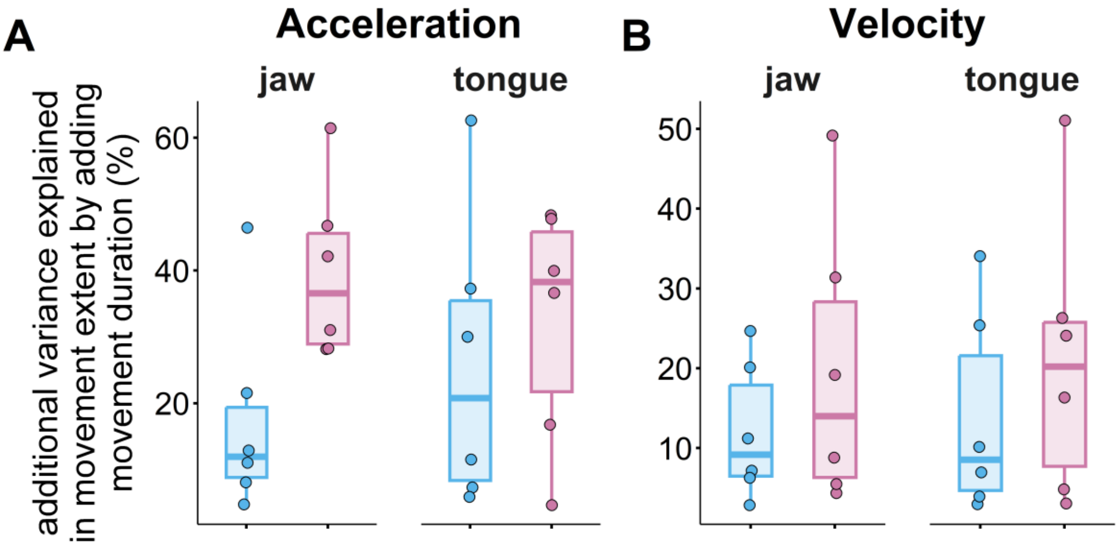

Figure 13

Group-level distributions of the additional variance explained in movement extent after the addition of movement duration as a predictor, our estimate of feedback-based contributions at the individual level. Data are shown for models across kinematic variables (peak tangential acceleration in A, peak tangential velocity in B) and by effector (jaw on left, tongue tip on right) for both Nonstuttering (blue) and Stuttering (pink) groups. Boxplot visualization details as in prior figures.

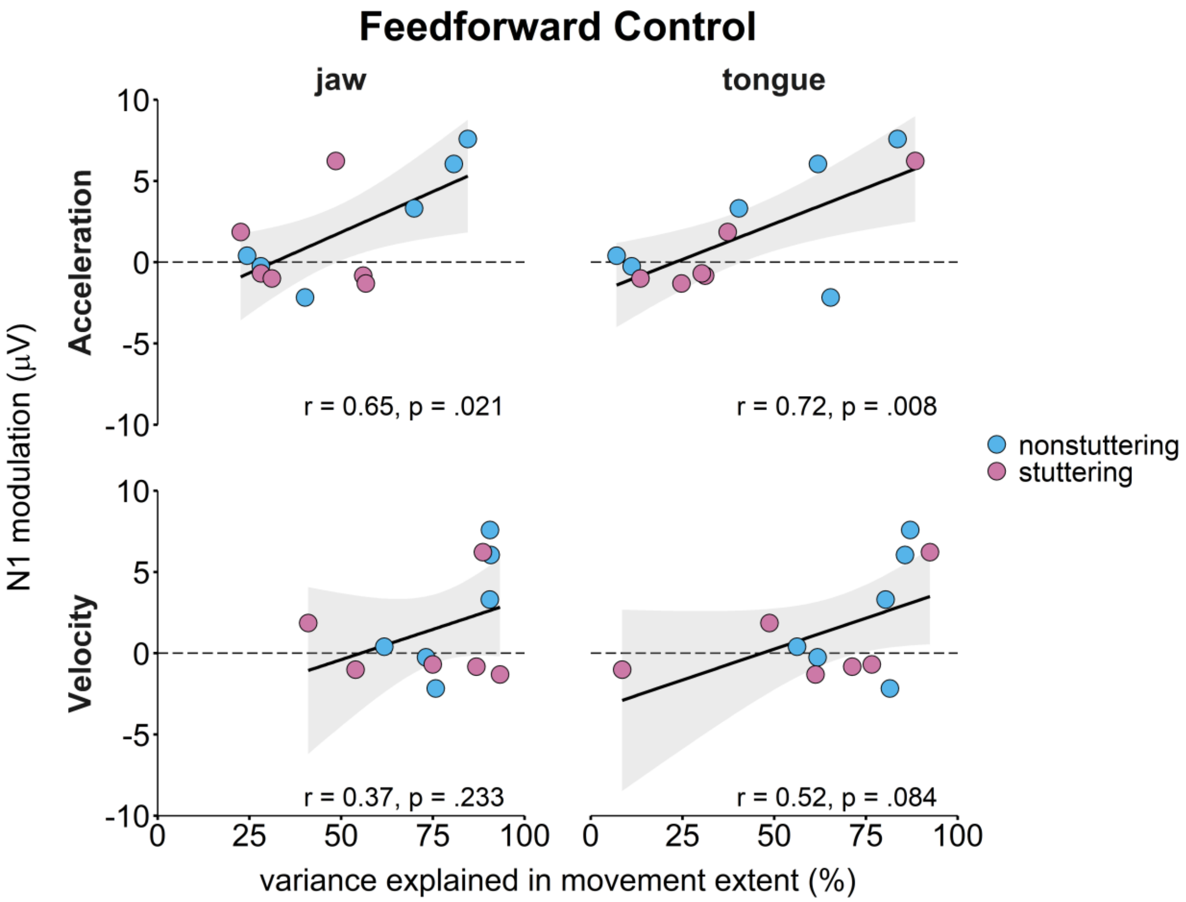

Figure 14

Correlations between pre-speech auditory modulation values and feedforward control estimates for all 4 models, divided by effector (jaw/tongue) and kinematic variable (peak tangential acceleration/velocity). Individual datapoints correspond to data from individual participants, colored based on group identity (Stuttering in pink, Nonstuttering in blue).

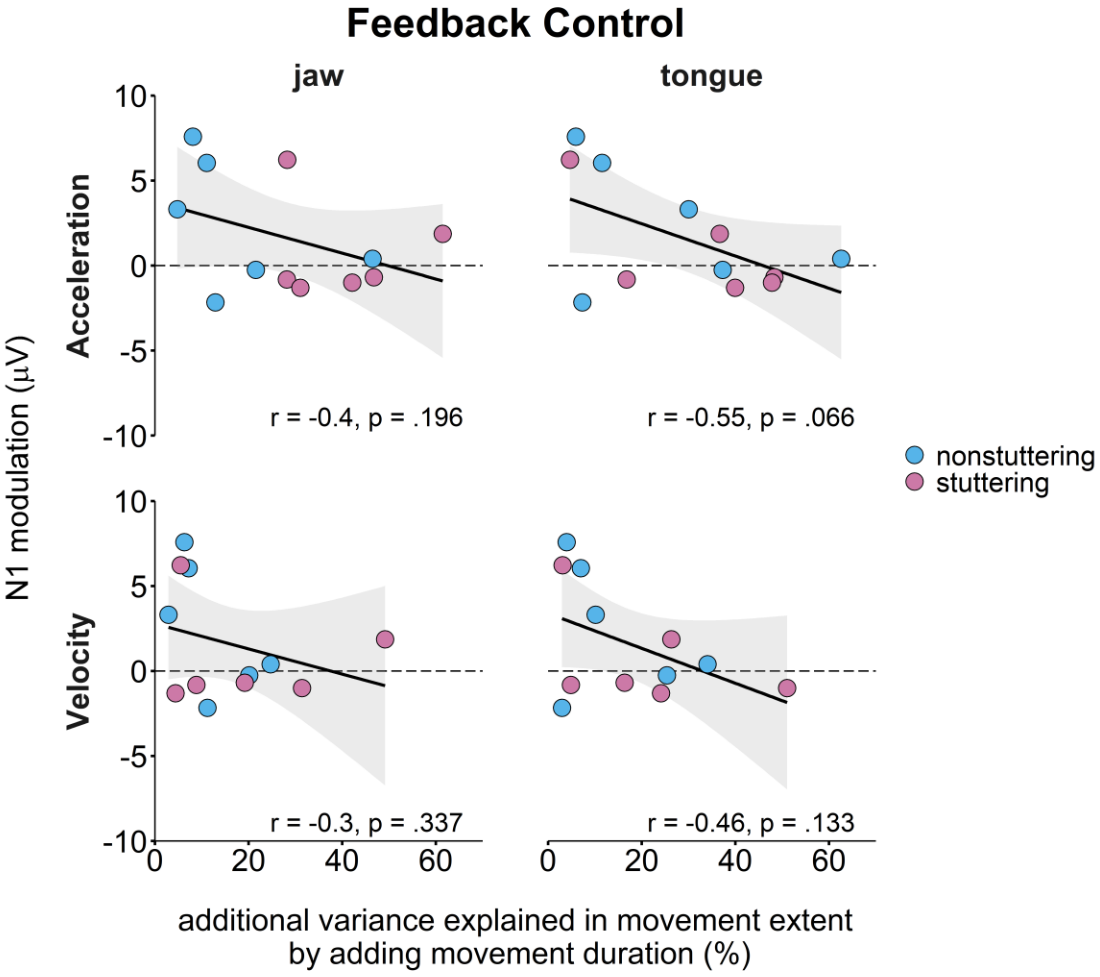

Figure 15

Correlations between pre-speech auditory modulation values and feedback control estimates for all 4 models. Visualization details as in previous figure.

10. TABLES

Table 1. *Demographic information for all Stuttering participants, sorted by their same-day mean stuttering frequency across both SSI tasks.*

Participant ID	Sex assigned at birth	Gender	Age (years)	Handedness	Reported age of onset (years)	Self-identifies as an IWS	SSI score	SSI severity	**Mean stuttering frequency (%SS)**
1	Female	Female	21	Right	"as long as I can remember"	Yes	17	very mild	**2.9**
2	Female	Female	20	Right	“when I began talking”	Yes	15	Very mild	**2.9**
3	Male	Male	24	Right	6-10	Yes	15	very mild	**3.7**
4	Male	Male	31	Right	9-10 (parent reports 4-5)	Yes	26	moderate	**8.2**
5	Male	Male	34	Right	"always"	Yes	28	moderate	**10.2**
6	Male	Male	19	Right	6	Yes	31	moderate	**15.0**

IWS = Individual Who Stutters. SSI = Stuttering Severity Instrument, 4th edition. %SS = percent stuttered syllables.

Table 2. *Mean bivariate correlation coefficients (Pearson's r) across both Stuttering and Nonstuttering groups for different kinematic relationships.*

	Across Vowels						**Within Each Vowel**					
	Tongue Tip			**Jaw**			**Tongue Tip**			**Jaw**		
	M	**SE**	**% sig.**	**M**	**SE**	**% sig.**	**M**	**SE**	**% sig.**	**M**	**SE**	**% sig.**
Nonstuttering												
acceleration-extent	0.624	0.109	100	0.718	0.077	100	0.235	0.058	33	0.421	0.065	56
velocity-extent	0.866	0.032	100	0.894	0.028	100	0.593	0.055	67	0.740	0.042	83
Stuttering												
acceleration-extent	0.587	0.078	100	0.627	0.049	100	0.338	0.072	44	0.433	0.081	56
velocity-extent	0.742	0.097	83	0.847	0.054	100	0.560	0.058	61	0.713	0.062	78

M = mean value. SE = standard error. % sig. = percent of correlations that are significant with *p* values < 0.05.

Section VI: General Discussion

1. Conclusions and future directions

There has been extensive discussion in the stuttering literature regarding underlying limitations in sensorimotor integration (Max, 2004; Max et al., 2004; Neilson & Neilson, 1987; Stromsta, 1972). Such limitations are supported by consistent, widespread neuroimaging findings across sensorimotor areas of atypical functional activation during speech, reduced white matter connectivity, and reduced gray matter volume (e.g., Belyk et al., 2015; Budde et al., 2014; Chang et al., 2008; Fox et al., 1996; Neef et al., 2015; Neef & Chang, 2024), as well as behavioral paradigms showing differences in sensorimotor learning and control as compared to individuals who do not stutter (Daliri et al., 2013; Kim et al., 2020; Kim & Max, 2021). Contemporary stuttering researchers thus often suggest that stuttering involves extensive network-level disruptions impacting speech sensorimotor control, with no evident locus of specific deficit (Brown et al., 2005; Gracco et al., 2022; Ludlow & Loucks, 2003; Neef & Chang, 2024).

This dissertation sought to extend our understanding of the sensorimotor mechanisms in stuttering by further investigating our established pre-speech auditory modulation (PSAM) paradigm. In adults who stutter, we have repeatedly documented a lack of modulation in auditory cortex responses immediately prior to speech production as compared with typically fluent adults (for a review, see Max & Daliri, 2019). Specifically, in nonstuttering speakers, the cortical auditory responses recorded in response to probe tones presented just prior to speech production are *reduced* in amplitude as compared to those recorded at an identical timepoint during a control condition with no speech movement planning. This difference in amplitude in auditory-evoked potentials due to the presence of speech movement planning is what we define as *pre-speech auditory modulation* – and was the main outcome of interest across this dissertation.

To further describe this behavior, *Study 1* investigated the feasibility of extending our PSAM paradigm to more complex utterances. We found that the same neurophysiological phenomenon could indeed be measured prior to sentence-level speech in typical speakers. Although it is still unknown to what extent motor planning varies across linguistic levels within the same paradigm (i.e., are we capturing identical pre-speech mechanisms despite subsequent production of sentences instead of monosyllabic words?), this task will allow for the future investigation of sensorimotor behaviors as a function of fluency *state*. In other words, we will be able to compare fluent speech with dysfluent speech *within* individuals who stutter to begin investigating the sensorimotor mechanisms that directly precede stuttered speech itself.

Study 2 explored pre-speech *somatosensory* mechanisms to investigate whether similar motor-to-sensory modulatory behaviors exist in the somatosensory domain as well. By experimentally applying orofacial skin stretches to index somatosensory processing, we did not observe any evidence of pre-speech somatosensory modulation in either typical speakers or speakers who stutter. We conclude that the specific mechanoreceptors that were probed by this skin stretch may not be under such motor-related modulatory effects. As evidence of movement-related somatosensory modulation has indeed been documented in the past and we did not observe such effects in the nonstuttering group in our paradigm, we could not conclude that there was a difference due to stuttering trait. Future work probing somatosensory processing by stimulating more speech-relevant effectors (e.g., proprioceptors at the level of the tongue or lip) could begin to investigate whether similar limitations also exist in pre-speech somatosensory modulation in stuttering.

Finally, *Study 3* was an initial exploration into the functional relevance of PSAM. By correlating kinematic estimates of reliance on feedback vs. feedforward control mechanisms with individual extents of modulation, we were able to start validating or refuting ongoing hypotheses for PSAM. Specifically, we previously considered PSAM to reflect a general priming of the auditory system in preparation for its upcoming role in auditory feedback monitoring (Max & Daliri, 2019). However, based on new findings showing worse auditory discrimination at the same timepoint prior to speech production (Wang et al., 2024) and evidence that PSAM correlates with reduced variability as early as the onset of speech (Li et al., 2024), our evolving alternative hypothesis suggests that PSAM may instead reflect feed*forward* processes. Indeed, *Study 3* provided preliminary evidence to corroborate this hypothesis, as increasing reliance on feedforward control correlated with increasing amounts of pre-speech auditory modulation.

Taken together, all three of these experiments incrementally advanced our understanding of movement-related sensory modulation in speech and generated new testable hypotheses to advance our understanding of the neural mechanisms of stuttering and speech sensorimotor control in general. Indeed, there is vast evidence of motor-to-sensory modulatory phenomena across human and nonhuman models, and many questions remain unanswered. Perhaps most intriguing for this specific line of research is the relationship of PSAM with the overt symptoms of stuttering themselves, as all prior studies have exclusively studied fluent speech *between* groups (i.e., comparing individuals who stutter with individuals who do not stutter). Future work extending sentence-level PSAM tasks to paradigms in which fluency status can be experimentally manipulated *within* stuttering participants would be an exciting next step towards understanding the speech behaviors so hallmark of the disorder. A deeper understanding of the neural mechanisms directly associated with overt dysfluencies could potentially inform novel therapeutic techniques aimed at bolstering a more fluent speech sensorimotor network (e.g., neurostimulation). Furthermore, clever paradigms that allow for a measure of PSAM and an index of feedback-based monitoring *within the same trial* instead of across experimental tasks (sometimes separated by weeks, as in *Study 3*) would enhance our understanding of the functional relevance of this mechanism and allow for the fine-tuning of our working hypothesis. Finally, future neurophysiological work should certainly focus on the developmental aspects of stuttering, as documented changes in neurobiological structure and function occur as children who stutter age – and, sometimes, spontaneously recover (Chang, 2014; Chang et al., 2015; Chow et al., 2023; Chow & Chang, 2017).

2. References

Belyk, M., Kraft, S. J., & Brown, S. (2015). Stuttering as a trait or state—An ALE meta-analysis of neuroimaging studies. *European Journal of Neuroscience*, *41*(2), 275–284. https://doi.org/10.1111/ejn.12765

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

Budde, K. S., Barron, D. S., & Fox, P. T. (2014). Stuttering, induced fluency, and natural fluency: A hierarchical series of activation likelihood estimation meta-analyses. *Brain and Language*, *139*, 99–107. https://doi.org/10.1016/j.bandl.2014.10.002

Chang, S.-E. (2014). Research updates in neuroimaging studies of children who stutter. *Seminars in Speech and Language*, *35*(2), 67–79. https://doi.org/10.1055/s-0034-1382151

Chang, S.-E., Erickson, K. I., Ambrose, N. G., Hasegawa-Johnson, M. A., & Ludlow, C. L. (2008). Brain anatomy differences in childhood stuttering. *NeuroImage*, *39*(3), 1333–1344. https://doi.org/10.1016/j.neuroimage.2007.09.067

Chang, S.-E., Zhu, D. C., Choo, A. L., & Angstadt, M. (2015). White matter neuroanatomical differences in young children who stutter. *Brain*, *138*(3), 694–711. https://doi.org/10.1093/brain/awu400

Chow, H. M., & Chang, S.-E. (2017). White matter developmental trajectories associated with persistence and recovery of childhood stuttering. *Human Brain Mapping*, *38*(7), 3345–3359. https://doi.org/10.1002/HBM.23590

Chow, H. M., Garnett, E. O., Koenraads, S. P. C., & Chang, S.-E. (2023). Brain developmental trajectories associated with childhood stuttering persistence and recovery. *Developmental Cognitive Neuroscience*, *60*, 101224. https://doi.org/10.1016/j.dcn.2023.101224