Auditory Modulation During Speech Planning in Stuttering and Nonstuttering Individuals

Ayoub Daliri

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

Doctor of Philosophy

University of Washington
2015

Reading Committee:
Ludo Max, Chair
David J. Perkel
Fred D. Minifie

Program Authorized to Offer Degree:
Speech and Hearing Sciences

University of Washington

Abstract

Auditory Modulation During Speech Planning in Stuttering and Nonstuttering Individuals

Ayoub Daliri

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

Stuttering is associated with atypical structural and functional connectivity among sensorimotor brain areas. However, it remains entirely unknown which specific mechanisms of sensorimotor control are affected by these neurological differences. In the program of research described here, I used a novel experimental paradigm and electroencephalographic (EEG) recordings to study motor-to-sensory interactions during speech movement planning in stuttering versus nonstuttering speakers.

Experiment 1 investigated whether stuttering adults are deficient in modulating the auditory system prior to speech initiation. Auditory modulation was examined by recording auditory evoked potentials in response to probe tones presented during movement planning in a delayed-response speaking condition as compared with no-speaking control conditions. Findings indicated that stuttering speakers did not show the modulation of auditory processing (reflected in reduced amplitude of the N1 component) that was observed in nonstuttering speakers.

This finding raised the question whether stuttering individuals have problems specifically with generating or evaluating a planning-related efference copy signal that can be used to predict upcoming self-generated sensory inputs or, more generally, with using any available information to make sensory predictions. In *Experiment 2*, probe tones were therefore delivered while participants anticipated either self-producing speech or hearing their own pre-recorded speech played back and in a control condition without auditory input. Results showed that auditory modulation differed between stuttering and normally fluent adults in both conditions with predictable auditory input.

Experiment 3 was designed to start exploring the functional significance of pre-speech auditory modulation in general, and the functional implications of stuttering speakers' lack of modulation. Participants in this experiment completed a sensorimotor adaptation task with formant-shifted auditory feedback, and the results served to estimate each speaker's reliance on auditory feedback. In a separate session, pre-speech auditory modulation was again assessed by means of probe tones, but this time N1 modulation relative to a no-speaking control condition was quantified both in a condition that allowed typical reliance on auditory feedback (non-delayed auditory feedback; NAF) and in a condition that did not allow reliance on auditory feedback (delayed auditory feedback; DAF). Results revealed that (a) stuttering speakers showed only limited adaptation to formant-shifted auditory feedback; (b) for nonstuttering speakers, DAF caused the amount of pre-speech auditory modulation to be reduced whereas for stuttering speakers, DAF enhanced pre-speech auditory modulation; and (c) across the two groups, there was a relationship between the effect of DAF on pre-speech auditory modulation and reliance on auditory feedback during the adaptation task.

These studies demonstrate that stuttering individuals have difficulties with using auditory predictions—both those related to active movement planning and those related to input that is not a consequence of one's own actions—to prime this sensory system with critical importance for speech production. Moreover, stuttering individuals showed not only a lack of modulation of the auditory system under normal speaking conditions (NAF) but also a lower reliance on auditory feedback as revealed here during a sensorimotor adaptation task with formant-shifted auditory feedback. Overall, findings suggest that stuttering is associated with deficits in auditory-motor integration, and that the auditory system may be not appropriately modulated for its role in online feedback control during speech production. I speculate that the inability to use predictive information for appropriately priming task-relevant sensory systems for their role in monitoring articulatory movements may lead to unnecessary and disruptive attempts at correcting ongoing movements. These maladaptive “repairs” may contribute to the fluency breakdowns that form the primary symptoms of stuttering.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER TWO: Modulation of auditory processing during speech movement planning is limited in adults who stutter

CHAPTER THREE: Electrophysiological evidence for a general auditory prediction deficit in adults who stutter

CHAPTER FOUR: Relationship between auditory modulation during speech planning and reliance on auditory feedback during speech production in adults who stutter

LIST OF TABLES

CHAPTER TWO: Modulation of auditory processing during speech movement planning is limited in adults who stutter

CHAPTER THREE: Electrophysiological evidence for a general auditory prediction deficit in adults who stutter

CHAPTER FOUR: Relationship between auditory modulation during speech planning and reliance on auditory feedback during speech production in adults who stutter

ACKNOWLEDGMENTS

I would like to express my most respectful gratitude to my advisor, Dr. Ludo Max, for his extraordinary guidance, patience, and kindness in educating me over the past years. Without his persistent help and guidance, this dissertation would not have been possible. I would like to thank my committee members, Dr. Fred Minifie, Dr. David Perkel, and Dr. Eberhard Fetz, for their extreme patience in guiding me throughout my general exam and dissertation.

I would like to thank all past and present members of MaxLab. I like to thank all my friends and classmates for all the great times we had. I would like to thank my family, especially my wife, for their continuous support. I would like to thank ASHAFoundation and Olswangs Endowed Graduate Student Conference Fund for their generous financial support.

DEDICATION

To my Family

CHAPTER ONE

Introduction

Developmental stuttering is a common speech disorder that is characterized by sound or syllable repetitions and audible or inaudible prolongations (Bloodstein & Bernstein-Ratner, 2008). Stuttering typically starts at 2–4 years of age and affects more boys than girls (Yairi & Ambrose, 2013). The prevalence of stuttering in the general population is approximately 1% (Bloodstein & Bernstein-Ratner, 2008).

Researchers have shown that stuttering is associated with breakdowns in fundamental processes of the sensorimotor system (for review, see Caruso, 1991; Max, 2004). Studies have used both neurological measures (such as brain structure and activation) and behavioral measures (such as parameters of speech and nonspeech movements) to investigate the sensorimotor system of stuttering individuals (Bloodstein & Bernstein-Ratner, 2008).

Neuroimaging studies have examined both structural and functional brain abnormalities in stuttering individuals. Functional imaging studies on stuttering have shown abnormal brain activation in a widespread network of pre-motor, motor, and sensory regions (Braun et al., 1997; Chang, Kenney, Loucks, & Ludlow, 2009; De Nil, Kroll, Kapur, & Houle, 2000; De Nil, Kroll, Lafaille, & Houle, 2003; Fox et al., 1996; Watkins, Smith, Davis, & Howell, 2008). Across individual stuttering subjects, the involvement of specific brain regions appears to vary considerably (Ingham, Wang, Ingham, Bothe, & Grafton, 2013; Wymbs, Ingham, Ingham, Paolini, & Grafton, 2013). Nevertheless, several structural brain abnormalities have been reported, and these abnormalities include atypical white matter in pathways suggested to connect speech motor and auditory regions (Chang, Erickson, Ambrose, Hasegawa-Johnson, & Ludlow, 2008; Chang, Horwitz, Ostuni, Reynolds, & Ludlow, 2011; Cykowski, Fox, Ingham, Ingham, &

Robin, 2010; Foundas et al., 2003; Sommer, Koch, Paulus, Weiller, & Buchel, 2002; Watkins et al., 2008).

Acoustic and physiological methods have been used to investigate several characteristics of the perceptually fluent speech of stuttering speakers (i.e., speech that does not include noticeable moments of stuttering). The purpose of these studies was to examine speech that was not contaminated by stuttering itself (or by the speaker's reactions to the stuttering) and to determine whether subtle neuromotor differences between stuttering and nonstuttering speakers can be detected even in the absence of overt stuttering. Acoustic analyses have shown several between-group differences such as longer voice onset times, longer stop gap durations, and longer vowel durations in stuttering individuals (Borden, Kim, & Spiegler, 1987; Healey & Gutkin, 1984; Max & Gracco, 2005). Using kinematic analyses of speech movements, studies have reported longer movement durations in stuttering individuals (Max, Caruso, & Gracco, 2003; McClean, Tasko, & Runyan, 2004; Zimmermann, 1980). In addition, studies have also demonstrated longer movement durations for stuttering individuals during various nonspeech movements (e.g., finger movements) (Bishop, Williams, & Cooper, 1991; Borden, 1983; Max et al., 2003; Webster & Ryan, 1991).

Given the results of these studies, other researchers and our own laboratory have proposed that stuttering may result from deficits in specific processes of sensorimotor integration that are critical for both early speech motor learning and mature speech motor control (Cai et al., 2012; Chang et al., 2011; Daliri, Prokopenko, & Max, 2013; Hickok, Houde, & Rong, 2011; Max, 2004; Watkins et al., 2008). To fully understand which specific aspects of sensorimotor integration are deficient in stuttering individuals, previous studies in our laboratory have examined the central nervous system's (CNS) prediction of the sensory consequences of planned limb and speech movements (Daliri, Prokopenko, Flanagan, & Max, 2014; Max, Daniels, Curet, & Cronin, 2008). This approach is based on theoretical models suggesting that, during movement planning, the CNS uses an efference copy (i.e., a copy of the prepared motor commands) and a forward internal model (i.e., a neural representation of the system's motor-to-sensory mapping) to predict the sensory consequences of planned movements (Shadmehr, Smith, & Krakauer, 2010; Wolpert, Diedrichsen, & Flanagan, 2011; Wolpert, Miall, & Kawato, 1998). Such predictions serve to implement anticipatory adjustments that take account of biomechanical effects (Flanagan & Lolley, 2001) as well as to prime task-relevant sensory systems for their role in processing incoming afferent signals (Creutzfeldt, Ojemann, & Lettich, 1989; Crone et al., 2001; Wasaka, Hoshiyama, Nakata, Nishihira, & Kakigi, 2003).

To examine the first of these roles of movement-related sensory prediction, we used a nonspeech task (fast arm movements to slide an object under the fingertip toward visual targets) to compare stuttering and nonstuttering individuals' ability to predict movement consequences related to the biomechanics of the upper limb (Daliri et al., 2014). We found that individuals who stutter accurately predict specific movement consequences in a ballistic reaching task (i.e., arm movements completed without relying on afferent feedback) in which those consequences could be fully compensated through anticipatory adjustments during movement planning. However, based on our overall theoretical framework (Max, 2004), stuttering individuals may be more likely to have difficulties with appropriately using such predictions to successfully "prime" task-relevant sensory systems for their subsequent role in (a) closely monitoring afferent inputs for online feedback control while (b) simultaneously preventing feedback-based motor responses that are undesirable during self-generated voluntary movements (note that the latter part of this hypothesis overlaps with ideas proposed by Zimmermann, 1980). Limited evidence consistent with this hypothesis was provided by McClean (1996) who demonstrated that, as compared with fluent speakers, stuttering adults show less attenuation of mechanically-evoked lip muscle reflexes prior to the onset of speech (with lip muscle activity measured in speech trials vs. no speech trials). When participating sensory systems (auditory, somatosensory) are insufficiently modulated in terms of their response to self-generated afferent inputs, the triggered motor responses may interfere with, and disrupt, ongoing movements. Given that McClean (1996) only measured peripheral lip activity, however, it remains unknown whether stuttering individuals' atypical sensorimotor responses at speech onset are in fact due to a lack of *central modulation* of sensory neural systems and whether such modulation problems can also be detected in the more dominant *auditory* system (Feng, Gracco, & Max, 2011). In the following chapters, I describe three related but independent experiments conducted to examine different aspects of pre-speech auditory modulation in stuttering adults versus nonstuttering adults.

Chapter Two (Study 1) describes the first study conducted to compare pre-speech auditory modulation in stuttering versus nonstuttering speakers. Auditory evoked potentials were recorded from 12 stuttering and 12 nonstuttering adults by presenting probe tones during the movement planning phase in a delayed-response speaking condition versus two no-speaking control conditions (silent reading and seeing nonlinguistic symbols).

Chapter Three (Study 2) describes a follow-up study designed to investigate whether the observed lack of pre-speech auditory modulation in stuttering speakers could be attributed specifically to problems with a movement planning-related efference copy signal or a more general problem with using any available information to make sensory predictions, even in the absence of movement planning. Auditory evoked potentials were recorded from 10 stuttering and 10 nonstuttering adults by presenting probe tones while subjects anticipated self-producing speech, while subjects anticipated hearing their own previously recorded speech being played back, and while subjects did not anticipate any auditory input.

Chapter Four (Study 3) describes a study that aimed to investigate whether stuttering speakers' observed lack of auditory modulation prior to both speaking and hearing one's own played-back speech is associated with either an increased or decreased reliance on auditory feedback during speech production. As a separate measure of 13 stuttering and 13 nonstuttering speakers' individual reliance on auditory feedback during speech, the amount of sensorimotor adaptation to formant-shifted auditory feedback was determined. Subsequently, auditory evoked potentials were recorded from the same adult participants in response to probe tones delivered prior to self-producing speech in a condition with unaltered auditory feedback, a condition with delayed auditory feedback, and a no-speaking control condition.

In Chapter Five, I summarize and discuss the integrated findings from all three experiments. I also present implications in the context of our overall theoretical framework regarding the sensorimotor control of speech movements in individuals who stutter, and I discuss promising future research directions.

References

Bishop, J. H.,Williams, H. G., & Cooper,W. A. (1991). Age and task complexity variables in motor performance of stuttering and nonstuttering children. *Journal of Fluency Disorders, 16*, 207–217.

Bloodstein, O., & Bernstein-Ratner, N. (2008). *A handbook on stuttering* (6th ed.). Clifton Park, NY: Delmar.

Borden, G. J. (1983). Initiation versus execution time during manual and oral counting by stutterers. *Journal of Speech and Hearing Research, 26,* 389–396.

Borden, G. J., Kim, D. H., & Spiegler, K. (1987). Acoustics of stop consonant vowel relationships during fluent and stuttered utterances. *Journal of Fluency Disorders, 12*(3), 175–184. doi:10.1016/0094-730X(87)90024-6

Braun, A., Varga, M., Stager, S., Schulz, G., Selbie, S., Maisog, J., . . . Ludlow, C. (1997). Altered patterns of cerebral activity during speech and language production in developmental stuttering - an H-2 O-15 positron emission tomography study. *Brain, 120,* 761-784. doi:10.1093/brain/120.5.761

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 motor-auditory transformation. *Plos One, 7*(7), e41830. doi:10.1371/journal.pone.0041830

Caruso, A. J. (1991). Neuromotor processes underlying stuttering. In H. F. M. Peters, W. Hulstijn & C. W. Starkweather (Eds.), *Speech motor control and stuttering* (pp. 101–116). Amsterdam, The Netherlands: Elsevier.

Chang, S., 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. doi:10.1016/j.neuroiiiiage.2007.09.067

Chang, S., Horwitz, B., Ostuni, J., Reynolds, R., & Ludlow, C. L. (2011). Evidence of left inferior frontal-premotor structural and functional connectivity deficits in adults who stutter. *Cerebral Cortex, 21*(11), 2507–2518. doi:10.1093/cercor/bhr028

Chang, S., Kenney, M. K., Loucks, T. M. J., & Ludlow, C. L. (2009). Brain activation abnormalities during speech and non-speech in stuttering speakers. *Neuroimage, 46*(1), 201-212. doi:10.1016/j.neuroimage.2009.01.066

Creutzfeldt, O., Ojemann, G., & Lettich, E. (1989). Neuronal-activity in the human lateral temporal-lobe .2. responses to the subjects own voice. *Experimental Brain Research, 77*(3), 476-489. doi:10.1007/BF00249601

Crone, N. E., Hao, L., Hart, J., Boatman, D., Lesser, R. P., Irizarry, R., & Gordon, B. (2001). Electrocorticographic gamma activity during word production in spoken and sign language. *Neurology, 57*(11), 2045–2053.

Cykowski, M. D., Fox, P. T., Ingham, R. J., Ingham, J. C., & Robin, D. A. (2010). A study of the reproducibility and etiology of diffusion anisotropy differences in developmental stuttering: A potential role for impaired myelination. *Neuroimage, 52*(4), 1495–1504. doi:10.1016/j.neuroimage.2010.05.011

Daliri, A., Prokopenko, R., & 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. doi:10.1044/1092-4388(2013/12-0134)

Daliri, A., Prokopenko, R. A., Flanagan, J. R., & Max, L. (2014). Control and prediction components of movement planning in stuttering versus nonstuttering adults. *Journal of*

Speech, Language, and Hearing Research, 57(6), 2131–2141. doi:10.1044/2014_JSLHR-S-13-0333

De Nil, L., Kroll, R., Kapur, S., & Houle, S. (2000). A positron emission tomography study of silent and oral single word reading in stuttering and nonstuttering adults. *Journal of Speech, Language, and Hearing Research, 43*(4), 1038–1053.

De Nil, L., Kroll, R., Lafaille, S., & Houle, S. (2003). A positron emission tomography study of short- and long-term treatment effects on functional brain activation in adults who stutter. *Journal of Fluency Disorders, 28*(4), 357–380. doi:10.1016/j.jfludis.2003.07.002

Flanagan, J. R., & Lolley, S. (2001). The inertial anisotropy of the arm is accurately predicted during movement planning. *The Journal of Neuroscience, 21*(4), 1361-1369.

Feng, Y., Gracco, V. L., & Max, L. (2011). Integration of auditory and somatosensory error signals in the neural control of speech movements. *Journal of Neurophysiology, 106*(2), 667–679. doi:10.1152/jn.00638.2010

Foundas, A. L., Corey, D. M., Angeles, V., Bollich, A. M., Crabtree-Hartman, E., & Heilman, K. M. (2003). Atypical cerebral laterality in adults with persistent developmental stuttering. *Neurology, 61*(10), 1378–1385.

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

Healey, E. C., & Gutkin, B. (1984). Analysis of stutterers voice onset times and fundamental-frequency contours during fluency. *Journal of Speech and Hearing Research, 27*(2), 219–225.

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

Ingham, R. J., Wang, Y., Ingham, J. C., Bothe, A. K., & Grafton, S. T. (2013). Regional brain activity change predicts responsiveness to treatment for stuttering in adults. *Brain and Language, 127*(3), 510–519.

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

Max, L., & Gracco, V. L. (2005). Coordination of oral and laryngeal movements in the perceptually fluent speech of adults who stutter. *Journal of Speech, Language, and Hearing Research, 48*(3), 524–542. doi:10.1044/1092-4388(2005/036)

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. doi:10.1044/1092-4388(2003/017)

Max, L., Daniels, J., Curet, K., & Cronin, K. (2008). Modulation of auditory and somatosensory processing during the planning of speech movements. *Paper presented at the Proceedings of the 8th International Seminar on Speech Production,* Strasbourg, France. 41–44.

McClean, M. D. (1996). Lip-muscle reflexes during speech movement preparation in stutterers. *Journal of Fluency Disorders, 21*(1), 49–60.

McClean, M. D., Tasko, S. M., & Runyan, C. M. (2004). Orofacial movements associated with fluent speech in persons who stutter. *Journal of Speech, Language, and Hearing Research, 47*(2), 294–303. doi:10.1044/1091-4388(2004/024)

Riley, G., & Riley, J. (1986). Oral motor discoordination among children who stutter. *Journal of Fluency Disorders, 11*(4), 335–344. doi:10.1016/0094-730X(86)90021-5

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

Sommer, M., Koch, M., Paulus, W., Weiller, C., & Buchel, C. (2002). Disconnection of speech-relevant brain areas in persistent developmental stuttering. *Lancet, 360*(9330), 380–383. doi:10.1016/S0140-6736(02)09610-1

Wasaka, T., Hoshiyama, M., Nakata, H., Nishihira, Y., & Kakigi, R. (2003). Gating of somatosensory evoked magnetic fields during the preparatory period of self-initiated finger movement. *Neuroimage, 20*(3), 1830–1838. doi:10.1016/S1053-8119(03)00442-7

Watkins, K. E., Smith, S. M., Davis, S., & Howell, P. (2008). Structural and functional abnormalities of the motor system in developmental stuttering. *Brain, 131*, 50–59. doi:10.1093/brain/awm241

Webster, W. G., & Ryan, C. R. L. (1991). Task complexity and manual reaction-times in people who stutter. *Journal of Speech and Hearing Research, 34*, 708–714.

Wolpert, D. M., Diedrichsen, J., & Flanagan, J. R. (2011). Principles of sensorimotor learning. *Nature Reviews Neuroscience, 12*(12), 739–751.

Wolpert, D. M., Miall, R. C., & Kawato, M. (1998). Internal models in the cerebellum. *Trends in Cognitive Sciences, 2*(9), 338–347. doi:10.1016/S1364-6613(98)01221-2

Wymbs, N. F., Ingham, R. J., Ingham, J. C., Paolini, K. E., & Grafton, S. T. (2013). Individual differences in neural regions functionally related to real and imagined stuttering. *Brain and Language, 124*(2), 153–164.

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

Zimmermann, G. (1980). Stuttering - a disorder of movement. *Journal of Speech and Hearing Research, 23*(1), 122–136.

CHAPTER TWO

Modulation of auditory processing during speech movement planning is limited in adults who stutter

Ayoub Daliri and Ludo Max

Manuscript accepted for publication in *Brain and Language*

Abstract

Stuttering is associated with atypical structural and functional connectivity in sensorimotor brain areas, in particular premotor, motor, and auditory regions. It remains unknown, however, which specific mechanisms of speech planning and execution are affected by these neurological abnormalities. To investigate pre-movement sensory modulation, we recorded 12 stuttering and 12 nonstuttering adults' auditory evoked potentials in response to probe tones presented prior to speech onset in a delayed-response speaking condition vs. no-speaking control conditions (silent reading; seeing nonlinguistic symbols). Findings indicate that, during speech movement planning, the nonstuttering group showed a statistically significant modulation of auditory processing (reduced N1 amplitude) that was not observed in the stuttering group. Thus, the obtained results provide electrophysiological evidence in support of the hypothesis that stuttering is associated with deficiencies in modulating the cortical auditory system during speech movement planning. This specific sensorimotor integration deficiency may contribute to inefficient feedback monitoring and, consequently, speech dysfluencies.

Introduction

Stuttering is a disorder of speech fluency associated with abnormal brain *activation* in a widespread network of pre-motor, motor, and sensory regions (Braun et al., 1997; Chang, Kenney, Loucks, & Ludlow, 2009; De Nil, Kroll, Kapur, & Houle, 2000; De Nil, Kroll, Lafaille, & Houle, 2003; Fox et al., 1996; Neumann et al., 2003; Watkins, Smith, Davis, & Howell, 2008). Across individual stuttering subjects, the involvement of specific brain regions appears to vary considerably (Ingham, Wang, Ingham, Bothe, & Grafton, 2013; Wymbs, Ingham, Ingham, Paolini, & Grafton, 2013). Nevertheless, several *structural* brain abnormalities have been reported, and these abnormalities include atypical white matter in pathways suggested to connect speech motor and auditory regions (Chang, Erickson, Ambrose, Hasegawa-Johnson, & Ludlow, 2008; Chang, Horwitz, Ostuni, Reynolds, & Ludlow, 2011; Cykowski, Fox, Ingham, Ingham, & Robin, 2010; Foundas et al., 2003; Sommer, Koch, Paulus, Weiller, & Buchel, 2002; Watkins et al., 2008). Accordingly, some of the most prominent contemporary theoretical views of stuttering suggest that the disorder may result from deficits in specific processes of sensorimotor integration that are critical for both early speech motor learning and mature speech motor control (Beal et al., 2010; Brown, Ingham, Ingham, Laird, & Fox, 2005; Cai et al., 2012; Chang et al., 2011; Daliri, Prokopenko, & Max, 2013; Hickok, Houde, & Rong, 2011; Kell et al., 2009; Liotti et al., 2010; Max, 2004; Watkins et al., 2008).

One aspect of sensorimotor integration that may be of particular theoretical importance in this regard is the central nervous system’s (CNS) *prediction* of the sensory consequences (or, more generally, movement outcomes) of planned motor commands. In a recent study (Daliri, Prokopenko, Flanagan, & Max, 2014), we found that individuals who stutter accurately predict specific movement consequences in a ballistic reaching task (i.e., arm movements completed without relying on afferent feedback) in which those consequences could be fully compensated through anticipatory adjustments during movement planning. However, based on our overall theoretical framework (Max, 2004), stuttering individuals may be more likely to have difficulties with appropriately using such predictions to successfully "prime" task-relevant *sensory systems* for their subsequent role in (a) closely monitoring afferent inputs for online feedback control while (b) simultaneously preventing feedback-based motor responses that are undesirable during self-generated voluntary movements (note that the latter part of this hypothesis overlaps with ideas proposed by Zimmermann, 1980). Limited evidence consistent with this hypothesis was already provided by McClean (1996) who demonstrated that, as compared with fluent speakers, stuttering adults show less attenuation of mechanically-evoked lip muscle reflexes prior to the onset of speech (with lip muscle activity measured in speech trials vs. no speech trials). When participating sensory systems (auditory, somatosensory) are insufficiently modulated in terms of their response to self-generated afferent inputs, the triggered motor responses may interfere with, and disrupt, ongoing movements. To date, however, it remains completely unknown (a) whether stuttering individuals' atypical sensorimotor responses at speech onset are in fact due to a lack of *central modulation* of sensory neural systems, and, if so, (b) whether stuttering individuals show a lack of pre-speech sensory modulation in the *auditory cortical regions* that have been implicated in several, although not all, neuroimaging studies (see above).

Here, we addressed both these questions directly by using electroencephalographical (EEG) data and auditory evoked potential analyses to investigate, in stuttering vs. nonstuttering adults, the modulation of auditory cortical activity in response to probe tones presented prior to speaking (i.e., during speech movement planning) and in control conditions without preparation for motor activity[1]. Using an experimental paradigm that we previously developed for work with

---

[1] Note that, by investigating sensory systems during movement planning, this paradigm addresses neural processes that are distinct from those investigated with another recent paradigm in which stuttering and

typically fluent speakers (Max et al., 2008), we recorded long latency auditory evoked potentials (LLAEPs) in response to auditory stimuli presented during the delay phase of a delayed-response *speaking* task (seeing a word on a monitor, silently reading the word, and saying it aloud after a *go* signal), a *silent reading* task (seeing a word and silently reading it), and a *seeing* task (seeing nonlinguistic symbols). We also recorded the same LLAEPs in a standard eyes-closed rest condition to compare both groups in terms of basic auditory processing in the absence of an active task, and to verify, through comparison with these reference data, the validity of data processing and analysis procedures used in the three active tasks. Analyses focused on the amplitude and latency of the LLAEP components N1 and P2[2]. We hypothesized that if stuttering is associated with a lack of modulation of auditory cortical regions prior to speech onset, the stuttering group would fail to show the typical N1 amplitude attenuation that we have previously documented for normally fluent speakers (Max et al., 2008).

Method

Participants

Twelve right-handed stuttering adults (eleven men and one woman; M_{age} = 27.32 years, age range: 18–46 years) and twelve right-handed nonstuttering adults (eleven men and one woman; M_{age} = 27.25 years, age range: 19–45 years) participated in the experiment after providing informed consent. Nonstuttering participants were individually matched with the stuttering participants based on age (± 3 years) and sex. All participants were naive to the

---

nonstuttering speakers have been compared in terms of auditory responses to their own speech during speech production (Beal et al., 2010, 2011; Liotti et al., 2010).

[2] The prominent LLAEP component that peaks ~70–130 ms after stimulus onset is typically labeled N100 or N1 in EEG recordings and M100 or N1m in magnetoencephalographic (MEG) recordings. This component is known to be generated by neuronal populations in the primary auditory cortex (Godey, Schwartz, De Graaf, Chauvel, & Liegeois-Chauvel, 2001; Näätänen & Picton, 1987; Zouridakis, Simos, & Papanicolaou, 1998). A similarly prominent component that peaks ~150–250 ms after stimulus onset is typically labeled P200 or P2 in EEG recordings and M200 or P2m in MEG recordings. The neural sources of the latter component have been shown to be located more anterior in auditory cortex (Papanicolaou, Baumann, Rogers, Saydjari, Amparo, & Eisenberg; Ross & Tremblay, 2009).

purpose of the study.

Eligibility criteria for all participants included (a) being a native speaker of American English, (b) self-reported absence of psychological, neurological, or communication disorders (other than stuttering in the stuttering group), (c) not taking any medications with possible effects on sensorimotor functioning, and (d) pure tone behavioral hearing thresholds at or below 20 dB HL at all octave frequencies from 250 Hz to 8 kHz in both ears. Using the Stuttering Severity Instrument, Fourth Edition (SSI–4; Riley, 2008), each stuttering participant's severity was determined by an American Speech-Language-Hearing Association-certified speech-language pathologist. Individual participant information for the stuttering group (age, sex, handedness, overall SSI–4 score, stuttering severity classification, and frequency of stuttering averaged across the SSI–4 speaking and reading tasks) are presented in Table 1.

Table 1. Individual participant information for the stuttering group (SSI score = Stuttering Severity Instrument 4th ed. overall score; Stuttering frequency (%SS) = percent stuttered syllables averaged across the SSI speaking and reading tasks).

Stuttering participant	Age (years)	Sex	Handedness	SSI score	SSI severity	Stuttering frequency (%SS)
1	18	male	right	22	mild	5.5
2	19	male	right	15	very mild	2.8
3	20	male	right	30	moderate	9.4
4	21	male	right	12	very mild	1.8
5	25	male	right	14	very mild	2.8
6	27	female	right	13	very mild	2.4
7	28	male	right	28	moderate	10.3
8	29	male	right	36	severe	21.1

10	30	male	right	26	moderate	15.3
9	31	male	right	25	moderate	9.3
11	33	male	right	19	mild	3.7
12	46	male	right	23	mild	5.4

Procedure and Instrumentation

The experiment was conducted inside a sound-attenuated room. Wearing an electrode cap (details given below), participants were seated approximately 1 m from a 23-inch liquid crystal display (LCD) monitor with a refresh rate of 60 Hz. Participants' speech output was transduced and amplified (WL185, Shure Incorporated, Niles, IL; DPS II, ART ProAudio, Niagara Falls, NY) and, after amplification by a headphones amplifier (S-phone, Samson Technologies Corp., Syosset, NY), played-back to the participant in real-time through insert earphones (ER-3A, Etymotic Research Inc., Grove Village, IL). The insert earphones were also used to deliver binaural auditory stimuli (1 kHz, 50 ms duration, 10 ms rise/fall time, 75 dB SPL) during some trials. Before each recording session, this feedback system was calibrated such that speech input with an intensity of 75 dB SPL at the microphone (approximately 15 cm from the participant's mouth) resulted in 72 dB SPL output in the earphones (Cornelisse, Gagné, & Seewald, 1991). For calibration, the intensity of the auditory feedback in the earphones was measured using a 2 cc coupler (Type 4946, Bruel & Kjaer Inc., Norcross, GA) connected to a sound level meter (Type 2250A Hand Held Analyzer with Type 4947 ½" Pressure Field Microphone, Bruel & Kjaer Inc., Norcross, GA).

Continuous EEG was recorded in three conditions: *speaking*, *reading*, and *seeing*. Each condition consisted of 270 trials (3 blocks of 90 trials). In each block, binaural auditory stimuli were delivered through the insert earphones during one third of the trials (tone trials) whereas no auditory stimuli were presented in the remaining trials (no-tone trials). The order of the conditions was counterbalanced across participants in each group.

In the *speaking* condition (Fig. 1A), each trial started with the presentation of a word in white characters on a black background on the display. After 600 ms, the color of the word changed to green. This change of color constituted a *go* signal for the participant to say the word aloud. In the *reading* condition (Fig. 1B), the procedure was the same as in the speaking condition except that participants were instructed to read the word silently without any movements—thus, the motor component of the task was eliminated. In the *seeing* condition (Fig. 1C), the procedure was the same as in the *reading* condition except that nonlinguistic symbols ("++++") were shown rather than words—thus, both the cognitive-linguistic activity associated with reading and the motor activity were eliminated. For the tone trials in all three conditions, auditory stimuli were delivered 400 ms after presentation of the word/symbols in white color (Fig. 1D). Each trial ended 500 ms after the color of the word/symbols changed to green. The inter-stimulus-interval from the end of a trial to the beginning of the next trial was randomly selected from a set of five possible intervals (1500, 2000, 2500, 3000, or 3500 ms).

The words presented in the *speaking* and *reading* conditions were from two different word lists (list A and list B). Half of the participants in each group used list A for the *speaking* condition and list B for the *reading* condition; the other half of the participants used list B for the *speaking* condition and list A for the *reading* condition. Each list contained 90 different words. Only monosyllabic consonant-vowel-consonant (CVC) words were included. All words were 3–5 letters long, and none of the words included consonant clusters.

In addition to the *speaking*, *reading*, and *seeing* conditions, we also recorded continuous EEG during an eyes-closed *rest* condition. The same binaural auditory stimuli (90 trials) were presented. Inter-stimulus intervals were randomly selected from the same intervals as listed above. These rest data mainly served to validate the morphology, amplitudes, and latencies of the other conditions' LLAEPs derived from the continuous EEG after the processing steps described below, but they also allowed a statistical comparison of the stuttering and nonstuttering groups' LLAEPs during rest.

Electroencephalographic Recordings

We used a Biosemi active-electrode EEG system, with electrodes mounted in a nylon head-cap (Biosemi Inc., Amsterdam, The Netherlands). EEG signals were recorded from 128 standard sites on the scalp (Fig. 1F) according to an extension of the international 10-10 electrode system (Gilmore, 1994; Oostenveld & Praamstra, 2001). Two electrodes were placed over the left and right mastoids for off-line re-referencing.

To record electrooculograms (EOGs), additional electrodes were placed below the lower eyelid and next to the outer canthus of the left eye. The EOG signals were used to detect and reject artifacts related to blinking and eye movements. Orofacial electromyograms (EMGs) were recorded by using four electrodes placed on the skin overlying muscles on the right side of the face (masseter: jaw elevation; anterior belly of the digastric: jaw depression; orbicularis oris: upper and lower lip elevation/rounding). The EMG signals were used to verify an absence of active muscle contraction prior to the *go* signal, and to estimate the onset time of speech movements in the speaking condition. The latter data showed that the average onset time for the stuttering group was not statistically significantly different from the average onset time for the nonstuttering group, $t(22) = 1.236$, $p = .229$.

The acoustic signal from an additional microphone (SM58, Shure, Niles, IL) was recorded together with the EEG data. All signals (EEG, EMG, EOG, and speech acoustics) were recorded with a sampling rate of 1024 Hz.

Data analysis

Data analyses were conducted using the EEGLAB toolbox (Delorme & Makeig, 2004) and custom-written MATLAB scripts (The MathWorks, Inc., Natick, MA). The EEG signals were first re-referenced to an off-line reconstructed average mastoids reference. The signals were then filtered off-line using a low-pass filter with cut-off frequency of 50 Hz. Next, the continuous signals were segmented into epochs ranging from 100 ms before to 400 ms after the onset of the auditory stimulus (and the equivalent time window in the no-tone trials). For each epoch, the mean amplitude of the pre-stimulus period (100 ms) was subtracted from the whole epoch to remove baseline differences. Epochs with EEG amplitudes greater than ±70 µV were excluded. Trials were visually inspected to exclude epochs contaminated by artifacts associated with (a) speech movements that started before the *go* signal—based on EMG activity, and (b) blinking or eye movements—based on the EOG signal.

After these processing steps, the averaged response for the tone trials reflected auditory as well as non-auditory activity (e.g., activity related to motor, linguistic, cognitive, and visual processes), whereas the averaged response for the no-tone trials reflected solely the non-auditory activity. Hence, to isolate the auditory activity evoked by the probe tones, each participant's averaged response for the no-tone trials from each condition was subtracted from the averaged response for the tone trials of the same condition (Fig. 1E). This subtracted signal was then used for all LLAEP analyses as it provides the best estimate of the true auditory response (Baess, Jacobsen, & Schroger, 2008; Baess, Horvath, Jacobsen, & Schroeger, 2011; Martikainen, Kaneko, & Hari, 2005; Max et al., 2008). In the rest condition, the LLAEPs were obtained directly by averaging epochs time-locked to the auditory stimuli. The final LLAEPs for each channel were low-pass filtered with a cut-off frequency of 15 Hz.

In the last step, overall LLAEPs for each of three regions of interests (ROIs) were obtained by averaging the LLAEPs from 6 electrodes over the left hemisphere (Left ROI: electrodes D11, D12, D19, D20, D27, D28), 6 electrodes over the central region (Central ROI: electrodes A1, A2, B1, C1, D1, D15), and 6 electrodes over the right hemisphere (Right ROI: electrodes B17, B18, B21, B22, B30, B31). The three ROIs and the included individual channels are illustrated in Fig. 1F. As dependent variables, the latency and amplitude of the N1 and P2 components were extracted for each participant's ROI LLAEPs. N1 was defined as the largest negative peak between 70 and 130 ms following the onset of the auditory stimulus. P2 was defined as the largest positive peak between 150 and 250 ms following the onset of the auditory stimulus. Surface distribution maps of the brain's auditory responses were also created, using the data from all 128 electrodes.

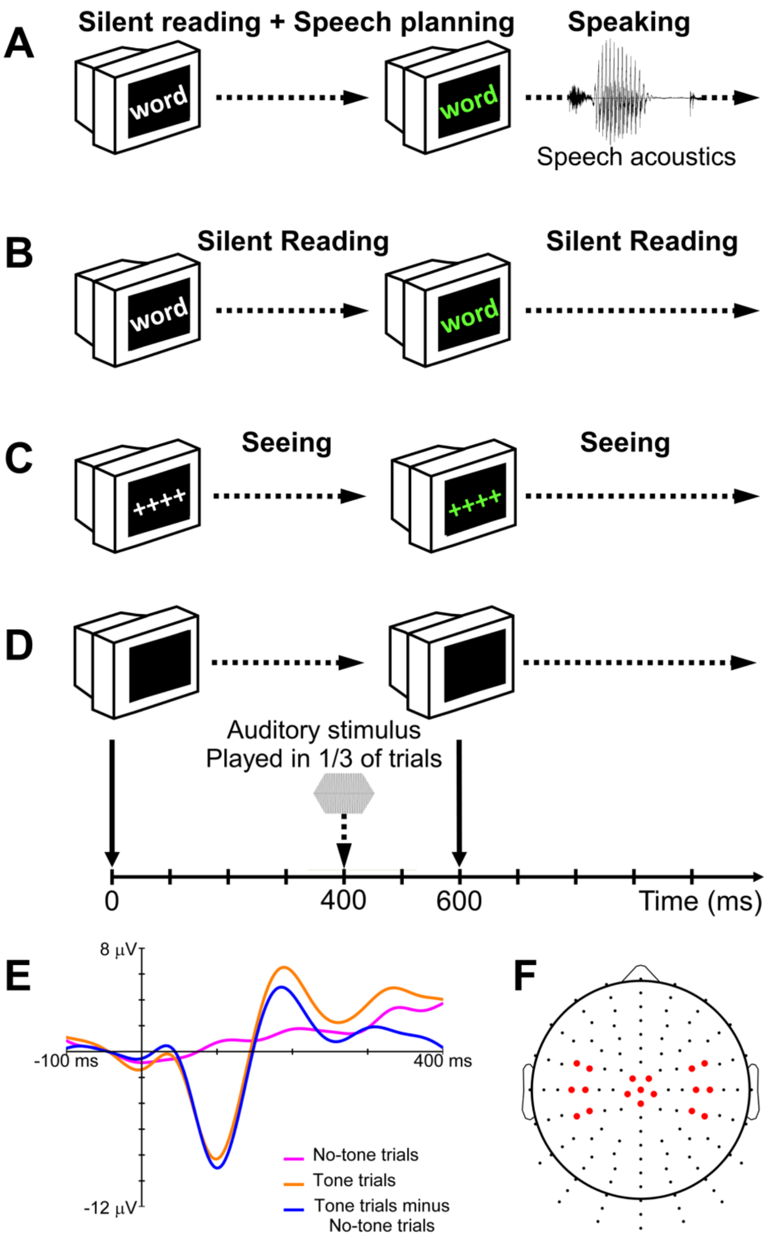

Fig. 1. Experimental procedures for *speaking* (A), *reading* (B), and *seeing* (C) conditions. In 1/3 of the trials (tone trials) for each condition, an auditory stimulus (1 kHz, 50 ms duration, 75 dB SPL) was presented (D). In the remaining trials (no-tone trials), no auditory stimulus was presented. To best estimate the auditory cortex response to the stimulus, each subject's averaged signal for no-tone trials was subtracted from that for tone trials (E). Subtracted signals were obtained for each of three ROIs, with each ROI's signal representing the average of 6 electrodes.

Statistical analyses

Statistical analyses were performed using the IBM SPSS Statistics 19 software package (IBM, Armonk, NY). N1 and P2 amplitude and latency data were subjected to analysis of variance (ANOVA) for repeated measures with Condition (speaking, reading, seeing) and ROI (left, central, right) as within-subjects variables and Group (stuttering, nonstuttering) as a between-subjects variable. Given the differences in recording procedure between these conditions and the rest condition (i.e., absence of visual stimuli), N1 and P2 amplitude and latency for the rest condition were analyzed separately. For the rest condition, we used repeated measures ANOVA with ROI as the within-subjects variable and Group as a between-subjects variable. For all within-subjects tests, degrees of freedom were adjusted using the Huynh-Feldt correction to account for potential violations of the sphericity assumption (Max & Onghena, 1999). Statistically significant interactions involving the Group variable were followed up with post-hoc analyses conducted by means of *t*-tests with Bonferroni corrections for multiple comparisons. Lastly, for the stuttering group only, we used Pearson correlation coefficients to investigate whether there was a relationship between participants' stuttering frequency (defined as the average percent stuttered syllables across the speaking and reading tasks of the SSI–4, see Table 1) and any of the dependent variables or the amount of N1 amplitude change in the speaking condition vs. the reading condition.

Results

Grand average LLAEPs for the different ROIs (Left ROI, Central ROI, Right ROI), conditions (*speaking*, *reading*, *seeing*), and groups (nonstuttering, stuttering) are shown in Fig. 2. Summarized group data (means and standard errors of the mean) for N1 and P2 amplitude and the distribution of individual subjects (boxplots with individual subject markers added) for the amount of N1 and P2 modulation in the *speaking* condition as compared with the *reading* condition are represented in Fig. 3.

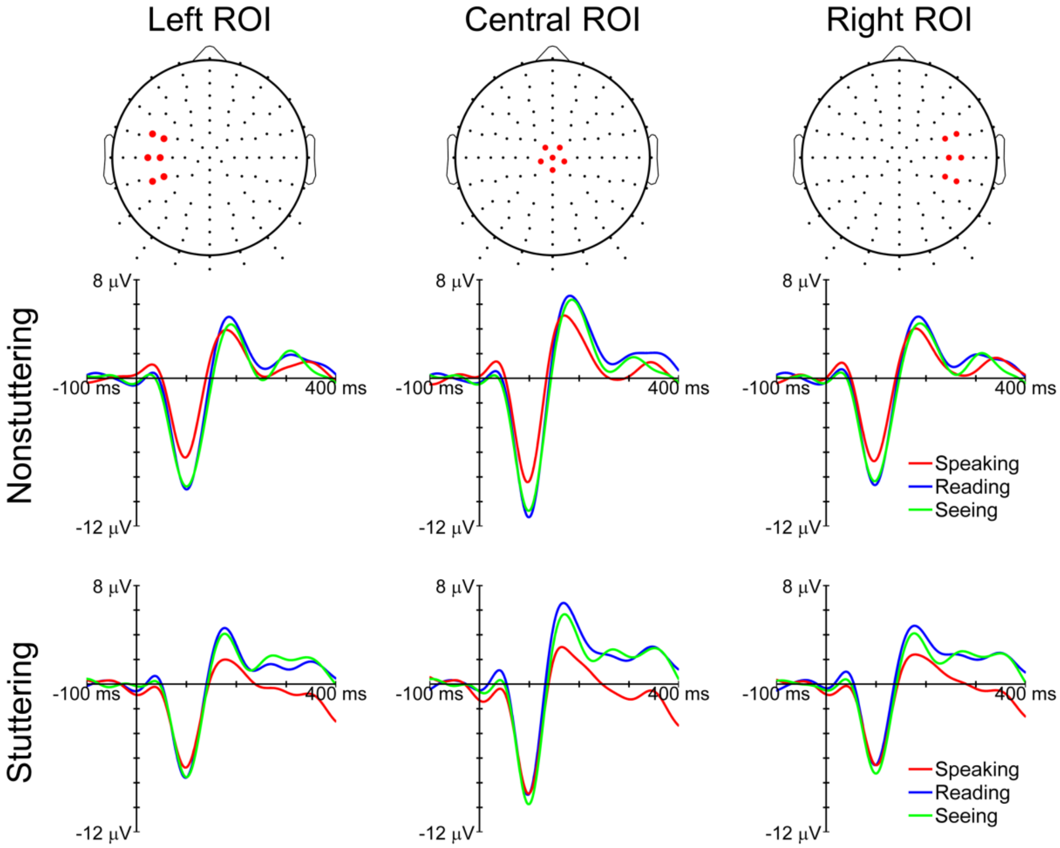

Fig. 2. Grand average LLAEPs from the Left ROI (left column), Central ROI (middle column), and Right ROI (right column) for individuals who stutter (bottom row) and individuals who do not stutter (middle row) during speech movement planning (red), silent reading (blue), and seeing nonlinguistic symbols (green). Individual electrode channels contributing to each ROI are colored red in the scalp representations (top row). Only the nonstuttering group showed a distinct attenuation of N1 amplitude prior to speaking as compared with both silent reading and seeing.

N1 amplitude

N1 amplitude data showed statistically significant main effects of ROI, *F(2, 44) = 38.081*, *p < .001*, and Condition, *F(1.983, 41.647) = 7.210*, *p = .002*, but not Group (*p = .452*). The ROI effect was associated with descriptively larger N1 amplitudes in the Central ROI (*M = 10.097 µV*, *SD = 3.442*) than in the Left ROI (*M = 8.125 µV*, *SD = 2.542*) and the Right ROI (*M = 7.753 µV*, *SD = 2.467*). The Condition effect was associated with overall smaller N1 amplitudes in the speaking condition (*M = 7.729 µV, SD = 2.687*) as compared with the reading condition (*M = 9.117 µV, SD = 3.209*) and the seeing condition (*M = 9.128 µV, SD = 3.126*). However, these main effects were modified by statistically significant interactions of Condition × ROI, *F(3.585, 78.873) = 3.955, p = .007,* and Condition × Group, *F(1.893, 41.647) = 3.738, p = .034*, in the absence of an ROI × Group interaction (*p = .884*) or an ROI × Condition × Group interaction (*p = .120*). The Condition × ROI interaction occurred because, averaged across groups, N1 amplitude differences between the speaking condition and the reading and seeing conditions were more extensive for the Left ROI and the Central ROI than for the Right ROI.

Most important for the hypothesis under investigation is the Condition × Group interaction. As can be seen in Fig. 3A, the N1 amplitude of the nonstuttering group was smaller in the *speaking* condition (*M = 7.538 µV, SD = 1.962*) than in both the *reading* condition (*M = 10.072 µV, SD = 2.533*), *t (11) = -7.802, p < .001,* and the *seeing* condition (*M = 9.659 µV, SD = 2.745*), *t = (11) =-3.695, p = .003*. For the stuttering group, on the other hand, the N1 amplitude in the *speaking* condition (*M = 7.921 µV, SD = 3.342*) was not statistically significantly different from the N1 amplitude in either the *reading* condition (*M = 8.163 µV, SD = 3.624, p = .680*) or the *seeing* condition (*M = 8.598 µV, SD = 3.505, p = .414*). When expressing the extent of auditory response modulation prior to *speaking* as a reduction in N1 amplitude relative to the *reading* condition, there was a complete separation of the interquartile-ranges of the stuttering and nonstuttering subject samples (Fig. 3C). In fact, as can be seen in the same figure, whereas all 12 nonstuttering speakers showed a reduction of the N1amplitude prior to speaking, 4 of the 12 stuttering speakers even showed an *increase* of the N1 amplitude prior to speaking. Topographic maps of both groups' N1 response in the three primary conditions, and particularly the comparison *reading* minus *speaking*, also clearly illustrate this difference between stuttering and nonstuttering speakers in N1 amplitude modulation prior to speech onset (Fig. 4).

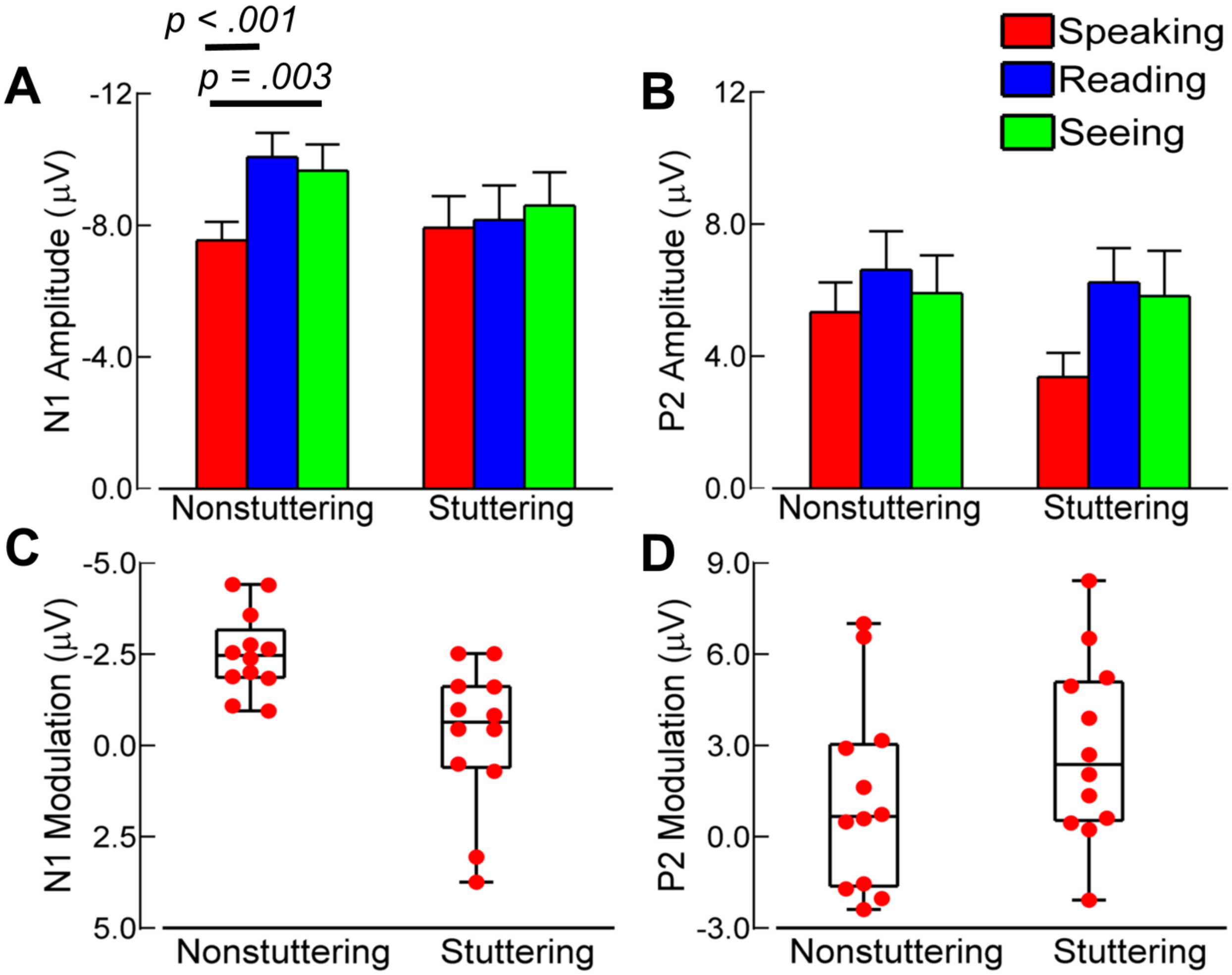

Fig. 3. Stuttering and nonstuttering group means and standard errors (error bars) of the mean for N1 amplitude (A) and P2 amplitude (B) during speech movement planning, silent reading, and seeing nonlinguistic symbols. Subject distributions in terms of the amount of response modulation (amplitude decrease in the speaking condition vs. the reading condition) are shown as box plots with overlaid individual subject data (red markers) for the N1 (C) and P2 (D) components of the auditory evoked potential. There was a statistically significant Group × Condition interaction for N1 amplitude: the nonstuttering group showed a reduction in N1 amplitude prior to speaking, whereas the stuttering group's N1 amplitude remained unchanged across conditions. The Group × Condition interaction for P2 was not statistically significant. Data in all panels have been averaged across Left, Central, and Right ROIs.

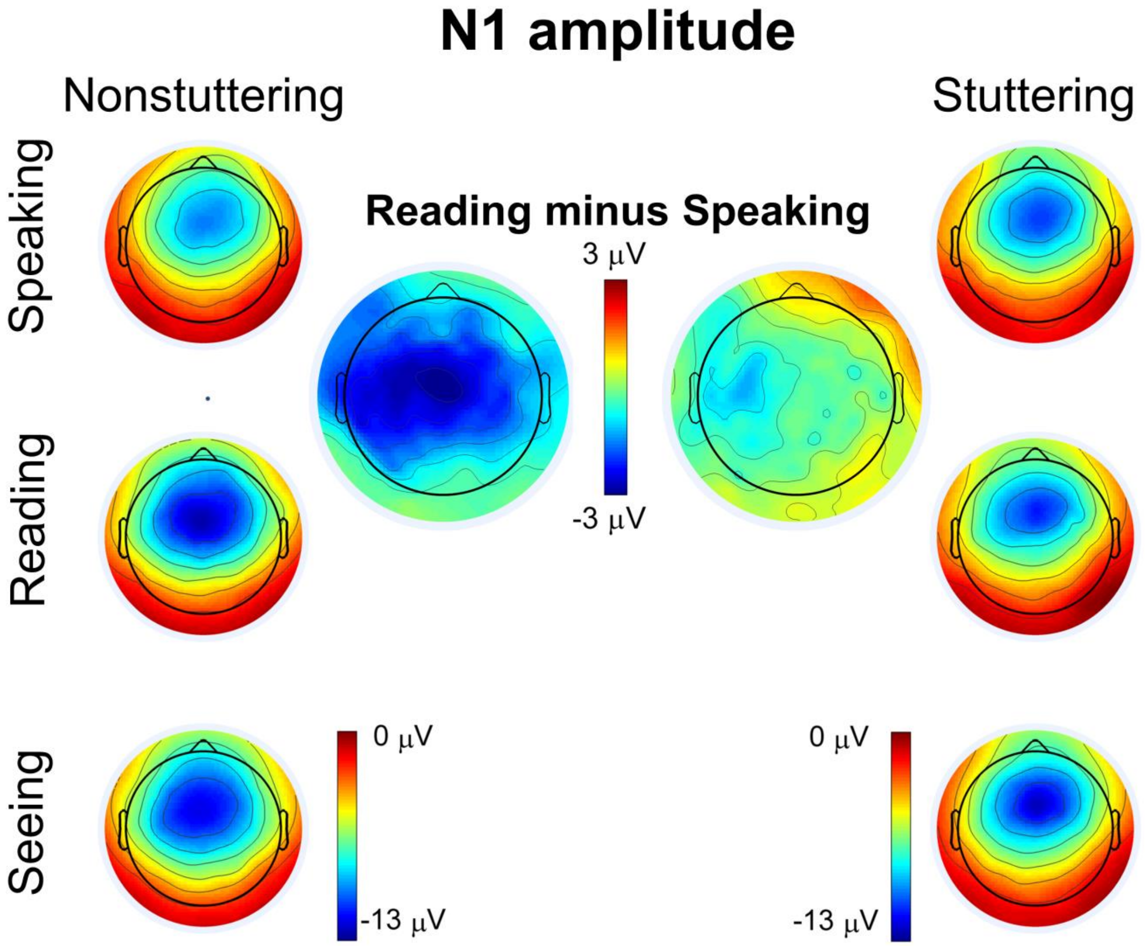

Fig. 4. Nonstuttering (left column) and stuttering (right column) subjects' topographic maps for N1 amplitude (maps created at the latency of each individual subject's N1 peak amplitude) in the *speaking*, *reading*, and *seeing* conditions. Additional topographic maps shown in the center were created by subtracting the N1 response in the speaking condition from that in the reading condition. The subtraction graphs clearly show a strong pre-speech modulation of the auditory response in the nonstuttering group but not in the stuttering group.

N1 latency

N1 latency data showed a statistically significant main effect of ROI, $F(2, 44) = 3.495$, $p = .039$, with overall longer latencies for the Left ROI ($M = 100.377$ ms, $SD = 8.308$) than for the Central ROI ($M = 98.450$ ms, $SD = 6.565$) and the Right ROI ($M = 99.047$ ms, $SD = 8.118$). The main effects of Condition and Group were not statistically significant ($p = .745$ and $p = .747$, respectively). There was a statistically significant ROI × Group interaction, *F(2, 44) = 3.722, p = .032*. For the nonstuttering group, the latency in the Left ROI (*M = 99.902 ms, SD = 7.992*) was longer than that in the Right ROI (*M = 97.514 ms, SD = 7.692*), *t(11) = 3.239, p = .007*. For the stuttering group, the latency in the Left ROI (*M = 100.852 ms, SD = 8.941*) was not statistically different (*p = .804*) from that in the Right ROI (*M = 100.580 ms, SD = 8.572*). In other words, N1 latencies of the stuttering group were more symmetric across the left and right hemispheres than those of the nonstuttering group. Of particular importance in the context of auditory modulation prior to speaking, however, there was no statistically significant Condition × Group interaction (*p = .203*). The Condition × ROI (*p = .151*) and Condition × Group × ROI (*p = .630*) interactions were also not statistically significant for the N1 latency.

P2 amplitude

Similar to the N1 amplitude data, P2 amplitude also showed statistically significant main effects of Condition, *F(1.756, 38.634) = 6.851, p = .004*, and ROI, *F(1.877, 41.295) = 29.755, p < .001*, but not Group (*p = .558*). Both effects followed the same trend as those described for N1: across conditions, P2 amplitudes were smaller in the speaking condition (*M = 4.353 µV, SD = 2.944*) than in both the reading condition (*M = 6.423 µV, SD = 3.750*) and the seeing condition (*M = 5.863 µV, SD = 4.269*); across ROIs, P2 amplitudes were larger in the Central ROI (*M = 6.720 µV, SD = 4.047*) than in the Left ROI (*M = 4.881 µV, SD = 2.958*) and the Right ROI (*M = 5.037 µV, SD = 3.042*). These main effects were modified by a statistically significant Condition × ROI interaction, *F(4,88) = 3.086, p = .020*: here, averaged across groups, P2 amplitude differences between the speaking condition and the reading and seeing conditions were more extensive for the Central ROI than for the Left ROI and the Right ROI. Also consistent with the N1 data, the ROI × Group and Condition × ROI × Group interactions for P2 amplitude data were not statistically significant (*p = .610* and *p = .278*, respectively). In this case of P2 amplitudes, however, there was no statistically significant Condition × Group interaction (*p = .234*, data included in Fig. 3B and 3D).

P2 latency

Measurements of P2 latency showed a statistically significant main effect of Condition, *F(1.636, 35.993) = 5.735, p = .010*, with shorter latencies in the speaking condition (*M = 181.214 ms, SD = 21.088*) than in the reading condition (*M = 188.799 ms, SD = 27.940*) and the seeing condition (*M = 201.867 ms, SD = 36.658*). In addition, there also was a statistically significant main effect of ROI, *F(1.258, 27.671) = 6.128, p = .014*, with shorter latencies in the Central ROI (*M = 186.167 ms, SD = 23.719*) than in the Left ROI (*M = 194.811 ms, SD = 27.816*) and Right ROI (*M = 190.903 ms, SD = 22.593*). None of the interactions Condition × Group, ROI × Group, Condition × ROI, or Condition × ROI × Group interactions were statistically significant (*p* > *.089* in all cases).

Rest condition

Fig. 5 illustrates the stuttering and nonstuttering groups' LLAEPs from the eyes-closed *rest* condition as recorded in all three ROIs. The main effect of ROI was statistically significant for both N1 amplitude, *F(2, 44) = 32.083, p < .001*, and P2 amplitude, *F(1.866, 41.052) = 26.654, p < .001*. In both cases, amplitudes were larger for the Central ROI than for the Left and Right ROIs. There were, however, no statistically significant effects on N1 or P2 amplitude for Group (*p* > *.390*) or ROI × Group (*p* > *.317*). For N1and P2 latencies, none of the main effects or interactions were statistically significant (*p* > *.078* in all cases).

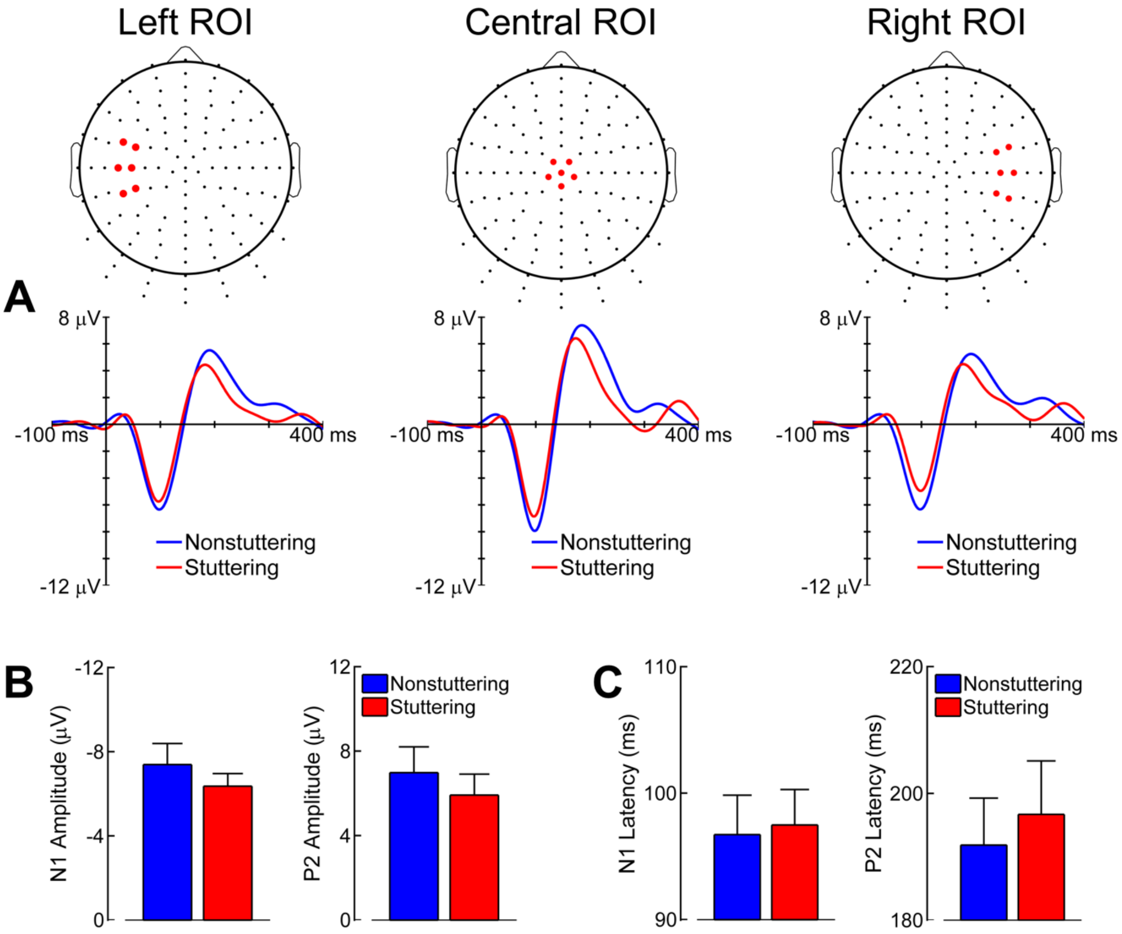

Fig. 5. Eyes-closed rest condition LLAEPs from three ROIs for stuttering (red) and nonstuttering (blue) subjects (A). There was no statistically significant Group main effect for N1 or P2 amplitudes (B) or for N1 or P2 latencies (C). Error bars indicate standard errors.

Correlational analyses

Pearson correlation coefficients revealed that there were no statistically significant relationships ($p > .159$ in all cases) between stuttering frequency and any of the dependent variables or between stuttering frequency and the amount of change in N1 amplitude in the *speaking* condition vs. the *reading* condition.

Discussion

Several research groups have previously suggested that stuttering may be associated with deficits in fundamental processes of sensorimotor integration underlying speech motor learning and speech motor control (Beal et al., 2010; Brown et al., 2005; Cai et al., 2012; Chang et al., 2011; Daliri et al., 2013; Hickok et al., 2011; Kell et al., 2009; Liotti et al., 2010; Max, 2004; Watkins et al., 2008). However, an answer to the question of which specific sensorimotor mechanisms are affected in this disorder of speech fluency has remained elusive.

Recent empirical and theoretical work on the neural control of movement has demonstrated that one important aspect of sensorimotor integration involves the prediction of a movement's sensory consequences (e.g., auditory feedback would be predicted during speech movement planning). Incorporating this information into our own theoretical framework for stuttering (Max, 2004), we have proposed that stuttering may be associated with difficulties in using such predictions to appropriately prime task-relevant sensory systems in advance of their subsequent role in online feedback monitoring and control. Here, we tested one type of sensory prediction directly by using probe tones and EEG recordings (focusing on the N1 and P2 components of the auditory evoked potential) to investigate adjustments in auditory processing during speech movement planning in stuttering vs. nonstuttering adults. In particular, we compared, for these two groups of speakers, the modulation of auditory evoked potentials in response to probe tones presented prior to speaking (i.e., during speech movement planning) and in control conditions without preparation for motor activity (i.e., silent reading and seeing nonlinguistic symbols). We hypothesized that, if stuttering is associated with deficient modulation of auditory cortical regions prior to speech onset, the stuttering group would fail to show the typical pre-speech N1 amplitude attenuation that we have previously documented for normally fluent speakers (Max et al., 2008).

The primary finding of this study is fully in agreement with the proposed hypothesis: the nonstuttering group showed the expected reduction in auditory N1 amplitude during speech movement planning relative to their own auditory N1 amplitude in the control conditions, but the age-, handedness-, and sex-matched stuttering group failed to show this modulation of N1 amplitude prior to speech initiation as compared with the control conditions. At the individual subject level, the majority of stuttering subjects showed an amount of auditory modulation that was smaller than that observed for any of the nonstuttering subjects, and there was no overlap between the inter-quartile ranges of both groups for this measure of modulation. Topographic maps based on data reflecting the difference in auditory response between the *reading* and *speaking* conditions also clearly confirmed the substantial between-group difference in auditory modulation prior to speech onset.

A second finding directly related to the phenomenon of auditory modulation indicates that both groups showed smaller P2 amplitudes during speech movement planning as compared with the no-speaking control conditions. Descriptively (based on group mean amplitudes), this P2 amplitude reduction prior to speech onset even appeared to be greater in the stuttering group than in the nonstuttering group. However, the condition by group interaction did not reach or approach statistical significance, and the corresponding box plots in Fig. 3 (panel D) make it clear that the range of P2 amplitude reductions across individual subjects was highly similar for the stuttering and nonstuttering groups. It should be noted also that the interpretation of descriptive data regarding such later components of the auditory evoked potential (P2 and later) is further complicated by two additional observations. First, a comparison of the box plots for the control subjects' amplitude modulation data reveals that the across-subjects range for P2 modulation (Fig. 3D) is more than twice as large as that for N1 modulation (Fig. 3C). Hence, the individual subject data for P2 are associated with substantially more variability. Second, the group grand average auditory evoked potentials (Fig. 2) show that the stuttering subjects' *speaking* condition, as compared with their non-speech control conditions, is not associated with a specific reduction in the P2 component but with a generally less positive/more negative potential from approximately 150 ms after stimulus onset until the end of the extracted epochs (400 ms after stimulus onset). Based on the data available to date, however, this generally downward-shifted potential appears due to random inter-subject variability, rather than systematic between-group differences, given that (a) our initial study on pre-speech auditory modulation found a similar effect in nonstuttering speakers (Max et al., 2008), and (b) a separate new study with *speaking*, *reading*, and *listening* conditions found this effect in neither stuttering nor nonstuttering subjects (Daliri & Max, 2014).

As a third finding—informative about stuttering individuals' auditory processing in general but not directly related to pre-speech sensory modulation—our analyses revealed that the nonstuttering group showed a small but statistically significant N1 latency difference between the Left and Right ROIs (2.4 ms longer on the left) across the *speaking*, *reading*, and *seeing* conditions, whereas such a lateralization difference was absent in the stuttering group. In the eyes-closed *rest* condition, the two groups did not differ in any measures of N1 or P2 amplitude or latency.

In summary, the overall conclusion that can be drawn from the present study is that stuttering speakers failed to show the typical pre-speech attenuation of auditory N1 responses to probe tones as documented for nonstuttering speakers both here and in Max et al. (2008). This conclusion based on data in the auditory domain is compatible with McClean's (1996) previous finding, in the somatosensory domain, that stuttering speakers show limited attenuation of lip muscle reflexes evoked prior to the onset of speech. Individual subject data from our own study show that the pre-speech auditory modulation of 10 out of 12 stuttering subjects (83%) was more limited than that observed at the 75th percentile of the nonstuttering group. In McClean's (1996) study, 10 out of 14 stuttering subjects (71%) failed to show significant modulation of the lip reflex prior to speech initiation. Hence, combined, these studies strongly suggest that stuttering is associated with deficiencies in predictively modulating relevant sensory systems for their crucial contributions to speech sensorimotor control. In the context of contemporary theoretical perspectives on the neural control of voluntary movements, those contributions of the auditory and somatosensory systems are likely to involve not only the online monitoring of actual feedback but also the fine-tuning of planned motor commands based on the movements' *predicted* sensory consequences (see Max, 2004). Disruptions in either of these two important functions (correcting movement planning and implementing error monitoring) may directly lead to the repetitive articulatory movements or sustained vocal tract postures that are identified as the primary characteristic of stuttering (Max, 2004).

This suggestion does raise the question why our stuttering subjects were able to produce the target utterances for the present experiment without perceptually noticeable dysfluencies despite the demonstration of atypical pre-speech auditory modulation in this situation. Although any answer to this question is merely speculative at this time, the apparent discrepancy may be related specifically to the fact that all target utterances consisted of monosyllabic CVC words that contained no consonant clusters and that were produced in isolation. It is well documented that both shorter words and words produced in isolation are less likely to result in stuttering than longer words and connected speech (Brown, 1938; Brown & Moren, 1942; Wingate, 1967; Wingate, 2002). Such short and isolated words may be only minimally taxing in terms of the CNS' reliance on auditory feedback. That is, if the production of such words depends almost entirely on feedforward, rather than feedback, control mechanisms (see Kim & Max, 2014), then the production of this type of utterances may not be directly impacted by limitations in modulating the auditory system prior to speech onset. We acknowledge, however, that the functional relevance of this pre-speech auditory modulation phenomenon itself remains to be determined in future studies. Specifically, a primary question that remains to be addressed is whether the phenomenon reflects increases vs. decreases in the weighting of feedback control mechanisms during speech production. If the decreased N1 amplitude actually corresponds to an overall attenuated auditory responsiveness (as opposed to a “fine-tuning” or “priming” of this sensory system for *enhanced* auditory monitoring), then it cannot be excluded that the stuttering speakers’ lack of modulation may reflect a compensatory mechanism that seeks to benefit from the availability of sensory feedback for online movement corrections (Max, 2004).

Similarly, explanations of the specific neural mechanisms leading to the observed absence of pre-speech auditory modulation in individuals who stutter will require further experimental and theoretical work, but at least two hypotheses can be formulated to guide this line of inquiry. First, given that (a) modulation of the earlier N1 component but not the later P2 component of the auditory evoked potential differentiated between stuttering and nonstuttering subjects and (b) atypical modulation in stuttering subjects has also been observed outside of the auditory domain (McClean, 1996), stuttering may be associated with a delay in sending the modulation signals themselves. That is, slower motor preparation in stuttering individuals may cause the modulating neural signals to reach sensory cortex at a later point in time, closer to speech onset. In our study, each auditory stimulus was delivered 400 ms after initial presentation of the target word. Thus, on average, the evoked N1 and P2 components occurred approximately 500 and 600 ms, respectively, into the movement planning phase. The first explanatory hypothesis suggests that stuttering speakers’ auditory modulation signals (associated with motor preparation) may have reached auditory cortex more than 500 ms after initial presentation of the target word. In contrast, a second hypothesis is that the deficient auditory modulation in stuttering subjects stems not from a functionally delayed motor system but from structural white matter abnormalities between speech (pre-) motor regions and the neural generators of the N1 component. We propose that these alternative (but certainly not exhaustive) hypotheses can be tested directly by varying the time of delivery of the auditory stimuli (e.g., 400, 500, 600 ms after presentation of the word), and examining whether stuttering subjects do show N1 modulation when the auditory stimuli are presented later into the movement planning phase. The first hypothesis predicts that stuttering speakers may show typical N1 modulation at time points closer to movement onset whereas the second hypothesis predicts that N1 modulation will remain limited in stuttering speakers regardless of the time of stimulus delivery.

It is worth noting here that the first of these hypotheses may be more consistent with the results of Beal et al. (2010, 2011) and Liotti et al. (2010) who found no differences between stuttering and nonstuttering speakers in the modulation of auditory cortex activity in response to one's *own auditory feedback after speech onset* as compared with played-back, pre-recorded speech (a paradigm developed by Houde, Nagarajan, Sekihara, & Merzenich, 2002). Given that the experimental procedures differ substantially, however, we speculate that the latter paradigm is likely to address neural processes that are distinct from those addressed by our own paradigm. In the future, more comprehensive studies should be designed to explore the relationship and potential overlap, if any, between the sensorimotor interactions responsible for each of these two forms of auditory modulation.

In conclusion, this study investigated pre-speech auditory modulation in stuttering and nonstuttering adults. Auditory modulation was examined by comparing N1 and P2 amplitude and latency measures in long-latency auditory evoked potentials elicited with pure tones during the delay phase of a delayed-response speaking condition as compared with no-speaking control conditions. During this speech movement planning phase, only the nonstuttering subjects showed a statistically significant reduction in the amplitude of the N1 component. For the stuttering subjects, N1 amplitude did not different between the speech and control conditions. Consequently, these electrophysiological data suggest that stuttering is associated with deficiencies in modulating the cortical auditory system, and potentially relevant sensory systems in general, prior to speech initiation.

Acknowledgements

This research was supported, in part, by grants R01DC007603 and P30DC004661 from the National Institute on Deafness and Other Communication Disorders. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute on Deafness and Other Communication Disorders or the National Institutes of Health. The authors thank Santosh Maruthy, Selena Galaviz, and Derek G. Maffett for assistance with subject recruitment and data collection.

References

Baess, P., Horvath, J., Jacobsen, T., & Schroeger, E. (2011). Selective suppression of self-initiated sounds in an auditory stream: An ERP study. *Psychophysiology, 48*(9), 1276–1283. doi:10.1111/j.1469-8986.2011.01196.x

Baess, P., Jacobsen, T., & Schroger, E. (2008). Suppression of the auditory N1 event-related potential component with unpredictable self-initiated tones: Evidence for internal forward models with dynamic stimulation. *International Journal of Psychophysiology, 70*(2), 137–143.

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. doi:10.1016/j.neuroimage.2010.04.277

Beal, D. S., Quraan, M. A., Cheyne, D. O., Taylor, M. J., Gracco, V. L., & De Nil, L. F. (2011). Speech-induced suppression of evoked auditory fields in children who stutter. *Neuroimage, 54*(4), 2994–3003. doi: 10.1016/j.neuroimage.2010.11.026

Braun, A., Varga, M., Stager, S., Schulz, G., Selbie, S., Maisog, J., . . . Ludlow, C. (1997). Altered patterns of cerebral activity during speech and language production in developmental stuttering - an H-2 O-15 positron emission tomography study. *Brain, 120,* 761–784. doi:10.1093/brain/120.5.761

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

Brown, S. F. (1938). A further study of stuttering in relation to various speech sounds. *Quarterly Journal of Speech, 24*(3), 390–397.

Brown, S. F., & Moren, A. (1942). A study of stuttering in relation to word length. *Journal of Speech Disorders, 7,* 153–159.

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 motor-auditory transformation. *Plos One, 7*(7), e41830. doi:10.1371/journal.pone.0041830

Chang, S., 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. doi:10.1016/j.neuroimage.2007.09.067

Chang, S., Horwitz, B., Ostuni, J., Reynolds, R., & Ludlow, C. L. (2011). Evidence of left inferior frontal-premotor structural and functional connectivity deficits in adults who stutter. *Cerebral Cortex, 21*(11), 2507–2518. doi:10.1093/cercor/bhr028

Chang, S., Kenney, M. K., Loucks, T. M. J., & Ludlow, C. L. (2009). Brain activation abnormalities during speech and non-speech in stuttering speakers. *Neuroimage, 46*(1), 201–212. doi:10.1016/j.neuroimage.2009.01.066

Cornelisse, L. E., Gagné, J., & Seewald, R. C. (1991). Ear level recordings of the long-term average spectrum of speech. *Ear and Hearing, 12*(1), 47–54.

Cykowski, M. D., Fox, P. T., Ingham, R. J., Ingham, J. C., & Robin, D. A. (2010). A study of the reproducibility and etiology of diffusion anisotropy differences in developmental stuttering: A potential role for impaired myelination. *Neuroimage, 52*(4), 1495–1504. doi:10.1016/j.neuroimage.2010.05.011

Daliri, A., & Max, L. (2014). *Electrophysiological evidence for a general sensory prediction deficit in adults who stutter*. Manuscript submitted for publication.

Daliri, A., Prokopenko, R., & 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. doi:10.1044/1092-4388(2013/12-0134)

Daliri, A., Prokopenko, R., Flanagan, J. R., & Max, L. (2014). Control and prediction components of movement planning in stuttering vs. nonstuttering adults. *Journal of Speech, Language, and Hearing Research, 57*(6), 2131–2141. doi:10.1044/2014_JSLHR-S-13-0333

De Nil, L., Kroll, R., Kapur, S., & Houle, S. (2000). A positron emission tomography study of silent and oral single word reading in stuttering and nonstuttering adults. *Journal of Speech, Language, and Hearing Research, 43*(4), 1038–1053.

De Nil, L., Kroll, R., Lafaille, S., & Houle, S. (2003). A positron emission tomography study of short- and long-term treatment effects on functional brain activation in adults who stutter. *Journal of Fluency Disorders, 28*(4), 357–380. doi:10.1016/j.jfludis.2003.07.002

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.

Foundas, A. L., Corey, D. M., Angeles, V., Bollich, A. M., Crabtree-Hartman, E., & Heilman, K. M. (2003). Atypical cerebral laterality in adults with persistent developmental stuttering. *Neurology, 61*(10), 1378–1385.

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

Gilmore, R. L. (1994). American-electroencephalographic-society guidelines in electroencephalography, evoked-potentials, and polysomnography. *Journal of Clinical Neurophysiology, 11*(1), 1–142.

Godey, B., Schwartz, D., De Graaf, J., Chauvel, P., & Liegeois-Chauvel, C. (2001). Neuromagnetic source localization of auditory evoked fields and intracerebral evoked potentials: A comparison of data in the same patients. *Clinical Neurophysiology, 112*(10), 1850–1859.

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

Houde, J. F., Nagarajan, S. S., Sekihara, K., & Merzenich, M. M. (2002). Modulation of the auditory cortex during speech: An MEG study. *Journal of Cognitive Neuroscience, 14*(8), 1125–1138. doi:10.1162/089892902760807140

Ingham, R. J., Wang, Y., Ingham, J. C., Bothe, A. K., & Grafton, S. T. (2013). Regional brain activity change predicts responsiveness to treatment for stuttering in adults. *Brain and language, 127*(3), 510–519.

Kell, C. A., Neumann, K., von Kriegstein, K., Posenenske, C., von Gudenberg, A. W., Euler, H., & Giraud, A. L. (2009). How the brain repairs stuttering. *Brain, 132*(10), 2747–2760. doi:10.1093/brain/awp185; 10.1093/brain/awp185

Kim, K. S., & Max, L. (2014). Estimating feedforward versus feedback control of speech production through kinematic analyses of unperturbed articulatory movements. *Frontiers in Human Neuroscience, 8*, 911.

Liotti, M., Ingham, J. C., Takai, O., Paskos, D. K., Perez, R., & Ingham, R. J. (2010). Spatiotemporal dynamics of speech sound perception in chronic developmental stuttering. *Brain and Language, 115*(2), 141–147. doi:10.1016/j.bandl.2010.07.007

Martikainen, M. H., Kaneko, K., & Hari, R. (2005). Suppressed responses to self-triggered sounds in the human auditory cortex. *Cerebral Cortex, 15*(3), 299–302. doi:10.1093/cercor/bhh131

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

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.

Max, L., Daniels, J., Curet, K., & Cronin, K. (2008). Modulation of auditory and somatosensory processing during the planning of speech movements. Paper presented at the *Proceedings of the 8th International Seminar on Speech Production*, Strasbourg, France (pp. 41–44).

McClean, M. D. (1996). Lip-muscle reflexes during speech movement preparation in stutterers. *Journal of Fluency Disorders, 21*(1), 49–60.

Näätänen, R., & Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: A review and an analysis of the component structure. *Psychophysiology, 24*(4), 375–425.

Neumann, K., Euler, H., von Gudenberg, A., Giraud, A., Lanfermann, H., Gall, V., & Preibisch, C. (2003). The nature and treatment of stuttering as revealed by fMRI - A within- and between-group comparison. *Journal of Fluency Disorders, 28*(4), 381–410. doi:10.1016/j.jfludis.2003.07.003

Oostenveld, R., & Praamstra, P. (2001). The five percent electrode system for high-resolution EEG and ERP measurements. *Clinical Neurophysiology, 112*(4), 713–719.

Papanicolaou, A. C., Baumann, S., Rogers, R. L., Saydjari, C., Amparo, E. G., & Eisenberg, H. M. (1990). Localization of auditory response sources using magnetoencephalography and magnetic resonance imaging. *Archives of Neurology, 47*(1), 33–37.

Riley G. (2008). *Stuttering Severity Instrument, Fourth Edition (SSI–4).* San Antonio, TX: Pearson.

Ross, B., & Tremblay, K. (2009). Stimulus experience modifies auditory neuromagnetic responses in young and older listeners. *Hearing Research, 248*(1), 48–59.

Sommer, M., Koch, M., Paulus, W., Weiller, C., & Buchel, C. (2002). Disconnection of speech-relevant brain areas in persistent developmental stuttering. *Lancet, 360*(9330), 380–383. doi:10.1016/S0140-6736(02)09610-1

Watkins, K. E., Smith, S. M., Davis, S., & Howell, P. (2008). Structural and functional abnormalities of the motor system in developmental stuttering. *Brain, 131,* 50–59. doi:10.1093/brain/awm241

Wingate, M. E. (1967).Stuttering and word length. *Journal of Speech and Hearing Research, 10,* 146–152.

Wingate, M. E. (2002). *Foundations of stuttering.* San Diego, CA: Academic Press.

Wymbs, N. F., Ingham, R. J., Ingham, J. C., Paolini, K. E., & Grafton, S. T. (2013). Individual differences in neural regions functionally related to real and imagined stuttering. *Brain and Language, 124*(2), 153–164.

Zimmermann, G. (1980). Stuttering—a disorder of movement. *Journal of Speech and Hearing Research*, *23*(1), 122–136.

Zouridakis, G., Simos, P. G., & Papanicolaou, A. C. (1998). Multiple bilaterally asymmetric cortical sources account for the auditory N1m component. *Brain Topography*, *10*(3), 183–189.

CHAPTER THREE

Electrophysiological evidence for a general auditory prediction deficit in adults who stutter

Ayoub Daliri and Ludo Max

Manuscript submitted to *Brain and Language*

Abstract

Previous work revealed that stuttering individuals do not show the typical auditory modulation that is observed during speech movement planning in nonstuttering individuals. We aimed to further elucidate this limitation by investigating whether stuttering speakers' atypical auditory modulation is observed only when sensory predictions are based on active movement planning or also when predictable auditory input is not a consequence of one's own actions. We recorded 10 stuttering and 10 nonstuttering adults' auditory evoked potentials in response to probe tones delivered while anticipating either speaking or hearing one's own speech played back and in a control condition without auditory input. N1 amplitude of nonstuttering speakers was reduced prior to both speaking and hearing versus the control condition. However, no N1 amplitude reduction occurred for stuttering speakers in either the speaking or hearing conditions. These findings suggest that stuttering speakers have general auditory prediction difficulties.

Introduction

Over the past decades, researchers have used both behavioral measures (such as acoustic parameters of speech or kinematic parameters of both speech and nonspeech movements) and neurological measures (such as brain structure and activation) to investigate differences between stuttering and nonstuttering individuals (Bloodstein & Bernstein-Ratner, 2008). Collectively, these studies provide evidence for the view that stuttering is associated with deficits in sensorimotor integration (for a review, see Max, 2004). To better understand which specific aspects of sensorimotor integration are deficient in stuttering individuals, we have previously examined the central nervous system's (CNS) prediction of the sensory consequences of planned limb and speech movements (Daliri & Max, 2014; Daliri, Prokopenko, Flanagan, & Max, 2014). This approach is based on theoretical models suggesting that, during movement planning, the CNS uses an efference copy (i.e., a copy of the prepared motor commands) and a forward internal model (i.e., a neural representation of the system's motor-to-sensory mapping) to predict the sensory consequences of planned movements (Shadmehr, Smith, & Krakauer, 2010; Sperry, 1950; Wolpert, Diedrichsen, & Flanagan, 2011; Wolpert, Miall, & Kawato, 1998). Such predictions serve to implement anticipatory adjustments that take account of biomechanical effects (Flanagan & Lolley, 2001) as well as to prime task-relevant sensory systems for their role in processing incoming afferent signals (Creutzfeldt, Ojemann, & Lettich, 1989; Crone et al., 2001; Mock, Foundas, & Golob, 2011; Wasaka, Hoshiyama, Nakata, Nishihira, & Kakigi, 2003).

First, we used a nonspeech task (fast arm movements to slide an object under the fingertip toward visual targets) to compare stuttering and nonstuttering individuals' ability to predict movement consequences related to the biomechanics of the upper limb. Although the fast reaching movements were performed with minimal reliance on sensory feedback, the planning of such movements in different directions is associated with anticipatory adjustments in vertical force applied to the object in order to compensate for biomechanically determined variation in movement acceleration (Daliri et al., 2014 Stuttering individuals were less accurate than nonstuttering individuals in reaching to the targets, but their compensatory adjustments in vertical force depending on movement direction were similar to those of the nonstuttering individuals. Thus, at least for this limb movement task that was performed in a ballistic manner without relying on feedback, stuttering individuals appeared to have no problems with predicting the biomechanical effects of the limb and appropriately implementing compensatory adjustments from the beginning of the movement.

In a second approach, we compared stuttering and nonstuttering speakers' modulation of the auditory system prior to the onset of speech movements (a task that does involve auditory feedback control). We recorded auditory evoked potentials (AEPs) in response to probe tones delivered during the movement planning phase prior to speech onset as compared with no-speaking control conditions (Daliri & Max, 2014; Max, Daniels, Curet, & Cronin, 2008). Results indicated that the group of nonstuttering speakers showed a statistically significant modulation of the N1 amplitude during speech planning whereas this modulation was lacking in the group of stuttering speakers. Thus, findings suggested that stuttering individuals may have difficulties with using sensory prediction mechanisms to prime the auditory system for its role in processing upcoming auditory feedback.

This intriguing finding raised the question whether pre-speech modulation of the auditory system in stuttering speakers is atypical due to problems specifically with generating or evaluating (by means of a forward internal model) an efference copy signal of the planned movements or, alternatively, a more general problem with using any available information to make sensory predictions. Indeed, in the absence of movement planning or execution, the CNS also builds associations among sensory events in a sequence, and then uses the association to predict future events—e.g., predicting the sound of thunder after seeing a bolt of lightning (for review, see Bubic, von Cramon, & Schubotz, 2010; Hughes, Desantis, & Waszak, 2013; O'Reilly, Mesulam, & Nobre, 2008; Schubotz, 2007; Waszak, Cardoso-Leite, & Hughes, 2012; Wolfensteller, Schubotz, & Von Cramon, 2007). We now address this specificity question by means of a study for which we again recorded AEPs (with a primary focus on the N1 component) of stuttering and nonstuttering adults in response to probe tones delivered while subjects were preparing to speak (involving prediction of the sensory consequences of planned speech movements). Here, however, we compared these pre-speech auditory responses with the auditory responses evoked by probe tones delivered while subjects were anticipating hearing a played-back version of their own pre-recorded speech (involving prediction of auditory input but in the absence of movement planning) or during a silent reading task (involving no prediction of auditory input). In each condition, participants saw a target word on a computer screen, and they either said the displayed word aloud after a short delay period (delayed-speaking condition), they heard a played-back version of their own pre-recorded production of the displayed word after the same delay period (delayed-hearing condition), or they silently read the displayed word (silent reading condition). We hypothesized that (a) if stuttering is associated with specific difficulties in generating sensory predictions based on an efference copy, then a difference between the two groups in auditory modulation would be limited to the delayed-speaking condition, whereas (b) if stuttering is associated with a more general sensory prediction deficit, then a difference between the two groups can be expected in both the delayed-speaking condition and the delayed-hearing condition.

Method

Participants

We obtained written consent from 10 right-handed stuttering adults (9 males and 1 female; M = 29.100 years, SD = 8.062) and 10 right-handed nonstuttering adults (9 males and 1 female; M = 29.400 years, SD = 9.383) prior to participation in the experiment. Nonstuttering participants were individually matched with the stuttering participants based on age (± 3 years) and sex. All participants were native speakers of American English, with no history of psychological, neurological, or communication disorders (other than stuttering in the stuttering group). All participants had normal binaural hearing thresholds (≤ 20 dB HL for octave frequencies 250–8000 Hz). To determine each stuttering participant's severity, we used the Stuttering Severity Instrument, Fourth Edition (SSI–4; Riley, 2008). Table 1 presents individual participant information for the stuttering group (age, sex, handedness, overall SSI–4 score, stuttering severity classification, and frequency of stuttering averaged across the SSI–4 speaking and reading tasks). Four stuttering participants had also participated in our previous study (Daliri & Max, 2014).

Table 1 Individual participant information for the stuttering group (SSI score = Stuttering Severity Instrument 4th ed. overall score; Stuttering frequency (%SS) = percent stuttered syllables averaged across the SSI delayed-speaking and reading tasks). The participants are rank ordered by stuttering frequency.

Stuttering participant	Age (years)	Sex	Handedness	SSI score	SSI severity	Stuttering frequency (%SS)
S1	25	male	right	12	very mild	1.85
S2	40	male	right	17	very mild	3.61
S3	31	male	right	22	mild	4.05

S4	34	male	right	17	very mild	5.50
S5	22	male	right	22	mild	5.52
S6	43	male	right	28	moderate	7.75
S7	20	male	right	21	mild	7.82
S8	19	female	right	23	mild	8.23
S9	28	male	right	32	severe	8.52
S10	31	male	right	28	moderate	10.38

Procedure and Instrumentation

Participants were seated approximately 1.5 m away from a 23-inch monitor, inside a sound-attenuated room. A microphone (WL185, Shure Incorporated, Niles, IL) was placed approximately 15 cm away from participant’s mouth to transduce speech signals. After amplification (microphone amplifier: DPS II, ART ProAudio, Niagara Falls, NY; headphone amplifier: S-phone, Samson Technologies Corp., Syosset, NY), speech signals were played-back to the participant in real-time through insert earphones (ER-3A, Etymotic Research Inc., Grove Village, IL). The insert earphones were also used to deliver pre-recorded speech signals and binaural probe tones (see below). Before each recording session, the amplifiers (microphone and headphone) were adjusted such that speech input with an intensity of 75 dB SPL at the microphone resulted in 73 dB SPL output in the earphones (Cornelisse, Gagné, & Seewald, 1991). We used a 2 cc coupler (Type 4946, Bruel & Kjaer Inc., Norcross, GA) connected to a sound level meter (Type 2250A Hand Held Analyzer with Type 4947 ½” Pressure Field Microphone, Bruel & Kjaer Inc., Norcross, GA) to calibrate the intensity of the auditory feedback in the earphones.

The experiment consisted of three conditions, hereafter referred to as *delayed-speaking*, *delayed-hearing*, and *silent reading* (Fig. 1A–C). Each condition consisted of 3 blocks of 90 trials. The first block was always a block from the *delayed-speaking* condition (such that the participant's speech output could be recorded, see below), but the order of the remaining blocks was randomized for each participant.

Each trial started with the presentation of a word in white character on a black background on the display (Fig. 1D). The word for each trial was randomly selected from a list of 45 monosyllabic consonant-vowel-consonant words in which the initial consonant was always a bilabial stop. All words were 3–5 letters long, and none of the words included consonant clusters. The color of the word changed to green after 600 ms (*go-signal*), and the word stayed on the display for an additional 500 ms. The inter-stimulus-interval from the end of a trial to the beginning of the next trial was randomly selected from a set of five possible intervals (1500, 2000, 2500, 3000, or 3500 ms).

In the *delayed-speaking* condition (Fig. 1A), participants said the word aloud after the color changed to green. In the *delayed-hearing* condition (Fig. 1B), the participant heard a self-produced version of the word played back with the same intensity and onset latency (relative to the beginning of the trial and thus also relative to the change in color) as produced in the first block of the *delayed-speaking* condition. In the *silent reading* condition (Fig. 1C), participants silently read the word without making any movements or hearing a word played back.

On 40% of the trials in each block (tone trials), probe tones were delivered 400 ms after presentation of the word in white characters (Fig. 1D). The probe tones were1 kHz pure tones with a duration of 40 ms and rise/fall times of 10 ms, presented at 75 dB SPL. No auditory stimuli were presented in the remaining trials (no-tone trials).

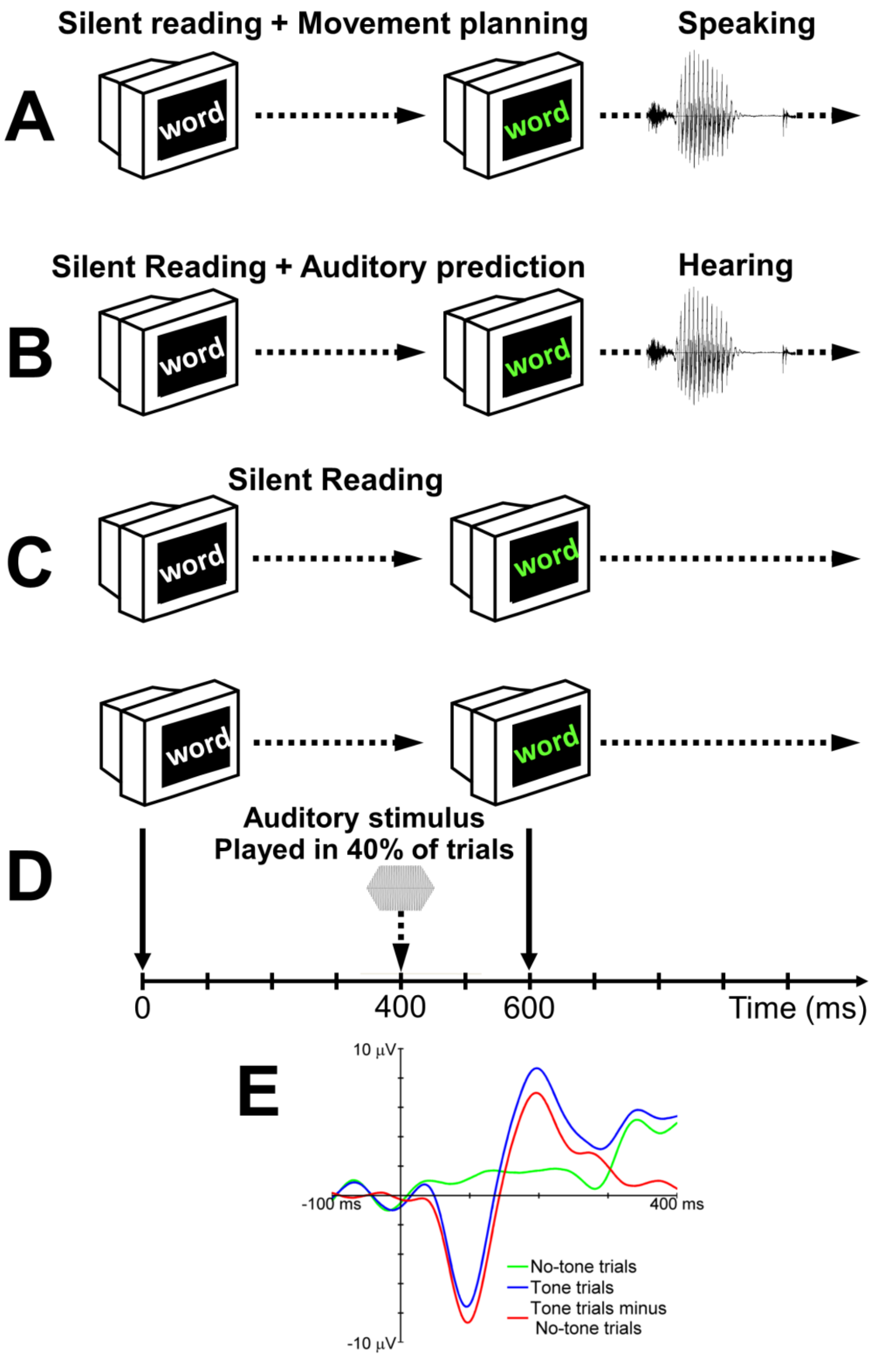

Fig. 1. Experimental procedures for delayed-speaking (A), delayed-hearing (B), and silent reading (C) conditions. A probe auditory stimulus (1 kHz; 40 ms duration; 75 dB SPL) was played on 40% of trials (tone trials) for each condition (D). No auditory stimulus was played in the remaining trials (no-tone trials). For each subject and in each condition, the average of no-tone trials was subtracted from the average of tone trials (E). This procedure was used to find the best estimate of the auditory cortex's response to the probe tones.

Electroencephalographic Recordings

Electroencephalographic (EEG), electromyographic (EMG), and electrooculographic (EOG) activity was recorded continuously throughout each block using an ActiveTwo BioSemi system (BioSemi, Amsterdam, The Netherlands). EEG signals were recorded from 128 standard sites on the scalp according to an extension of the international 10-10 electrode system (Gilmore, 1994; Oostenveld & Praamstra, 2001). Signals from two additional electrodes, placed over the left and right mastoids, were recorded for offline re-referencing of the EEG data (see Data Analysis). EMG signals were recorded from four electrodes placed on the skin overlying upper and lower lip muscles (i.e., right and left orbicularis oris superior and inferior which contributed to lip closing for the word-initial stop consonant). The EMG signals were used to verify an absence of muscle contraction prior to the *go*-signal (color of the displayed word changing to green) in the delayed-speaking condition. EMG signals were also used to estimate the onset time of speech movements, and this onset time was found to be not statistically significantly different between the stuttering and nonstuttering groups, $t(18) = -1.848, p = .081$. EOG signals were recorded using two electrodes placed below the lower eyelid and next to the outer canthus of the left eye. The EOG signals were used to reject trials with artifacts related to blinking and eye movements. Lastly, the acoustic signal from an additional microphone (SM58, Shure, Niles, IL) was recorded together with the EEG data. All signals (EEG, EMG, EOG, mastoid electrodes, and speech acoustics) were recorded with a sampling rate of 1024 Hz.

Data analysis

Data analysis was completed using the EEGLAB toolbox (Delorme & Makeig, 2004) and custom-written MATLAB scripts (The MathWorks, Inc., Natick, MA). Signals from the two mastoid electrodes were mathematically averaged to reconstruct a reference signal. EEG signals were re-referenced to this reconstructed reference signal. Using a low-pass filter (cut-off frequency 50 Hz), the EEG signals were filtered off-line. The filtered EEG signals were then segmented into epochs ranging from 100 ms before to 400 ms after the onset of the probe tone in tone trials, and the equivalent time window in no-tone trials. The average amplitude in the pre-stimulus window (i.e., first 100 ms of each epoch) was subtracted from the whole epoch to remove inter-trial baseline differences. The EEG data were then visually inspected to reject epochs contaminated with (a) muscle artifact, (b) blink or eye-movement artifacts, and (c) muscle activity before the go-signal. Next, the artifact free epochs for tone trials and no-tone trials were averaged separately. The averaged response for the tone trials reflected auditory (as a result of the probe tones) as well as non-auditory activity (e.g., related to motor, linguistic, cognitive, and visual processes), whereas the averaged response for the no-tone trials reflected only the non-auditory activity. To dissociate the auditory activity evoked by the probe tones from the non-auditory activity, each subject’s average signal obtained for the no-tone trials in a given condition was subtracted from the average signal obtained for tone trials in the same condition (Fig. 1E). This subtracted signal was then used for all AEP analyses as it provides the best estimate of the true auditory response (Baess, Horvath, Jacobsen, & Schroeger, 2011; Baess, Jacobsen, & Schroger, 2008; Daliri & Max, 2014; Martikainen, Kaneko, & Hari, 2005). The final AEPs for each channel were low-pass filtered with a cut-off frequency of 15 Hz.

Prior to statistical analysis, we calculated the average AEP across electrodes within each of three regions of interests (ROIs) over the left hemisphere (Left ROI: electrodes D11, D12, D19, D20, D27, D28), the central region (Central ROI: electrodes A1, A2, B1, C1, D1, D15), and the right hemisphere (Right ROI: electrodes B17, B18, B21, B22, B30, B31). The three ROIs and the included individual channels are illustrated in the top row of Fig. 2. From these

ROI-based AEPs, we extracted the latency and amplitude of the N1 and P2 components for each participant. N1 was defined as the largest negative peak between 70 and 130 ms following the onset of the auditory stimulus. P2 was defined as the largest positive peak between 150 and 250 ms following the onset of the auditory stimulus.

Statistical analyses

As dependent variables, N1 and P2 amplitude and latency were entered into the statistical analyses. We used analysis of variance (ANOVA) for repeated measures with Condition (delayed-speaking, delayed-hearing, and silent reading) and ROI (left, central, and right) as within-subjects variables and Group (stuttering and nonstuttering) as a between-subjects variable. To account for potential violations of the sphericity assumption, degrees of freedom were adjusted using the Huynh-Feldt correction (Max & Onghena, 1999). Statistically significant interactions relevant to the aims of the study were followed up with post-hoc analyses conducted by means of *t*-tests with Bonferroni corrections for multiple comparisons. All statistical analyses were conducted in the IBM SPSS Statistics 19 software package (IBM, Armonk, NY).

Results

Fig. 2 illustrates the grand average AEP waveforms for all ROIs (Left, Central, and Right), conditions (delayed-speaking, delayed-hearing, and silent reading), and groups (nonstuttering and stuttering). The following sections present results from the statistical analyses carried out for N1 amplitude and latency and P2 amplitude and latency, respectively.

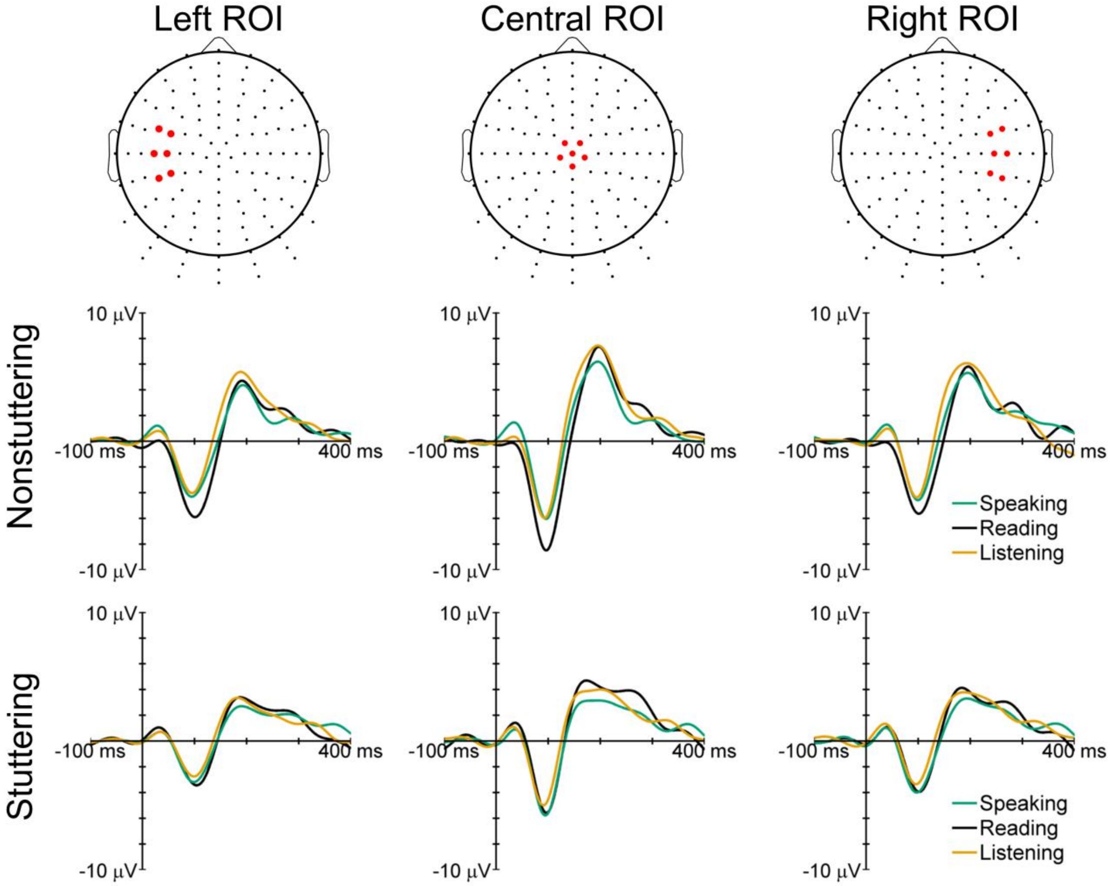

Fig. 2. The top row shows the three regions of interests with each consisting of six electrodes. AEPs from all six electrodes in each ROI were averaged. The middle row shows the grand average AEPs for the nonstuttering group from the Left ROI (left column), Central ROI (middle column), and Right ROI (right column) in the delayed-speaking (green), delayed-hearing (orange), and silent reading (black). The bottom row shows the AEPs for the stuttering group. AEPs of the nonstuttering group show a distinct attenuation of N1 amplitude in the delayed-speaking and delayed-hearing conditions.

N1 amplitude

Examining N1 amplitude, we found statistically significant effects of Condition, $F(1.794, 32.290) = 10.852, p < .001$, Group, $F(1, 18) = 5.140, p = .036$, and Condition × Group interaction, $F(1.794, 32.290) = 4.757, p = .018$. We also found a significant main effect of ROI, $F(1.977, 35.592) = 55.403, p < .001$, with larger N1 amplitude in the central ROI than left ROI, $t(19) = -8.502, p < .001$, and right ROI, $t(19) = -9.983, p < .001$. All other two-way and three-way interactions (ROI × Group, Condition × ROI, and Condition × ROI × Group) were not statically significant ($p > .106$ in all cases).

Most relevant for the purpose of the present study is the statistically significant Condition × Group interaction. The data from the nonstuttering group indicated that the N1 amplitudes in the delayed-speaking condition, $t(9) = 4.856, p < .001$, and in the delayed-hearing condition, $t(9) = 6.064, p < .001$, were smaller than the N1 amplitude in the silent reading condition. On the other hand, the data from the stuttering group indicated no significant differences between the N1 amplitude from either the delayed-speaking or the delayed-hearing condition as compared with the silent reading condition ($p > .321$). Additionally, in the silent reading condition only, the N1 amplitude for the stuttering group was smaller than the N1 amplitude for the nonstuttering group, $t(18) = 2.947, p = .008$. Fig. 3A shows the average N1 amplitude for the stuttering and nonstuttering groups in all three conditions of delayed-speaking, delayed-hearing, and silent reading. The boxplots in Fig. 3B illustrate the stuttering and nonstuttering individual subject distributions with regard to amount of modulation in N1 amplitude (i.e., reduction in N1 amplitude relative to silent reading) in the delayed-speaking (green) and delayed-hearing conditions (orange).

N1 latency

N1 latency data showed a statistically significant main effect of Condition, $F(2,36) = 5.724, p = .007$, with latencies varying in the order delayed-hearing (*96.026 ms*), delayed-speaking (*98.404 ms*), and silent reading (*101.368 ms*). We also found a statistically significant main effect of ROI, $F(1.630, 29.343) = 4.758, p = .022$, with shorter latencies for the central ROI (*95.440 ms*) than for the left ROI (*99.642 ms*) and the right ROI (*100.716 ms*). No other main effects (Condition, Group) or two-way and three-way interactions (Condition × Group, ROI

× Group, Condition × ROI, Condition × ROI × Group) were statistically significant (*p* > *.323* in all cases).

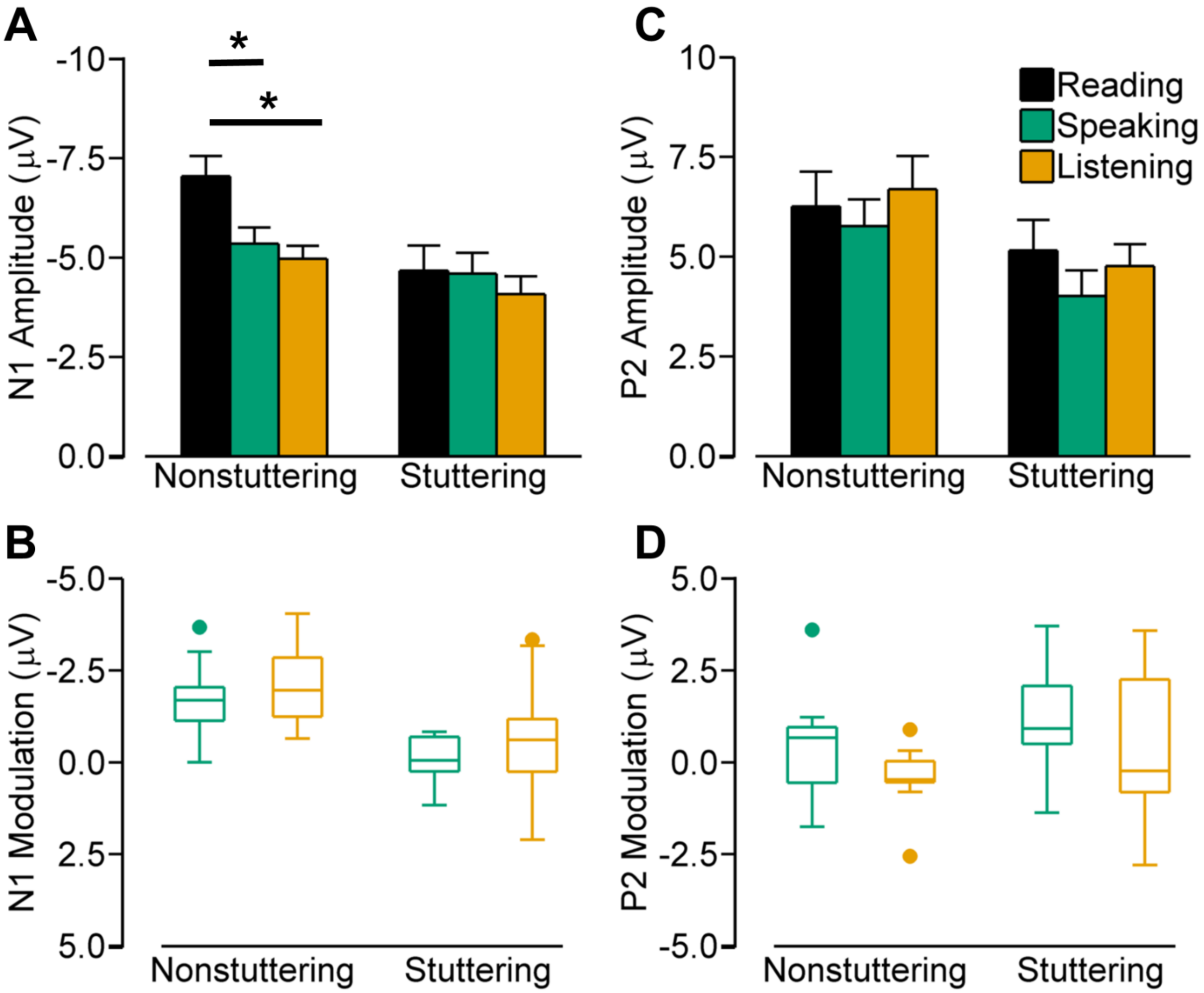

Fig. 3. The group average N1 amplitude (A) and P2 amplitude (C) for the stuttering and nonstuttering groups during the delayed-speaking condition (green), delayed-hearing condition (orange), and silent reading condition (black). The distributions of modulation (amplitude decrease in the delayed-speaking and delayed-hearing conditions vs. the silent reading condition) for N1 amplitude (B) and P2 amplitude (D) are shown as boxplots. We found a statistically significant Group × Condition interaction for N1 amplitude, indicating significant modulation in delayed-speaking and delayed-hearing conditions for the nonstuttering group, but no significant modulation for the stuttering group. The Group × Condition interaction for P2 was not statistically significant. The bar graph in each condition was calculated by averaging the data in all ROIs in that condition. Error bars correspond to standard errors. The asterisk corresponds to p < .001.

P2 amplitude

P2 amplitude showed significant main effects of Condition, *F(2, 36) = 4.496, p = .018*, and ROI, *F(2, 36) = 31.217, p < .001*, but these effects were modified by a statistically significant Condition × ROI interaction, *F(4, 72) = 4.264, p = .004*. For the central ROI, P2 amplitude was smaller in the delayed-speaking condition than in the silent reading condition, *t(19) = 6.396, p < .001*, and the delayed-hearing condition, *t(19) = 6.853, p <.001*. For the right and left ROIs, there were no statistically significant differences for the same comparisons (*p > .201*). The main effect of Group and other two-way and three-way interactions (Condition × Group, ROI × Group, and Condition × ROI × Group) were not statistically significant (*p > .115*). Fig. 3C shows the average P2 amplitude in all conditions for both the stuttering and nonstuttering groups, and the individual subject distributions for amount of P2 amplitude modulation during the delayed-speaking and delayed-hearing conditions are shown as boxplots in Fig. 3D.

P2 latency

P2 latency measures showed no statistically significant main effects of Condition, ROI, or Group (*p > .114*). There was, however, a statistically significant ROI × Group interaction, *F(2, 36) = 4.073, p = .025*: P2 latency for the stuttering group was shorter in the central ROI than in the left ROI, *t(9) = 2.834, p = .020,* and not in the right ROI (*p = .070*), whereas P2 latency for the nonstuttering group did not differ between pairs of conditions (*p > .170*). None of the other two-way and three-way interactions (Condition × Group, Condition × ROI, and Condition × ROI × Group) were statistically significant (*p > .370*).

Discussion

Previously, we have reported that the typically observed modulation of auditory processing during speech movement planning is limited in adults who stutter as compared with adults who do not stutter (Daliri & Max, 2014). This finding (and a compatible result for the somatosensory domain (McClean, 1996)) suggests that fluency breakdowns in individuals who stutter may be related to difficulties with appropriately priming the auditory (and somatosensory) system for upcoming inputs, which, in turn, may lead to feedback-driven actions that are disruptive rather than corrective (Max, 2004). The finding that stuttering speakers have problems with appropriately preparing for the sensory consequences of their own speech movements raised the new question whether an underlying core deficit can be attributed (a) specifically to the process of generating sensory predictions on the basis of an efference copy signal during movement planning, or (b) to a more global problem with using any available information to generate correct predictions regarding upcoming afferent inputs regardless of whether they are self-generated or externally generated.

To address this question, we recorded AEPs of stuttering and nonstuttering adults in response to probe tones delivered while anticipating either speaking (delayed-speaking condition) or hearing one's own speech played back (delayed-hearing condition) and a control condition without auditory input (silent reading condition). The delayed-speaking condition involved auditory predictions based on the speaker's own movement planning (in each trial, subjects heard their own speech feedback); the delayed-hearing condition involved auditory predictions in the absence of movement planning (in each trial, subjects heard their own pre-recorded speech); and the silent reading condition involved no predictions of auditory input. We hypothesized that (a) if stuttering is associated with specific difficulties in generating sensory predictions based on an efference copy, then a difference between the two groups in auditory modulation would be limited to the delayed-speaking condition, whereas (b) if stuttering is associated with a more general sensory prediction deficit, then a difference between the two groups would be expected in both the delayed-speaking and delayed-hearing conditions.

The primary findings relevant to the hypothesis under investigation revealed that (a) nonstuttering speakers showed a modulation of the N1 amplitude in both the delayed-speaking condition and the delayed-hearing condition as compared with the silent reading condition, but (b) stuttering speakers did not show a modulation of the N1 amplitude in either the delayed-speaking or the delayed-hearing conditions as compared with their silent reading condition. In other words, whereas the nonstuttering group modulated their auditory system in both conditions in which auditory input was predictable (either based on movement planning or in the absence of movement planning), the stuttering group failed to show this modulation for either of those two conditions. As another finding with direct relevance to the aims of the study, we observed a modulation of the P2 amplitude that was specific to the delayed-speaking condition (i.e., reduced P2 amplitude relative to both the delayed-hearing and silent reading conditions), but this P2 modulation—which occurred in the central ROI only—was observed in both the stuttering and the nonstuttering group. This absence of a between-group difference in P2 modulation is consistent with our previous results (Daliri & Max, 2014).

A number of additional findings may warrant further exploration in follow-up studies, but are not directly related to our primary hypotheses regarding AEP amplitude modulation in stuttering versus nonstuttering speakers. First, examining the latencies of the separate AEP components, we found a shorter N1 latency in the delayed-hearing condition in comparison with the delayed-speaking and silent reading conditions. Second, N1 latency was shorter in the central ROI than in the left and right ROIs. Third, whereas P2 latency of the stuttering individuals in the central ROI was shorter than that in the left ROI (but not right ROI), there were no significant differences between P2 latencies in the three ROIs for the nonstuttering individuals.

Thus, the main conclusion to be drawn from the present study is that whereas nonstuttering individuals predictively modulated auditory processing in *both* conditions in which auditory input was expected (i.e., both while anticipating self-producing speech and while anticipating hearing a played-back version of their own pre-recorded speech), stuttering individuals failed to show auditory modulation in both those conditions. In other words, the overall results suggest that stuttering adults' previously demonstrated failure to modulate the auditory system prior to speech onset (Daliri & Max, 2014) is part of a more global auditory prediction deficit that is present regardless of whether the prediction is based on self-generating a movement plan or on available contextual cues. Here, in the delayed-hearing condition, the sequence of events in each trial provided the cues that allowed subjects to predict when a previously recorded production of the word shown on the screen would be played back to them. In the delayed-speaking condition and during speech production in general, their own planning of movement sequences allows speakers to predict when the resulting auditory consequences will be heard. Our novel finding that stuttering speakers, in contrast to normally fluent speakers, do not modulate the auditory system in either of those two scenarios suggests that their lack of modulation during speech planning is most likely due to difficulties with rapidly and predictively *using* the information regarding expected auditory consequences, that results from generating and evaluating an efference copy signal rather than with the generation of this signal itself. We speculate that, during speech production, this inability to use such predictive information for appropriately priming task-relevant sensory systems for their role in monitoring articulatory movements (see Kim & Max, 2014) leads to unnecessary, and in fact disruptive, attempts at correcting ongoing movements; these "repairs" may contribute to the fluency breakdowns that form the primary symptoms of stuttering (Max, 2004).

Further research will be necessary to uncover the underlying functional or structural foundations for the observed auditory modulation deficit in stuttering individuals. However, given that the current study fully replicated our previous finding of atypical N1 modulation but typical P2 modulation, the data are consistent with an underlying problem whereby the central nervous system of stuttering individuals produces modulating signals at time points closer to movement onset (Daliri & Max, 2014). All probe tones were presented 400 ms after initial presentation of the target word and 200 ms before the visual change (the word's color changing from white to green) indicating that the word had to be spoken or would be played back. Thus, the evoked N1 and P2 components occurred approximately 500 and 600 ms, respectively, after initial presentation of the word. In stuttering individuals, the central neural signals modulating the auditory cortex may arrive only by the end of this 600 ms time window. However, as we have suggested previously (Daliri & Max, 2014), it is also possible that the underlying problem relates to suboptimal structural connectivity between the brain regions involved in generating predictive modulating signals and specific regions of auditory cortex (those that are the neural generators of the N1 component). Several studies on neurologically healthy individuals have shown that motor and premotor cortex are involved in the processing of predictions during movement planning as well as in the absence of movement (for review, see Schubotz & Von Cramon, 2003; Schubotz, 2007), and it is well known that there are abnormalities in structural and functional connectivity between motor/premotor and auditory regions in individuals who stutter (Chang, Erickson, Ambrose, Hasegawa-Johnson, & Ludlow, 2008; Chang, Kenney, Loucks, & Ludlow, 2009; Chang, Horwitz, Ostuni, Reynolds, & Ludlow, 2011; Cykowski et al., 2008; Kell et al., 2009; Sommer, Koch, Paulus, Weiller, & Buchel, 2002; Watkins, Smith, Davis, & Howell, 2008).

In summary, we recorded stuttering and nonstuttering adults' AEPs in response to probe tones delivered while anticipating speaking, hearing one's own speech played back, and a control condition with no auditory input. Nonstuttering speakers showed modulation of the N1 amplitude prior to both speaking and hearing one's own speech as compared with the control condition. Stuttering speakers showed no N1 modulation in either of these conditions. Hence, this work suggests that the previously demonstrated atypical pre-speech auditory modulation in stuttering speakers is due to a general auditory prediction deficit.

Acknowledgements

This research was supported, in part, by grants R01DC007603 and P30DC004661 from the National Institute on Deafness and Other Communication Disorders and grant MOP-137001 from the Canadian Institutes of Health Research. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute on Deafness and Other Communication Disorders, the National Institutes of Health, or the Canadian Institutes of Health Research.

References

Baess, P., Horvath, J., Jacobsen, T., & Schroeger, E. (2011). Selective suppression of self-initiated sounds in an auditory stream: An ERP study. *Psychophysiology, 48*(9), 1276-1283. doi:10.1111/j.1469-8986.2011.01196.x

Baess, P., Jacobsen, T., & Schroger, E. (2008). Suppression of the auditory N1 event-related potential component with unpredictable self-initiated tones: Evidence for internal forward models with dynamic stimulation. *International Journal of Psychophysiology, 70*(2), 137-143.

Bloodstein, O., & Bernstein-Ratner, N. (2008). *A handbook on stuttering* (6th ed.). Clifton Park, NY: Delmar.

Bubic, A., von Cramon, D. Y., & Schubotz, R. I. (2010). Prediction, cognition and the brain. *Frontiers in Human Neuroscience, 4,* 25. doi:10.3389/fnhum.2010.00025

Chang, S., 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. doi:10.1016/j.neuroiiiiage.2007.09.067

Chang, S., Horwitz, B., Ostuni, J., Reynolds, R., & Ludlow, C. L. (2011). Evidence of left inferior frontal-premotor structural and functional connectivity deficits in adults who stutter. *Cerebral Cortex, 21*(11), 2507-2518. doi:10.1093/cercor/bhr028

Chang, S., Kenney, M. K., Loucks, T. M. J., & Ludlow, C. L. (2009). Brain activation abnormalities during speech and non-speech in stuttering speakers. *Neuroimage, 46*(1), 201-212. doi:10.1016/j.neuroimage.2009.01.066

Cornelisse, L. E., Gagné, J., & Seewald, R. C. (1991). Ear level recordings of the long-term average spectrum of speech. *Ear and Hearing, 12*(1), 47-54.

Creutzfeldt, O., Ojemann, G., & Lettich, E. (1989). Neuronal-activity in the human lateral temporal-lobe .2. responses to the subjects own voice. *Experimental Brain Research, 77*(3), 476-489. doi:10.1007/BF00249601

Crone, N. E., Hao, L., Hart, J., Boatman, D., Lesser, R. P., Irizarry, R., & Gordon, B. (2001). Electrocorticographic gamma activity during word production in spoken and sign language. *Neurology, 57*(11), 2045-2053.

Cykowski, M. D., Kochunov, P. V., Ingham, R. J., Ingham, J. C., Mangin, J., Riviere, D., . . . Fox, P. T. (2008). Perisylvian sulcal morphology and cerebral asymmetry patterns in adults who stutter. *Cerebral Cortex, 18*(3), 571-583. doi:10.1093/cercor/bhm093

Daliri, A., & Max, L. (2014). *Modulation of auditory processing during speech movement planning is limited in adults who stutter.* Manuscript submitted for publication.

Daliri, A., Prokopenko, R., Flanagan, J. R., & Max, L. (2014). Control and prediction components of movement planning in stuttering vs. nonstuttering adults. *Journal of Speech, Language, and Hearing Research, 57*(6), 2131–2141. doi:10.1044/2014_JSLHR-S-13-0333

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.

Flanagan, J. R., & Lolley, S. (2001). The inertial anisotropy of the arm is accurately predicted during movement planning. *The Journal of Neuroscience, 21*(4), 1361-1369.

Gilmore, R. L. (1994). American-electroencephalographic-society guidelines in electroencephalography, evoked-potentials, and polysomnography. *Journal of Clinical Neurophysiology, 11*(1), 1-142.

Hughes, G., Desantis, A., & Waszak, F. (2013). Mechanisms of intentional binding and sensory attenuation: The role of temporal prediction, temporal control, identity prediction, and motor prediction. *Psychological Bulletin, 139*(1), 133.

Kell, C. A., Neumann, K., von Kriegstein, K., Posenenske, C., von Gudenberg, A. W., Euler, H., & Giraud, A. L. (2009). How the brain repairs stuttering. *Brain, 132*(10), 2747-2760. doi:10.1093/brain/awp185

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*, 911.

Martikainen, M. H., Kaneko, K., & Hari, R. (2005). Suppressed responses to self-triggered sounds in the human auditory cortex. *Cerebral Cortex, 15*(3), 299-302. doi:10.1093/cercor/bhh131

Max, L. (2004). Stuttering and internal models for sensorimotor control: A theoretical perspective to generate testable hypotheses. In B. Maassen, R. Kent, H. F. M. Peters, P. van Lieshout & W. Hulstijn (Eds.), *Speech motor control in normal and disordered speech* (pp. 357-387). Oxford, UK: Oxford University Press.

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.

Max, L., Daniels, J., Curet, K., & Cronin, K. (2008). Modulation of auditory and somatosensory processing during the planning of speech movements. Paper presented at the *Proceedings of the 8th International Seminar on Speech Production*, Strasbourg, France. 41-44.

McClean, M. D. (1996). Lip-muscle reflexes during speech movement preparation in stutterers. *Journal of Fluency Disorders, 21*(1), 49-60.

Mock, J. R., Foundas, A. L., & Golob, E. J. (2011). Modulation of sensory and motor cortex activity during speech preparation. *European Journal of Neuroscience, 33*(5), 1001-1011. doi:10.1111/j.1460-9568.2010.07585.x

Oostenveld, R., & Praamstra, P. (2001). The five percent electrode system for high-resolution EEG and ERP measurements. *Clinical Neurophysiology, 112*(4), 713-719.

O'Reilly, J. X., Mesulam, M. M., & Nobre, A. C. (2008). The cerebellum predicts the timing of perceptual events. *The Journal of Neuroscience, 28*(9), 2252-2260. doi:10.1523/JNEUROSCI.2742-07.2008

Riley, G. D. (2008). *SSI-4, Stuttering severity instrument for children and adults* (4th ed.). Austin, TX: Pro Ed.

Schubotz, R. I. (2007). Prediction of external events with our motor system: Towards a new framework. *Trends in Cognitive Sciences, 11*(5), 211-218.

Schubotz, R. I., & Von Cramon, D. Y. (2003). Functional–anatomical concepts of human premotor cortex: Evidence from fMRI and PET studies. *NeuroImage, 20*, S120-S131.

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

Sommer, M., Koch, M., Paulus, W., Weiller, C., & Buchel, C. (2002). Disconnection of speech-relevant brain areas in persistent developmental stuttering. *Lancet, 360*(9330), 380-383. doi:10.1016/S0140-6736(02)09610-1

Sperry, R. W. (1950). Neural basis of the spontaneous optokinetic response produced by visual inversion. *Journal of Comparative and Physiological Psychology, 43*(6), 482-489. doi:10.1037/h0055479

Wasaka, T., Hoshiyama, M., Nakata, H., Nishihira, Y., & Kakigi, R. (2003). Gating of somatosensory evoked magnetic fields during the preparatory period of self-initiated finger movement. *Neuroimage, 20*(3) doi:10.1016/S1053-8119(03)00442-7

Waszak, F., Cardoso-Leite, P., & Hughes, G. (2012). Action effect anticipation: Neurophysiological basis and functional consequences. *Neuroscience and Biobehavioral Reviews, 36*(2), 943-959. doi:10.1016/j.neubiorev.2011.11.004

Watkins, K. E., Smith, S. M., Davis, S., & Howell, P. (2008). Structural and functional abnormalities of the motor system in developmental stuttering. *Brain, 131,* 50-59. doi:10.1093/brain/awm241

Wolfensteller, U., Schubotz, R. I., & Von Cramon, D. Y. (2007). Understanding non-biological dynamics with your own premotor system. *NeuroImage, 36*, T33-T43.

Wolpert, D. M., Diedrichsen, J., & Flanagan, J. R. (2011). Principles of sensorimotor learning. *Nature Reviews Neuroscience, 12*(12), 739-751.

Wolpert, D. M., Miall, R. C., & Kawato, M. (1998). Internal models in the cerebellum. *Trends in Cognitive Sciences, 2*(9), 338-347. doi:10.1016/S1364-6613(98)01221-2

CHAPTER FOUR

Relationship between auditory modulation during speech planning and reliance on auditory feedback during speech production in adults who stutter

Ayoub Daliri and Ludo Max

Manuscript in preparation

Previously, we have shown that modulation of the auditory system during speech movement planning is limited in adults who stutter. However, the functional relevance of pre-speech auditory modulation is unknown. Therefore, we conducted to experiments to investigate the relationship between the auditory modulation during speech planning and reliance on auditory feedback during speech production in both nonstuttering and stuttering adults. The first experiment was designed to estimate the reliance on auditory feedback during speech production using a sensorimotor adaptation paradigm. We measured the change in formant frequencies of speech output (adaptation) when participant has experienced formant-shifted auditory feedback for several trials. In the second experiment, we recorded auditory evoked potentials in response to probe tones delivered prior to speech in a speaking condition with typical auditory feedback, a speaking condition with delayed auditory feedback (DAF; a manipulation selected to reduce the usefulness of the feedback signal), and in a no-speaking control condition (silent reading). Results of the first experiment showed that the extent of the adaptation to formant-shifted auditory feedback was limited for stuttering participants. Results of the second experiment showed that DAF caused a reduction in the modulation of N1 amplitude for nonstuttering speakers; however, DAF caused an increase in the modulation of N1 amplitude for stuttering speakers. Across the two groups, we found an overall relationship between change in N1 modulation in DAF condition and the extent of adaptation; when speaking with DAF as compared with typical feedback (NAF condition), the amount of N1 modulation decreased for participants who adapted more, but it increased for participants who adapted less. Overall, the obtained results suggest that auditory modulation during speech movement planning may be related to the extent of reliance on auditory feedback during speech production.

Introduction

Structural neuroimaging studies have revealed abnormalities in pathways suggested to be connecting speech motor regions and auditory regions of stuttering individuals (Chang, Erickson, Ambrose, Hasegawa-Johnson, & Ludlow, 2008; Chang, Horwitz, Ostuni, Reynolds, & Ludlow, 2011; Cykowski, Fox, Ingham, Ingham, & Robin, 2010; Sommer, Koch, Paulus, Weiller, & Buchel, 2002; Watkins, Smith, Davis, & Howell, 2008). Behavioral studies have shown reduced compensatory motor responses to unexpected auditory feedback perturbations in stuttering versus nonstuttering individuals (Cai, Beal, Ghosh, Guenther, & Perkell, 2014; Cai et al., 2012; Loucks, Chon, & Han, 2012). Together, these lines of research provide convincing evidence that stuttering is associated with deficits in the integration of auditory and motor information for speech production.

Sensorimotor integration involves, among others, processes related to (a) generating optimal control signals that achieve a desired movement goal, and (b) predicting the sensory consequences of these control signals (Kawato, 1999; Shadmehr, Smith, & Krakauer, 2010; Wolpert, Diedrichsen, & Flanagan, 2011; Wolpert, Miall, & Kawato, 1998). Predictions are used to optimize the motor commands and also to prepare task-relevant sensory systems for their subsequent roles in processing the sensory outcomes of the motor commands. To better understand the nature of deficits of the sensorimotor integration of stuttering individuals, we have recently investigated the prediction component by comparing the extent of modulation of the auditory system during speech movement planning (Daliri & Max, 2014). We recorded long latency auditory evoked potentials (LLAEPs) in response to probe tones delivered during the speech movement planning phase prior to speech onset as compared with no-speaking control conditions. Nonstuttering speakers showed significant modulation of the auditory system (i.e., observable as a reduction in amplitude of the N1 component in the LLAEP) during speech movement planning whereas stuttering speakers showed limited or no modulation of their auditory system. These findings suggested that stuttering is associated with deficits in using sensory prediction mechanisms to prime the auditory system for its role in processing upcoming auditory feedback. However, it remained unknown whether this limitation in pre-speech modulation also implies differences in stuttering speakers' reliance on auditory feedback during speech production. Therefore, the aim of this study was to investigate, for both stuttering and nonstuttering individuals, the relationship between the previously documented auditory modulation during speech planning and the speaker's reliance on auditory feedback during speech production.

To examine the relationship between a speaker's pre-speech auditory modulation and reliance on auditory feedback, we conducted two new experiments. In the first experiment, we used a sensorimotor adaptation paradigm to quantify each speaker's reliance on auditory feedback. In this paradigm, individuals' auditory feedback was digitally altered in real time (i.e., all formant frequencies were shifted up) while they produced monosyllabic words. Over the course of the experiment, participants learned to adapt to the feedback alteration (i.e., they lowered the formant frequencies in their speech output). We used the extent of auditory-motor adaptation as an estimate of the speaker's reliance on auditory feedback. In the second experiment, we then recorded the same speakers' LLAEPs in response to probe tones presented during speech movement planning in three conditions: (a) a speaking condition with typical auditory feedback (non-altered auditory feedback, NAF), (b) a speaking condition in which the usefulness of the auditory feedback signal was experimentally reduced (delayed auditory feedback, DAF), and (c) a no-speaking control condition (silent reading). Here, we hypothesized that if reliance on auditory feedback during speech production is related to auditory modulation during planning, the change in auditory modulation under DAF versus NAF would be larger for individuals with higher reliance on auditory feedback. Moreover, we investigated whether stuttering speakers' limited modulation of the auditory system during speech movement planning relates to a reduced reliance on auditory feedback during speech production.

Experiment I

Participants

After giving informed consent, 13 stuttering adults (11 males and 2 females; age 19–43 years, M = 28.71 years, SD = 7.96) and 13 nonstuttering adults (11 males and 2 females; age 20–45 years, M = 28.43 years, SD = 8.59) participated. Four stuttering participants had participated in our previous study (Daliri & Max, 2014). All participants were native speakers of American English, right-handed, and naive to the purpose of the study. Nonstuttering participants were individually matched with the stuttering participants based on age (± 3 years) and sex. Eligibility criteria for all participants included (a) no history of psychological, neurological, or communication disorders (other than stuttering in the stuttering group), (b) not taking medications with a possible effect on sensorimotor functioning, and (c) pure-tone hearing thresholds of 20 dB HL or less at all octave frequencies 250–4000 Hz in both ears.

An American Speech-Language-Hearing Association (ASHA)-certified speech-language pathologist or a graduate student with training and experience in the evaluation of stuttering confirmed the diagnosis of stuttering. All stuttering participants reported that their stuttering started during childhood (< 8 years of age). We used the Stuttering Severity Instrument, Fourth Edition (SSI–4; Riley, 2008), to determine each stuttering participant's severity. Individual stuttering participant's age, sex, SSI score, SSI stuttering severity classification, stuttering frequency (i.e., average frequency of stuttering across the SSI speaking and reading tasks) are listed in Table 1.

Table 1 Individual participant information for the stuttering group (age, sex, SSI score, SSI severity, and stuttering frequency). Participants are rank-ordered based on stuttering frequency.

Stuttering Participant	Age	Sex	SSI score	SSI severity	Stuttering frequency
S1	38	male	16	very mild	1.77
S2	26	male	12	very mild	1.85
S3	22	male	20	mild	2.36
S4	41	male	15	very mild	3.51
S5	26	male	27	moderate	5.29
S6	34	male	17	very mild	5.50
S7[a]	22	female	29	moderate	5.67
S8	19	female	29	moderate	6.59
S9	43	male	28	moderate	7.75

S10	20	male	21	mild	7.82
S11	28	male	32	severe	8.52
S12	32	male	28	moderate	10.38
S13	22	male	45	very severe	20.96

[a] *For participant S7, no recording of the reading task was available; S7's SSI score, SSI severity, and stuttering frequency are based on the speaking task only.*

Apparatus

Participants were seated on a chair inside a sound-attenuated booth (Fig.1A), approximately 1.5 m away from a computer monitor. The monitor was used to present target words and to provide visual feedback about participant's speech intensity. A microphone (SM58, Shure Incorporated, Niles, IL) was placed approximately 15 cm away from participants' mouth to transduce speech output. After amplification (DPS II, ART ProAudio, Niagara Falls, NY), the microphone signal was routed through a speech processor (VoiceOne, TC Helicon, Victoria, BC, Canada). The speech processor was used to alter all formant frequencies of the speech signal in real time (~10 ms delay) during some trials in the experiment (see Procedure). The output of the speech processor was then amplified (S-phone, Samson Technologies Corp., Syosset, NY) and played-back to the participant through insert earphones (ER-3A, Etymotic Research Inc., Grove Village, IL). A custom-written MATLAB script (The MathWorks, Inc., Natick, MA) was used to visually present target words and color-coded feedback about speech intensity (see Procedure), to control the speech processor, and to record the microphone signal and the speech processor's output at a sampling rate of 44100 Hz (16 bit). Fig.1A illustrates the apparatus for this experiment.

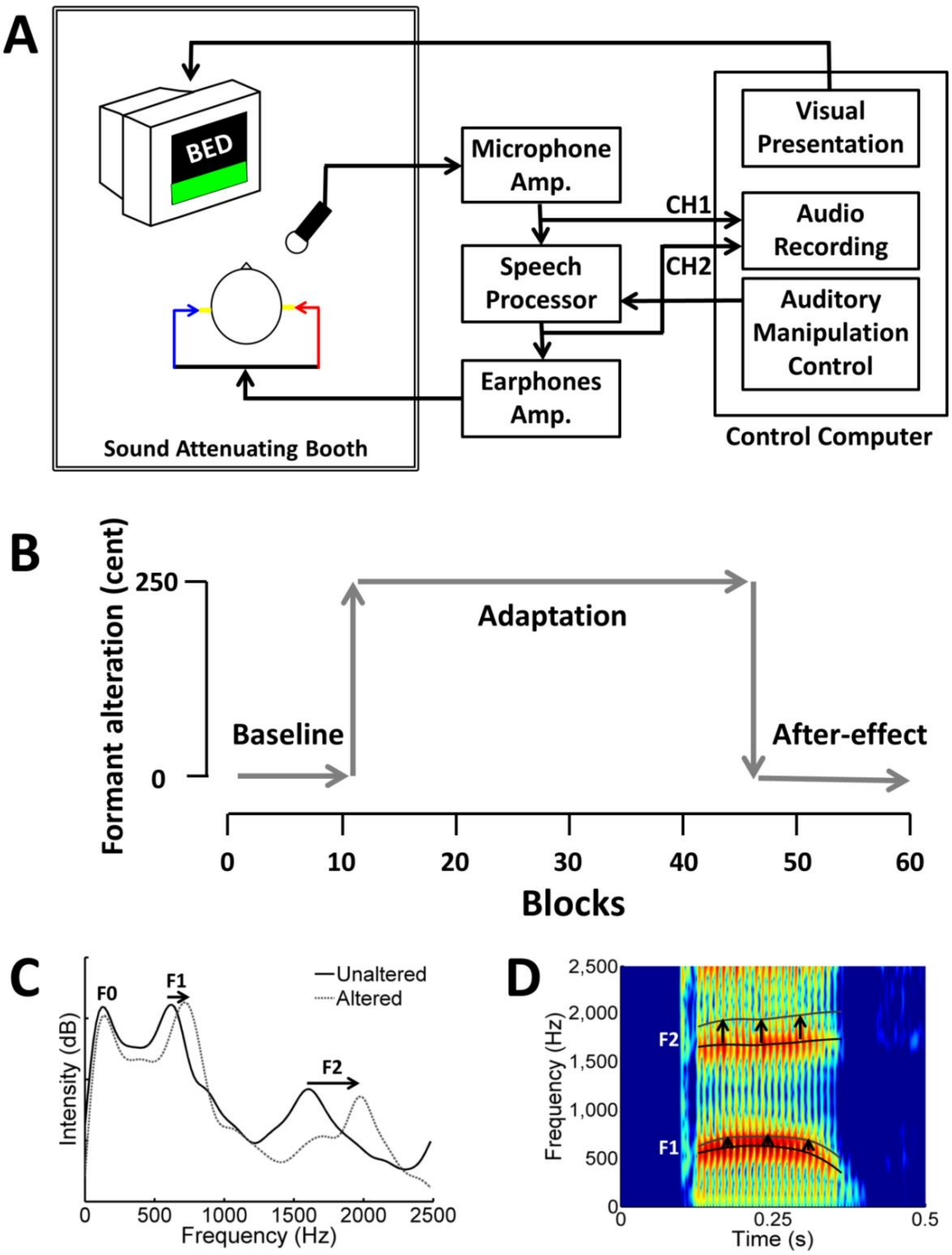

Fig 1. Instrumentation (A) and procedure (B) for Experiment I. During the adaptation phase, all formant frequencies were shifted upward by 250 cents. Panel C and D show a spectrum and a spectrogram of a sample speech signal, respectively. Black lines indicate the original un-altered spectrum (C) and formant trajectories (D) for the produced speech signal; grey lines indicate the same information for the formant-shifted auditory feedback signal (altered).

Before each recording session, the amplification level of the microphone and headphone amplifiers were calibrated such that speech input with a sound level of 68 dB SPL at the microphone resulted in 73 dB SPL output in the earphones. For calibration, the intensity of the auditory feedback in the earphones were measured using a 2 cc coupler (Type 4946, Bruel & Kjaer Inc., Norcross, GA) connected to a sound level meter (Type 2250A Hand Held Analyzer with Type 4947 ½” Pressure Field Microphone, Bruel & Kjaer Inc., Norcross, GA).

Procedure

The experiment consisted of two identical sessions (with a few minutes rest between the two sessions). In each session, participants completed 60 consecutive blocks of 3 trials. Each trial started with the presentation of a target word (/bɛd/, /bʌd/, or /bɔd/) on the monitor. Order of the words within each block was randomized. The presented word stayed on the monitor for 2.5 s. Participants were instructed to produce the word aloud with a sound level between 65 dB SPL and 71 dB SPL. After each trial, visual feedback (a colored horizontal bar at the bottom of the screen, see Fig.1A) of participant’s sound level was provided (red for > 71 dB SPL, blue for < 65 dB SPL, and green for > 65 dB SPL and < 71 dB SPL). To familiarize participants with the task and the visual feedback, they completed 10 practice blocks (30 trials) prior to the two recording sessions.

Each session consisted of three succeeding phases of trials (Fig.1B): (a) 10 blocks in a baseline phase (30 trials) in which participants produced the target words with unaltered auditory feedback, (b) 35 blocks in an adaptation phase (105 trials) in which participants produced the target words with while hearing auditory feedback in which all formant frequencies were shifted 250 cents up (1 octave equals 1200 cents) (Fig.1C–D), and (c) 15 blocks in an after-effects phase (45 trials) in which participants produced the target words again with unaltered auditory feedback.

Data analysis and statistical analysis

Prior to data analysis, the acoustic signals were downsampled to 22050 Hz. Based on visual inspection of the spectrogram, vowels were extracted manually. Then, the trajectories of the first two formant frequencies (F1, F2) were extracted automatically, using two formant tracking modules in Praat (Boersma, 2002) and The Snack Sound Toolkit (Sjölander, 2004). Formant trajectories from the two modules were visually inspected to select trajectories with fewer errors (miss tracking). Then, a custom-written MATLAB code was used to search the spectrogram to find the peaks in the spectrums (corresponding to the final formant trajectories) within 100 Hz from the selected trajectories. The final formant trajectories were smoothed using robust local regressions, visually inspected, and if necessary, manually corrected based on the spectrogram. The steady-state first formant (F1) and the second formant (F2) values for each trial were calculated by averaging estimates in a time window 40–60 % into the vowel.

To eliminate the effect of individual differences in formant frequencies in in the baseline phase, all formant values were converted from Hz to cents for each speaker separately, using the following formula: $F_{cents} = 1200\, Log_2(\frac{F_{Hz}}{B_{Hz}})$, where F_{Hz} on the left side of the equation corresponds to the F1 or F2 value (in Hz) for a given vowel, and B_{Hz} on the left side corresponds to the average formant values (in Hz) for all productions of the same vowel in the baseline phase (i.e., averaged across 10 trials). Thus, a speaker's extent of auditory-motor adaptation (i.e., change in speech output relative to baseline) was expressed in the same unit (i.e., cent) as the auditory feedback alteration implemented by the signal processor.

Our preliminary analysis did not reveal significant differences between the formant values in the first session and the second session; therefore, individual data for a given formant and a given vowel in the first session and the second session were averaged together. The average of formant values in the last five trials of the adaptation phase and the average of formant values in the last five trials of the after-effect phase (for each vowel and each subject) were entered into the statistical analysis.

We used the IBM SPSS Statistics 19 software package (IBM, Armonk, NY) for all analyses. To examine whether there were significant changes (compared to zero) for each of the formants in the adaptation phase and the after-effect phase (averaged across all three vowels) for each group, we included *a priori* planned comparisons using uncorrected one sample t-tests. For each formant separately, repeated measures analysis of variance (ANOVA) was carried out with Group (stuttering and nonstuttering) as a between-subjects factor, and with Vowel (/ɛ/, /ʌ/, and /ɔ/) and Phase (adaptation and after-effect) as repeated measures. Given that all formant values were normalized based on formants in the baseline phase, baseline phase was not included in the analysis. To account for potential violations of the sphericity assumption in conditions with more than two levels, the degrees of freedom were adjusted using the Greenhouse-Geisser correction (Max & Onghena, 1999). Statistically significant interactions were followed up with post-hoc analyses, using t-tests with Bonferroni corrections for multiple comparisons. For the stuttering group, we used Pearson correlation coefficients to examine relationships between participants' stuttering frequency (the average percent stuttered syllables across the speaking and reading tasks of the SSI–4, see Table 1) and the dependent variables.

Results

Fig.2A shows both groups' average change in F1 throughout the experiment. To ensure that participants adapted to the alteration, we conducted four planned comparison one-sample t-tests to compare extent of change in F1 in each phase (averaged across the three vowels). Our analysis showed that the F1 for the nonstuttering group statistically significantly changed in response to the 250-cent upward formant alteration in both the adaptation phase and the after-effect phase ($p < .001$); however, this was not the case for the stuttering individuals in either of the phases ($p > .128$). For the first formant, we found statistically significant main effects of Phase, $F(1, 24) = 12.630, p = .002$, and Group, $F(1, 24) = 11.447, p = .002$. These main effects were modified by a statistically significant Phase × Group interaction, $F(1, 24) = 7.749, p = .010$. This interaction indicated that whereas F1 for the nonstuttering group in the adaptation phase was different from F1 in the after-effect phase, $t(12) = -4.186, p = .001$, F1 for the stuttering group in the adaptation phase did not differ from that in the after-effect phase ($p = .571$). No statistically significant main effect of Vowel, two-way interactions (Vowel × Group and Phase × Vowel), or three-way interactions (Phase × Vowel × Group) were found ($p > .169$). Fig.2B shows the average F1 in each of the phases for the two groups.

The overall change in F2 throughout the experiment is shown in Fig.2C. The planned comparisons showed significant F2 change for nonstuttering individuals in both phases ($p < .001$); however, this was not the case for the stuttering group ($p > .060$). Examining F2, we found statistically significant main effect of Phase, $F(1,24) = 19.994$, $p < .001$, and significant Phase × Vowel interaction, $F(1.612, 38.683) = 4.219, p = .029$. This interaction indicated that, in the adaptation phase, F2 for vowel /ɔ/ was larger than F2 for vowel /ɛ/ (one-tail t-test: $t(25) = -2.648, p = .007$) and F2 for vowel /ʌ/ was larger than F2 for vowel /ɛ/ (one-tail t-test: $t(25) = -2.272, p = .016$); however, no significant difference between vowels were found in the after-effect phase ($p > .147$). We found statistically significant main effect of Group, $F(1, 24) = 4.517, p = .044$, indicating larger F2 for the nonstuttering group than the stuttering group. We also found a near significant main effect of Vowel ($p = .057$). All other two-way and three-way interactions were not statistically significant ($p > .220$).

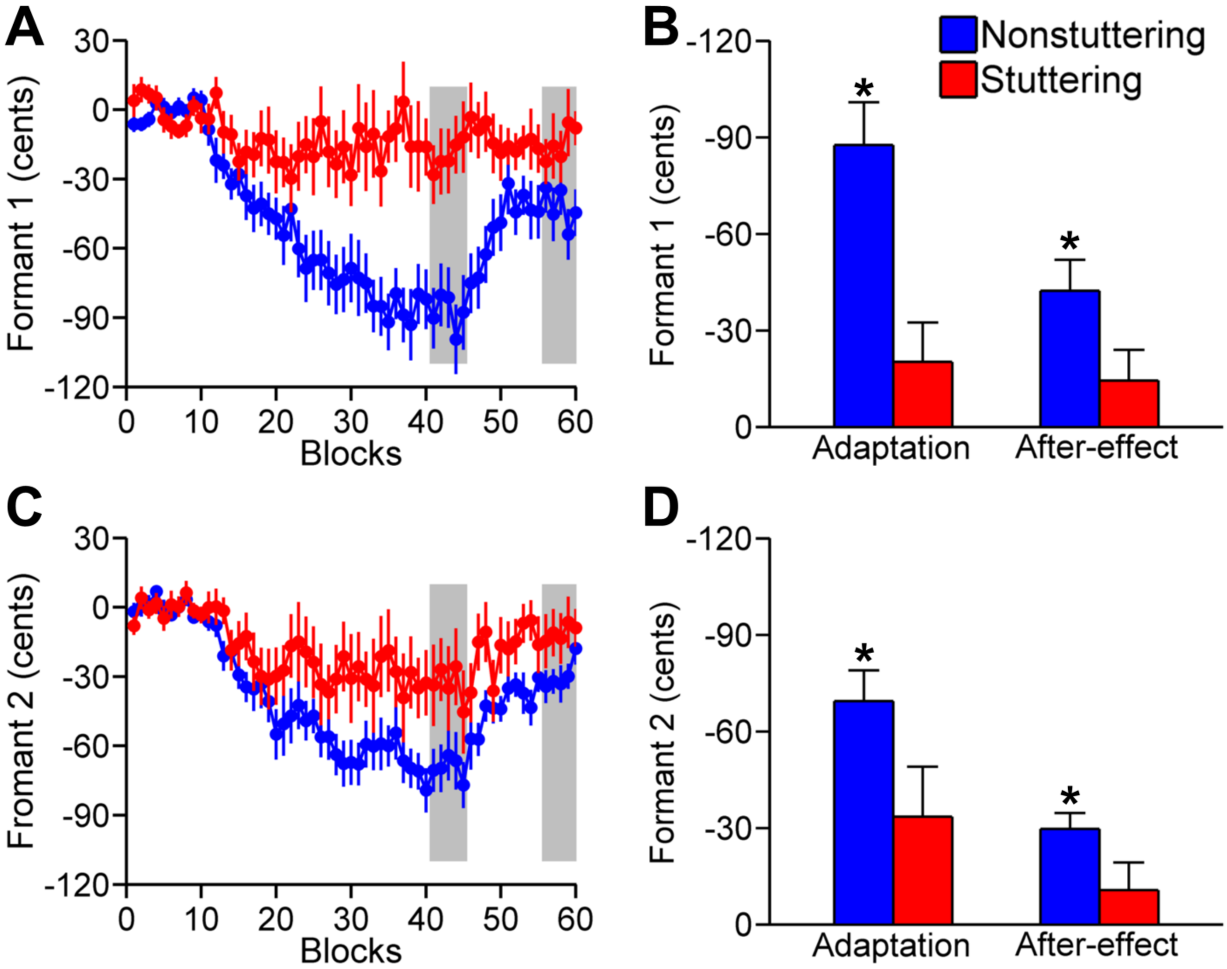

Fig.2. Panel A and C show the average formant values in cents (for all three vowels) during successive phases of baseline, adaptation, and after-effect. The average data in the shaded areas that correspond to the end of adaptation phase and after-effect phase were entered into the statistical analysis. Panel B and D show the average F1 and F2 for each group in the two phases. The stuttering group is shown in Red and the nonstuttering group is shown in cyan. Whereas the nonstuttering group adapted to the alteration, the stuttering group did not adapt. Error bars correspond to standard errors, and asterisks correspond to $p < .05$.

Examining the relationship between F1 and F2 with stuttering frequency, we found significant correlation between F1 in the adaptation phase with stuttering severity, $r = .605$, $p = .028$, indicating more severe individuals adapted less to the alteration. None of the other relationships were significant ($p > .096$).

Experiment II

Participants

The same stuttering and nonstuttering groups participated in the second experiment. Stuttering participant S11 and his matched nonstuttering participant were excluded from the analysis due to excessive noise in the data. Four stuttering participants had participated in our previous study (Daliri & Max, 2014).

Apparatus

The apparatus was similar to the one in the experiment I. Participants were seated on a chair inside a sound-attenuated booth (Fig.1A), approximately 1.5 m away from a computer monitor. A microphone (WL185, Shure Incorporated, Niles, IL) was placed approximately 15 cm away from participants' mouth to transduce speech output. After amplification with the microphone amplifier, the microphone signal was entered into the speech processor. The speech processor in this experiment was used to delay the signal by 100 ms (including the 10 ms inherent delay) in some blocks. The output of the speech processor was then amplified and played-back to participants through the insert earphones. The insert earphones were used to deliver binaural auditory stimuli (1 kHz, 40 ms duration, 10 ms rise/fall time, 75 dB SPL). Before each recording session, the amplifiers (microphone and headphone) were adjusted such that speech input with a sound level of 68 dB SPL at the microphone resulted in 73 dB SPL output in the earphones.

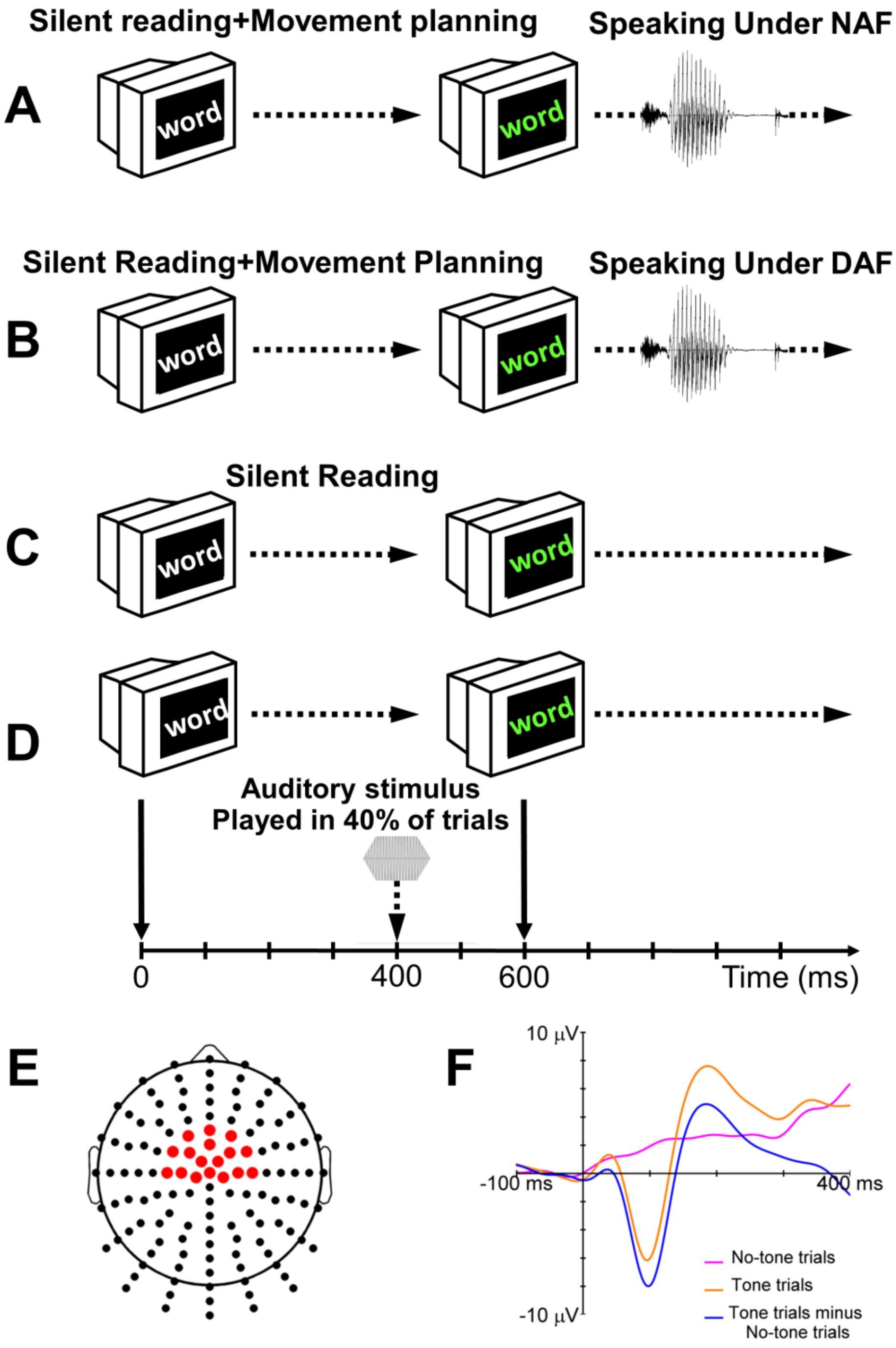

Fig.3. Experimental procedures for speaking under non-delayed auditory feedback (NAF) (A), speaking under delayed auditory feedback (DAF) (B), and silent reading (C) conditions. A probe auditory stimulus (1 kHz; 40 ms duration; 75 dB SPL) was played on 40% of trials (tone trials) for each condition (D). No auditory stimulus was played in the remaining trials (no-tone trials). The average response in 17 electrodes (E) were calculated. For each subject and in each condition, the average of no-tone trials was subtracted from the average of tone trials (F). This procedure was used to find the best estimate of the auditory cortex's response to the probe tones.

Procedure

For most participants, the experiments I and II were conducted one week apart. The experiment included three conditions (Fig.3A–C) with each condition consisting of 3 blocks of 90 trials. Each trial started with presentation of a word in white character on a black background; after 600 ms, the color of the word changed to green. The change in color served as a go-signal. The word in each trial was randomly selected from a list of 90 3–5 letters long, monosyllabic, consonant-vowel-consonant (CVC), bilabial words with no consonant cluster. In the first condition, participants read the target words aloud after the go-signal while receiving non-delayed auditory feedback (NAF). In the second condition, participants read the words aloud after the go-signal while receiving delayed auditory feedback (DAF). The third condition served as a control condition in which participant silently read the words. As shown in Fig.3D, 400 ms after presentation of the word in white characters, binaural auditory stimuli (1 kHz tone) were delivered on only 40% of trials in each block (tone trials) and no auditory stimuli were presented in the rest of trials (no-tone trials). The time interval from the end of a trial to the beginning of the next trial was randomly selected from a set of five possible intervals (1500, 2000, 2500, 3000, or 3500 ms). The order of conditions/blocks was randomized for each participant.

Electroencephalographic Recordings

We used Active Two BioSemi system (BioSemi, Amsterdam, The Netherlands) to record continuous electroencephalographic (EEG), electromyographic (EMG), and electrooculographic (EOG) activity. EEG recordings were taken from 128 standard sites on the scalp according to an extension of the international 10-10 electrode system (Gilmore, 1994; Oostenveld & Praamstra, 2001). In addition, signals from two electrodes placed over the left and right mastoids were recorded. Mastoid electrodes were used for off-line re-referencing. The EOG signals generated from blinks and eye movements were recorded using two electrodes placed below the lower eyelid and next to the outer canthus of the left eye. Surface EMG signals generated from upper and lower lip muscles were recorded by means of four electrodes placed around the lips (right and left orbicularis oris superior and inferior: upper and lower lip elevation/rounding). Acoustic signals from an additional microphone (SM58, Shure, Niles, IL), together with EEGs, EOGs, EMGs, and mastoid electrodes were recorded with a sampling rate of 1024 Hz.

Data analysis

First, EEG signals were re-referenced to an average of the mastoid signals, and low-pass filtered at 50 Hz. Data were then epoched into segments from 100 ms before the auditory stimulus to 400 ms after the onset of the auditory stimulus in tone trials, and the equivalent time window in the no-tone trials. Epochs were baseline-corrected by subtracting the average of pre-stimulus baseline (100 ms) from the whole epoch. Epochs with amplitudes greater than ±100 µV were excluded from the analysis. Epochs were further visually inspected to exclude trials with artifacts associated with excessive muscle activity, muscle activity before the go-signal, blinking, or eye movements.

Next, for each condition, the artifact free epochs for all tone trials and all no-tone trials were averaged, separately. The average of tone trials reflected brain activity related to auditory processing (as a result of the auditory stimuli) and non-auditory processing (e.g., activity related to motor, linguistic, cognitive, and visual processes), whereas the average of no-tone trials reflected brain activity related to the non-auditory activity only. To correct for non-auditory activity present in the average of tone trials, the average of non-tone trials was subtracted from the average of tone trials (Fig.3F). The resultant signal, which was used as the final LLAEP, provided a good estimate of the auditory activity only (Baess, Horvath, Jacobsen, & Schroeger, 2011; Baess, Jacobsen, & Schroger, 2008; Daliri & Max, 2014; Martikainen, Kaneko, & Hari, 2005). The final LLAEPs were further low-pass filtered at 15 Hz.

For the data analysis, we focused on attenuation of N1 and P2 components of the LLAEP. Our previous studies have shown that the extent of attenuation is the largest in frontocentral region (Daliri & Max, 2014); therefore, for our analysis, we included 17 electrodes located in this region of interest (ROI; Fig.3E). LLAEPs in these 17 channels were then averaged in each condition and entered into the statistical analysis. The amplitude and latency of N1 and P2 components from the average LLAEPs over the ROI were extracted for each participant. N1 was defined as the largest negative peak between 70 and 130 ms following the onset of the auditory stimulus. P2 was defined as the largest positive peak between 150 and 250 ms following the onset of the auditory stimulus. The data analysis was carried out using the EEGLAB toolbox (Delorme & Makeig, 2004) and custom-written MATLAB scripts.

Statistical analyses

For each dependent variable, we conducted a repeated measure of ANOVA with Group (stuttering and nonstuttering) as a between-subjects variable, and Condition (NAF, DAF, and reading) as within-subjects variable. Using Greenhouse-Geisser correction, degrees of freedom for the within-subjects comparisons were adjusted to account for potential violations of the sphericity assumption (Max & Onghena, 1999). Given the hypothesis under investigation, we included *a priori* planned comparisons into the repeated ANOVA to examine whether or not there were significant N1 modulation (N1 amplitude in reading condition minus N1 amplitude in in speaking under NAF or DAF) for each group, using (uncorrected) one sample t-tests. To explore statistically significant interactions, we conducted post-hoc analyses by means of t-tests with Bonferroni corrections for multiple comparisons. For the stuttering group, Pearson correlation coefficients was used to investigate possible relationships between participants' stuttering frequency and any of the dependent variables or the magnitude of N1 and P2 modulations.

Results

Fig.4 shows the grand-average LLAEPs in all conditions for the nonstuttering group (A) and the stuttering group (B). Statistical analysis of N1 amplitude revealed statistically significant main effect of Condition, $F(1.571, 34.568) = 28.594, p < .001$, and Group by Condition interaction, $F(1.571, 34.568) = 15.142, p < .001$. The main effect of Group was not statistically significant ($p = .992$). Fig.4C shows the average N1 amplitude in each of the conditions for the two groups. As shown in Fig.4D, follow-up planned comparisons indicated statistically significant N1 modulation in the speaking under NAF, $t(11) = 13.405, p < .001$, and in the speaking under DAF, $t(11) = 8.011, p < .001$. For the stuttering group on the other hand, follow-up planned comparisons showed statistically significant N1 modulation in speaking under DAF, $t(11) = 2.555, p = .027$, but not in the speaking under NAF ($p = .798$). Interestingly, N1 modulation of nonstuttering group in the NAF condition was greater than N1 modulation in the DAF condition (one tail t-test: $p = .023$). In contrast, N1 modulation of stuttering group in the NAF condition was smaller than N1 modulation in the DAF condition (one tail t-test: $p = .015$).

Examining P2 amplitude, we found significant main effect of Condition, $F(1.990, 43.775) = 3.549, p = .037$, with smaller P2 amplitude in the speaking under NAF (M = 6.0145 µV, SD = 2.828), and speaking under DAF (M = 6.014 µV, SD = 3.434) in comparison to P2 amplitude in the reading condition (M = 7.443 µV, SD = 4.327). The main effect of Group and Group by Condition interaction were not statistically significant ($p > .089$). Our analysis did not reveal statistically significant main effects of Group and Condition, and Group by Condition interaction for both N1 latency and P2 latency ($p > .181$).

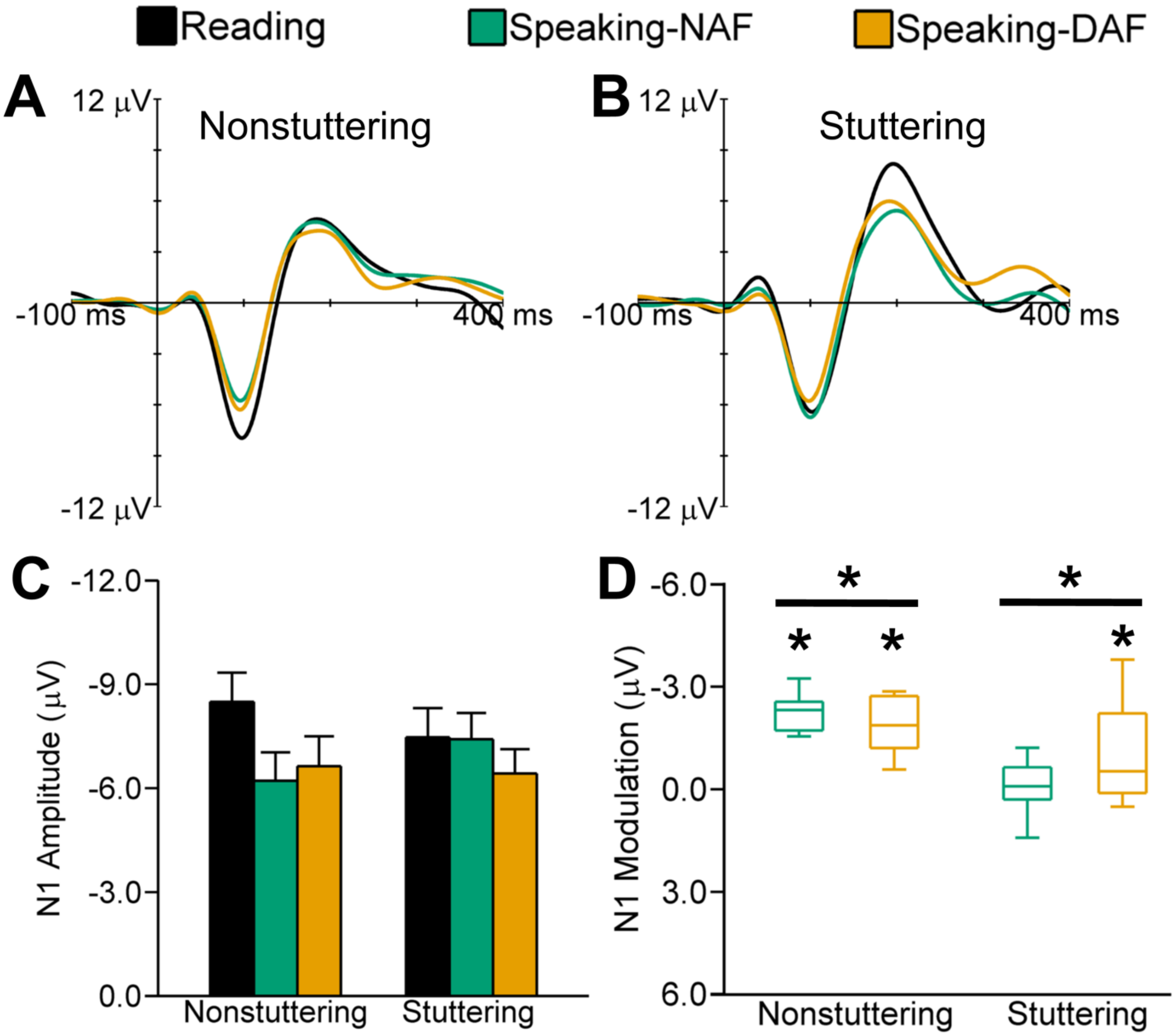

Fig.4. LLAEPs of the nonstuttering group (A) and the stuttering group (B) in the speaking under NAF (green), speaking under DAF (orange), and silent reading (black). The group average N1 amplitude and the distributions of N1 modulation (amplitude decrease in the speaking conditions vs. the reading condition) are shown in C and D, respectively. We found a statistically significant Group × Condition interaction for N1 amplitude, indicating significant modulation in speaking under NAF and DAF for the nonstuttering group, and only significant modulation in speaking under DAF for the stuttering group. Error bars correspond to standard errors. The asterisks correspond to $p < .050$.

We found no statistically significant relationship between dependent variables and stuttering frequency ($p > .072$). We also did not find a significant relationship between N1 modulation in the speaking under NAF and stuttering frequency ($p = .472$); however, there was a significant correlation between stuttering frequency and N1 modulation in the speaking under DAF, $r = .655$, $p = .021$ (i.e., more N1 modulation under DAF in more severe stuttering individuals). No significant relationship between P2 modulation in the speaking under NAF or DAF and stuttering frequency were found ($p = .567$). For the nonstuttering group, there was a positive correlation between N1 modulation in speaking under NAF and N1 modulation in speaking under DAF, $r = .610$, $p = .035$, but no such a relationship for P2 modulation ($p = .085$). There was no relationship between N1 modulations in the two conditions for stuttering individuals ($p = .304$), and only a near significant relationship for P2 modulation ($r = .517$, $p = .050$).

Exploring the relationship between auditory-motor adaptation and auditory modulation

As a measure of reliance on auditory feedback, we calculated an index of adaption for each individual by averaging formant values at the end of the adaptation phase in Experiment I (averaged over all vowels and formants) divided by -250. As shown in Fig.5A, for the nonstuttering group, we did not find statistically significant correlations between the adaptation index and N1 modulation in conditions of speaking under NAF or DAF ($p > .129$). For the stuttering group (Fig.5B), however, we found significant correlations between the adaptation index and N1 modulation in speaking under NAF, $r = 0.655$, $p = 0.021$, and N1 modulation in speaking under DAF, $r = 0.590$, $p = 0.043$. No significant correlations were found between adaptation index and P2 modulations ($p > .501$).

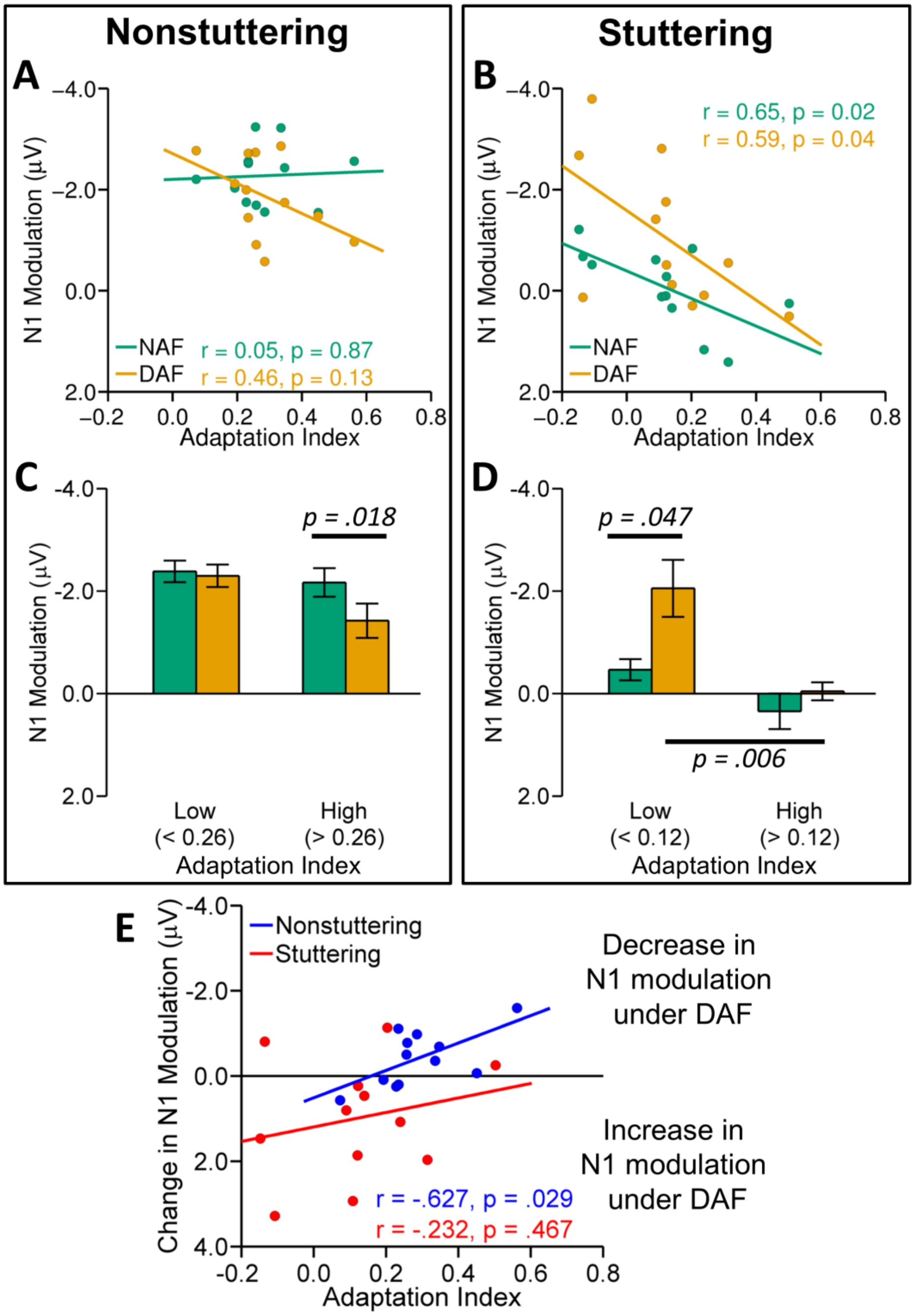

Fig.5. Panel A and B show the relationship between N1 modulation in speaking under NAF (green) and speaking under DAF (orange) and index of adaptation for the nonstuttering individuals and stuttering individuals, respectively. Each group was split to two sub-groups based on the index of adaptation (below and above median). The average N1 modulation for the nonstuttering sub-groups and stuttering sub-groups are shown in C and D. Error bars correspond to standard errors. Panel E shows the relationship between the change in N1 modulation in DAF versus NAF for nonstuttering (blue) and stuttering participants (red).

To further explore the relationship between adaptation index and N1 modulation, each group was split into two sub-groups based on median adaptation index for that group (i.e., High and Low: above median and below median). Our analysis showed that the extent of N1 modulation of, only, high adapting nonstuttering participants in the NAF condition differed from that in the DAF condition, *t(5) = -3.464, p = .018* (Fig.5C). As shown in Fig.5D, for the stuttering group, we found that (a) N1 modulation of low adapting participants in the NAF condition differed from that in the DAF condition, *t(5) = 2.611, p = .047,* and (b) N1 modulation of low adapting participants in the DAF condition differed from the N1 modulation of high adapting participants in the DAF condition, *t(10) = -3.445, p = .006*. It should be noted that, descriptively, the average stuttering frequency for low adapting stuttering participants (M = 8.7183%, SD = 6.572) was higher than stuttering frequency for high adapting stuttering participants (M = 4.5233%, SD = 2.520). There was a near significant correlation between the index of adaptation and stuttering frequency for the stuttering participants, *r = -0.563, p = .056,* indicating more severe individuals had lower index of adaptation.

To examine whether the change in auditory modulation under DAF versus NAF is higher for individuals with higher reliance on auditory feedback, we calculated the difference between N1 modulation under NAF and N1 modulation under DAF. We found a statistically significant relationship between index of adaptation and the change in N1 modulation for the nonstuttering group, *r = .627, p = .029*, but not for the stuttering group (*p = .467*). As shown in Fig.5E, whereas the N1 modulation under DAF for most nonstuttering individuals (8 out of 12) was smaller than N1 modulation under NAF, the N1 modulation under DAF for most stuttering individuals (9 out of 12) was greater than N1 modulation under NAF, *t(22) = 3.187, p = .004*.

Discussion

It has been suggested that stuttering is associated with deficits in integration of auditory system and motor information (Cai et al., 2012; Loucks et al., 2012; Max, 2004). Examining the sensorimotor integration in stuttering, in our previous study (Daliri & Max, 2014), we have shown that stuttering individuals have difficulties with modulating auditory system during speech planning. However, it remained unknown whether or not the pre-speech auditory modulation is associated with reliance on auditory feedback control mechanisms during speech production. Therefore, we conducted two experiments to investigate the relationship between the previously documented auditory modulation during speech planning and reliance on auditory feedback during speech production in both stuttering and nonstuttering adults.

In the first experiment, we used a sensorimotor adaptation paradigm in which participants experienced a sudden alteration in auditory feedback (250 cents upward shift in all formant frequencies) while producing monosyllabic CVC words. As a measure of reliance on auditory feedback, we estimated the extent of adaptation (i.e., change in produced formants over the course of experiment). We found significant between-group differences in the extent of adaptation. The adaptation for the nonstuttering group (overall adaptation: M = 31.5%, SD = 14.8; range: 7.3–60.2%) was larger than the adaptation for the stuttering group (overall adaptation: M = 10.6%, SD = 18.9; range: -14.8–50.3%). These results suggested that stuttering individuals may rely less on auditory feedback, and therefore, less likely to change their production to adapt to the alteration. Our findings are in agreement with previous studies that have shown reduced compensatory motor response to alteration in auditory feedback in stuttering speakers versus nonstuttering speakers (Cai et al., 2014; Cai et al., 2012; Loucks et al., 2012).

In the second experiment, we recorded LLAEPs in response to probe tones presented during speech movement planning in three conditions: (a) a speaking condition with normal reliance on auditory feedback—speaking while receiving non-delayed auditory feedback (NAF), (b) a speaking condition with low reliance on auditory feedback—speaking while receiving delayed altered auditory feedback (DAF), and (c) a no-speaking control condition—silent reading. Examining N1 modulation—N1 amplitude in reading condition minus N1 amplitude in speaking conditions—in the NAF and DAF conditions, we found significant N1 modulation in both conditions for nonstuttering participants and only significant N1 modulation for stuttering participants in DAF condition. N1 modulation of the nonstuttering group in the NAF condition was larger than the N1 modulation in the DAF condition; in contrast, N1 modulation of the stuttering group in the NAF condition was smaller than the N1 modulation in DAF condition. Supporting our hypothesis, these findings suggested that pre-speech auditory modulation may be related to the reliance on auditory feedback. The lack of auditory modulation of the stuttering group in the NAF condition was consistent with our previous study (Daliri & Max, 2014).

Examining the relationship between auditory modulation and reliance on auditory feedback, we found a significant correlation between the adaptation index and the change in N1 modulation in the DAF condition versus NAF condition for the nonstuttering individuals but not for the stuttering individuals. These results indicated that there was a larger reduction in N1 modulation in the DAF condition compared to the NAF condition for nonstuttering individuals. Whereas the N1 modulation in the DAF condition decreased for 8 out of 12 nonstuttering participants, the N1 modulation in the DAF condition increased for 9 out of 12 stuttering participants. These results provided evidence for the hypothesis that the change in auditory modulation under DAF versus NAF would be larger for individuals with higher reliance on auditory feedback. We also examined the relationship between adaptation index and N1 modulation in each condition. Whereas no significant relationships were found for the nonstuttering group, there were significant correlations for the stuttering group in both conditions of NAF and DAF, indicating larger adaptation index is associated with smaller N1 modulation. Further examination of the relationship between average adaptation for each participant and his/her level of N1 modulation in speaking under NAF and DAF, we found that N1 modulation in the DAF condition was smaller than N1 modulation in the NAF condition for those nonstuttering participants who adapted more. In addition, N1 modulation of low adapting stuttering individuals in the DAF condition was significantly improved to the level of N1 modulation seen in nonstuttering individuals. This result and a near significant ($p = .056$) relationship between adaptation index and stuttering frequency indicated larger improvement in N1 modulation in the DAF condition for more severe stuttering individuals.

One interpretation of these results is that, decrease in N1 modulation is associated with decrease in the weight of the auditory feedback when auditory feedback is less reliable (i.e., delayed auditory feedback). Therefore, individuals who rely more on auditory feedback decrease the N1 modulation more when auditory feedback is less reliable. Moreover, the low adapting individuals rely less on auditory feedback and, therefore, less sensitive to the delay in auditory feedback, and less likely to decrease the modulation. However, if decrease in N1 modulation is associated with decrease in reliance on auditory feedback, one may expect to find a relationship between N1 modulation in NAF condition and adaptation index. The lack of correlation (i.e., similar level of N1 modulation for high and low adapting individuals in the NAF condition) may seem inconsistent with the proposed interpretation. The lack of correlation may be explained by the fact that the task (producing monosyllabic words) was simple and needed low reliance on auditory feedback in general. To address this explanation, future research need to examine the relationship between auditory modulation and adaption with longer words that require more reliance on auditory feedback. Together, our findings partially suggest that decrease in auditory modulation during speech movement planning is associated with decrease in reliance on auditory feedback during speech production.

Based on this framework (i.e., less auditory modulation is associated with a decrease in the weight of auditory feedback), the lack of N1 modulation in stuttering may be interpreted as decrease in weight on auditory feedback. This interpretation may also be used to explain the lower level of adaptation for stuttering individuals (i.e., low reliance on auditory feedback). However, it remains unknown why N1 modulation in the DAF condition was improved for the low adapting stuttering individuals and not the high adapting stuttering individuals. Our analysis showed that low adapting stuttering individuals were more severe with greater stuttering frequency. In fact, more improvement in N1 modulation in the DAF condition for more severe stuttering individuals is consistent with studies that have shown more improvement in speech fluency under DAF for more severe stuttering individuals (for a review, see Lincoln et al., 2006).

In sum, we examined the relationship between sensorimotor adaptation and pre-speech auditory modulation when participants received normal auditory feedback and when they received delayed auditory feedback. Whereas nonstuttering individuals who adapted more to the auditory alteration were more likely to decrease the extent of auditory modulation under delayed auditory feedback, nonstuttering individuals who adapted less did not change the extent of auditory modulation under delayed auditory feedback. Results showed that stuttering individuals who adapted more failed to modulate their auditory system in the DAF condition; however, those who adapted less benefited from the delayed auditory feedback and modulated their auditory system. Overall, our results suggested that auditory modulation during movement planning may be associated with reliance on auditory feedback during speech production.

References

Baess, P., Horvath, J., Jacobsen, T., & Schroeger, E. (2011). Selective suppression of self-initiated sounds in an auditory stream: An ERP study. *Psychophysiology, 48*(9), 1276–1283. doi:10.1111/j.1469-8986.2011.01196.x

Baess, P., Jacobsen, T., & Schroger, E. (2008). Suppression of the auditory N1 event-related potential component with unpredictable self-initiated tones: Evidence for internal forward models with dynamic stimulation. *International Journal of Psychophysiology, 70*(2), 137–143.

Boersma, P. (2002). Praat, a system for doing phonetics by computer. *Glot International, 5*(9/10), 341–345.

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*, 24–29.

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), e41830. doi:10.1371/journal.pone.0041830

Chang, S., 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. doi:10.1016/j.neuroiiiiage.2007.09.067

Chang, S., Horwitz, B., Ostuni, J., Reynolds, R., & Ludlow, C. L. (2011). Evidence of left inferior frontal-premotor structural and functional connectivity deficits in adults who stutter. *Cerebral Cortex, 21*(11), 2507–2518. doi:10.1093/cercor/bhr028

Cykowski, M. D., Fox, P. T., Ingham, R. J., Ingham, J. C., & Robin, D. A. (2010). A study of the reproducibility and etiology of diffusion anisotropy differences in developmental stuttering: A potential role for impaired myelination. *Neuroimage, 52*(4), 1495–1504. doi:10.1016/j.neuroimage.2010.05.011

Daliri, A., & Max, L. (2014). *Modulation of auditory processing during speech movement planning is limited in adults who stutter.* Manuscript submitted for publication.

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.

Gilmore, R. L. (1994). American-electroencephalographic-society guidelines in electroencephalography, evoked-potentials, and polysomnography. *Journal of Clinical Neurophysiology, 11*(1), 1–142.

Kalinowski, J., Armson, J., Stuart, A., & Gracco, V. L. (1993). Effects of alterations in auditory feedback and speech rate on stuttering frequency. *Language and Speech, 36*(1), 1–16.

Kawato, M. (1999). Internal models for motor control and trajectory planning. *Current Opinion in Neurobiology, 9*(6), 718–727. doi:10.1016/S0959-4388(99)00028-8

Lincoln, M., Packman, A., & Onslow, M. (2006). Altered auditory feedback and the treatment of stuttering: A review. *Journal of Fluency Disorders, 31*(2), 71–89.

Loucks, T., Chon, H., & Han, W. (2012). Audiovocal integration in adults who stutter. *International Journal of Language & Communication Disorders, 47*(4), 451–456.

Martikainen, M. H., Kaneko, K., & Hari, R. (2005). Suppressed responses to self-triggered sounds in the human auditory cortex. *Cerebral Cortex, 15*(3), 299–302. doi:10.1093/cercor/bhh131

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

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.

Max, L., Daniels, J., Curet, K., & Cronin, K. (2008). Modulation of auditory and somatosensory processing during the planning of speech movements. Paper presented at the *Proceedings of the 8th International Seminar on Speech Production*, Strasbourg, France. 41–44.

Oostenveld, R., & Praamstra, P. (2001). The five percent electrode system for high-resolution EEG and ERP measurements. *Clinical Neurophysiology, 112*(4), 713–719.

Riley, G. D. (2008). *SSI-4, Stuttering severity instrument for children and adults (4th ed.)*. Austin, TX: Pro Ed.

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

Sjölander, K. (2004). The snack sound toolkit. KTH Stockholm, Sweden.

Sommer, M., Koch, M., Paulus, W., Weiller, C., & Buchel, C. (2002). Disconnection of speech-relevant brain areas in persistent developmental stuttering. *Lancet, 360*(9330), 380–383. doi:10.1016/S0140-6736(02)09610-1

Watkins, K. E., Smith, S. M., Davis, S., & Howell, P. (2008). Structural and functional abnormalities of the motor system in developmental stuttering. *Brain, 131*, 50–59. doi:10.1093/brain/awm241

Wolpert, D. M., Diedrichsen, J., & Flanagan, J. R. (2011). Principles of sensorimotor learning. *Nature Reviews Neuroscience, 12*(12), 739–751.

Wolpert, D. M., Miall, R. C., & Kawato, M. (1998). Internal models in the cerebellum. *Trends in Cognitive Sciences, 2*(9), 338–347. doi:10.1016/S1364-6613(98)01221-2

Yates, A. J. (1963). Delayed auditory feedback. *Psychological Bulletin, 60*(3), 213–232.

CHAPTER FIVE

Conclusion and future directions

Evidence from behavioral studies and neuroimaging studies suggests that stuttering is associated with deficits in sensorimotor integration (see Max, 2004). However, it remains unknown which specific aspects of sensorimotor integration are deficient in stuttering individuals. Sensorimotor integration involves, among other aspects, neural processes related to (a) generating motor commands that will achieve a desired movement goal given the current sensory state of the system, and (b) predicting the sensory outcomes that will result when the planned motor commands are executed. Such predictions are used both to optimize the motor commands and to prepare task-relevant sensory systems for their subsequent role in processing the sensory feedback generated by the executed movements. In this dissertation, I studied different aspects of the latter role of sensory prediction by investigating modulation of the auditory system during speech movement planning in stuttering and nonstuttering adults.

In the first study, I investigated modulation of the auditory system during speech movement planning. I recorded auditory evoked potentials in response to probe tones that were presented during speech movement planning in a speaking condition and during control conditions that did not involve movement planning (silent reading and seeing nonlinguistic symbols). Auditory modulation was examined by comparing the N1 amplitude in response to probe tones presented during the speaking condition with the N1 amplitude in response to probe tones presented in the no-speaking control conditions. The nonstuttering subjects showed a statistically significant modulation of the auditory system (i.e., reduced N1 amplitude in the speaking condition). However, the stuttering subjects showed only limited, statistically nonsignificant auditory modulation. Consequently, these electrophysiological data suggest that stuttering is associated with deficiencies in modulating the auditory system during speech movement planning. Note that although the N1 component reflects activation of neural populations in cortical auditory regions, the present results cannot rule out the possibility that modulatory processes operated at lower-level structures, or at multiple levels, of the auditory system (cf. Seki, Perlmutter, & Fetz, 2003).

In the second study, I investigated whether pre-speech modulation of the auditory system in stuttering speakers is limited due to problems specifically with making auditory predictions based on a planning-related efference copy or a more general problem with making auditory predictions based on *any* available information. I recorded auditory evoked potentials in response to probe tones delivered (a) while subjects anticipated hearing their self-produced speech (speaking condition), (b) while subjects anticipated hearing a played-back version of their own pre-recorded speech (hearing condition), and (c) during a control condition without auditory input (silent reading condition). For the nonstuttering group, there was a statistically significant modulation of the N1 amplitude in both the speaking and hearing conditions vs. the silent reading condition. For the stuttering group, however, there was no statistically significant modulation of the N1 amplitude in either the speaking condition or the hearing condition. Thus, these results from the second experiment suggest that stuttering individuals have difficulties with making auditory predictions in general; that is, both during movement planning and in the absence of movement planning.

In the third study, I investigated whether stuttering speakers’ limited modulation of the auditory system during speech movement planning relates to a reduced reliance on auditory feedback during speech production. First, each participant’s reliance on auditory feedback during speech production was estimated by measuring the extent of change in speech output after being exposed to formant-shifted auditory feedback (i.e., sensorimotor adaptation). Second, auditory modulation during speech planning was examined by recording auditory evoked potentials in response to probe tones delivered (a) prior to speaking with typical auditory feedback, (b) prior to speaking with delayed auditory feedback (DAF; i.e., auditory feedback was delayed by a signal processor to render it less useful), and (c) in a no-speaking control condition (silent reading). Results from the formant-shift paradigm showed that the extent of auditory-motor adaptation was limited for stuttering subjects in comparison with nonstuttering subjects. In other words, stuttering subjects relied less on auditory feedback from previous trials to adjust their articulatory movements in future trials. Results from the auditory modulation paradigm again replicated the finding from Studies I and II that stuttering subjects show less pre-speech auditory modulation than nonstuttering subjects. Interestingly, DAF caused a reduction in auditory modulation for nonstuttering subjects but an increase in auditory modulation for stuttering subjects. Furthermore, across the two groups, there was a relationship between the amount of *change* in auditory modulation when speaking with DAF (in comparison with non-delayed feedback) and the reliance on auditory feedback during the formant-shift adaptation task: with DAF, the amount of auditory modulation tended to decrease for those participants whose speech showed relatively large changes in response to formant-shifted auditory feedback whereas it tended to increase for those participants whose speech showed relatively small changes in response to the formant shifts.

Overall, the results obtained in the third experiment suggest that auditory modulation during speech movement planning is related to the extent of reliance on auditory feedback during speech production, and that stuttering individuals show both less pre-speech auditory modulation and a reduced reliance on auditory feedback. One potential limitation of this work is that it focused only on the use of prior auditory feedback to adjust future productions. That is, to estimate the extent of reliance on auditory feedback, I used a sensorimotor adaptation paradigm. In this paradigm, participants learn to use errors in auditory feedback from previous trials to adjust their articulatory movements in future trials. Given that it has been suggested that stuttering individuals may have deficits in sensorimotor learning itself (e.g., Namasivayam & van Lieshout, 2008), and to investigate also online corrections based on real-time auditory feedback, future studies should add additional paradigms that estimate the individual participants' reliance on auditory feedback for making adjustments *within* productions with perturbed auditory feedback.

Taken together, the results from these studies suggest that the sensorimotor system uses auditory prediction to modulate auditory processing during speech planning and that this modulation varies with the extent to which a speaker relies on prior auditory feedback when generating motor commands for articulatory movements. In particular, I speculate that the sensorimotor system uses auditory prediction to prime the auditory system for its role in processing auditory feedback during speech production. From this perspective, stuttering individuals' limited use of auditory prediction to prime the auditory system prior to speech onset may results in the observed lower reliance on auditory feedback during speech production. This, in turn, may lead to unnecessary and disruptive attempts at correcting ongoing movements, and such repairs may contribute to the fluency breakdowns that form the primary symptoms of stuttering.

It is intriguing that speaking with DAF led to a decrease in auditory modulation for nonstuttering individuals (as feedback presumably became less useful) but an increase in auditory modulation for stuttering individuals. Based on the abovementioned speculation, this finding would suggest that the presence of DAF causes *less* priming of the auditory system for feedback monitoring in nonstuttering speakers but *more* priming in stuttering speakers. It is known that speaking with DAF enhances fluency of stuttering individuals and causes disfluencies in nonstuttering individuals (e.g., Lincoln, Packman, & Onslow, 2006), although the mechanisms underlying both these effects have remained unknown. For the first time, the present result provide direct electrophysiological evidence relating the effects of DAF on speech fluency to changes in modulation of the auditory system prior to speech onset. Thus, it can speculated that methods to increase pre-speech auditory modulation in stuttering individuals may also lead to increases in speech fluency. Based on this speculation, one suggestion for future clinical research is to explore novel treatments that aim at enhancing such auditory modulation in stuttering individuals. Potential approaches toward this goal could use neuro-feedback training or neuro-stimulation techniques (e.g., transcranial magnetic stimulation, TMS, or transcranial direct-current stimulation, tDCS). These techniques can be used not only to influence activation levels in selected brain regions but also to enhance functional connectivity between brain regions.

With regard to the underlying neurobiological foundation, one could speculate that the deficient pre-speech auditory modulation in stuttering individuals is linked to documented structural white matter abnormalities in pathways suggested to connect speech motor regions and auditory regions (e.g., Cykowski, Fox, Ingham, Ingham, & Robin, 2010). Future studies can test this hypothesis by examining the relationship between pre-speech auditory modulation and the strength of structural connectivity between motor and auditory regions in stuttering versus nonstuttering individuals.

Lastly, with regard to the experimental procedures that were used here to examine pre-speech auditory modulation, I recommend that future investigations address (a) the relationship between pre-speech auditory modulation and the specific speech utterance, (b) the temporal relationship between pre-speech auditory modulation and the time of onset of auditory feedback, and (c) the relationship between pre-speech auditory modulation and the presumably increasing accuracy of prediction signals during speech motor learning.

References

Cykowski, M. D., Fox, P. T., Ingham, R. J., Ingham, J. C., & Robin, D. A. (2010). A study of the reproducibility and etiology of diffusion anisotropy differences in developmental stuttering: A potential role for impaired myelination. *Neuroimage, 52*(4), 1495-1504. doi:10.1016/j.neuroimage.2010.05.011

Lincoln, M., Packman, A., & Onslow, M. (2006). Altered auditory feedback and the treatment of stuttering: A review. *Journal of Fluency Disorders, 31*(2), 71-89.

Max, L. (2004). Stuttering and internal models for sensorimotor control: A theoretical perspective to generate testable hypotheses. In B. Maassen, R. Kent, H. F. M. Peters, P. van Lieshout & W. Hulstijn (Eds.), *Speech motor control in normal and disordered speech* (pp. 357-387). Oxford, UK: Oxford University Press.

Namasivayam, A. K., & van Lieshout, P. (2008). Investigating speech motor practice and learning in people who stutter. *Journal of Fluency Disorders, 33*(1), 32-51. doi:10.1016/j.jfludis.2007.11.005

Seki, K., Perlmutter, S. I., & Fetz, E. E. (2003). Sensory input to primate spinal cord is presynaptically inhibited during voluntary movement. *Nature Neuroscience, 6*(12) doi:10.1038/nn1154