Published in final edited form as:

Brain Lang. 2015 November ; 150: 37–44. doi:10.1016/j.bandl.2015.08.008.

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

Ayoub Daliri^{a,b} and Ludo Max^{a,c}

^aUniversity of Washington, Seattle, WA, United States

^bNow at Boston University, Boston, MA, United States

^cHaskins Laboratories, New Haven, CT, United States

Abstract

We previously found that stuttering individuals do not show the typical auditory modulation observed during speech planning in nonstuttering individuals. In this follow-up study, we further elucidate this difference by investigating whether stuttering speakers' atypical auditory modulation is observed only when sensory predictions are based on 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 random probe tones delivered while anticipating either speaking aloud or hearing one's own speech played back and in a control condition without auditory input (besides probe tones). N1 amplitude of nonstuttering speakers was reduced prior to both speaking and hearing versus the control condition. Stuttering speakers, however, showed no N1 amplitude reduction in either the speaking or hearing condition as compared with control. Thus, findings suggest that stuttering speakers have general auditory prediction difficulties.

Keywords

stuttering; speech; sensorimotor integration; auditory cortex; movement planning

1. Introduction

Stuttering is associated with deficits in sensorimotor integration (for a review, see Max, 2004). To elucidate which specific aspects of sensorimotor integration are deficient in this disorder of speech fluency, we have previously examined the central nervous system's (CNS) prediction of the sensory consequences of planned limb and speech movements (Daliri & Max, 2015; Daliri, Prokopenko, Flanagan, & Max, 2014). This approach is based on theoretical models suggesting that, during movement planning, the CNS uses an

Correspondence should be addressed to Ludo Max, Ph.D., Department of Speech and Hearing Sciences, University of Washington, 1417 N.E. 42nd St., Seattle, WA 98105-6246, USA. LudoMax@uw.edu. Phone: 206 543 2674. Fax: 206 543 1093..

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

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 (Creutzfeldt, Ojemann, & Lettich, 1989; Crone et al., 2001; Mock, Foundas, & Golob, 2011; Wasaka, Hoshiyama, Nakata, Nishihira, & Kakigi, 2003).

We investigated the first of these roles with a nonspeech task in which subjects performed fast arm movements to slide an object under the fingertip toward visual targets. This task allowed us 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; Flanagan & Lolley, 2001). 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 (Daliri et al., 2014). 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 biomechanical effects of the limb and appropriately implementing compensatory adjustments from the beginning of the movement.

To investigate the second role of sensory predictions, we compared stuttering and nonstuttering speakers' modulation of the auditory system prior to the onset of speech movements (i.e., a task that does involve feedback control) (Daliri & Max, 2015). 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 (a paradigm first described in Max, Daniels, Curet, & Cronin, 2008). Results indicated that the group of nonstuttering speakers showed a statistically significant modulation of N1¹ amplitude during speech planning whereas this modulation was lacking in the group of stuttering speakers (Daliri & Max, 2015). Thus, findings suggested that stuttering individuals have difficulties with using sensory prediction mechanisms to prime the auditory system for its role in processing upcoming auditory feedback.

This result raised the question whether pre-speech modulation of the auditory system in stuttering speakers is atypical due to problems specifically with generating an efference copy of planned motor commands (i.e., inaccurate input to the prediction system) or,

¹N1 is an evoked potential negative peak with a latency of 70–130 ms relative to the onset of an auditory stimulus. The neural generators of this component are located in primary auditory cortex (Godey, Schwartz, De Graaf, Chauvel, & Liegeois-Chauvel, 2001; Näätänen & Picton, 1987). P2, referred to later in this work, is a positive peak with a latency of 150–250 ms relative to stimulus onset. Its neural generators are located more anterior in auditory cortex (Papanicolaou et al., 1990; Ross & Tremblay, 2009). Both amplitude and latency of the N1 and P2 components vary with the stimuli's physical characteristics (Small & Dimitrijevic, 2012). Whereas N1 reflects primarily processes related to stimulus detection, P2 reflects higher-level perceptual processing (Small & Dimitrijevic, 2012).

alternatively, a more general problem with *using* available information to make sensory predictions (i.e., the predictive system fails regardless of input source). To address this question, we relied here on the fact that even in the absence of movement planning or execution, the CNS builds associations among sensory events in a sequence, and then uses these associations to predict future events (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 therefore designed a new experiment that again involved, in one condition, recording AEPs evoked by probe tones delivered while stuttering and nonstuttering subjects were preparing to speak (a *delayed-speaking condition* in which the prediction of auditory input was based on speech movement planning). In a separate condition, however, we also recorded AEPs evoked by probe tones delivered while subjects were anticipating hearing a played-back version of their own pre-recorded speech (a *delayed-hearing condition* in which auditory input prediction occurred in the absence of movement planning). Given that sensory modulation can only be defined relative to a reference condition, we also recorded subjects' AEPs in response to probe tones delivered during silent reading (a *silent reading condition* that involved no prediction 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 (i.e., in extent of within-subject changes in AEP amplitude in an experimental condition vs. the reference condition) would be limited to the delayed-speaking condition, whereas (b) if stuttering is associated with a more general auditory prediction deficit, then a difference in auditory modulation between the two groups can be expected in both the delayed-speaking condition and the delayed-hearing condition.

2. Methods

2.1. Participants

Ten 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, $SD = 9.383$) provided written consent 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 the octave frequencies 250–8000 Hz). To determine stuttering participants' 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 of the stuttering participants had also participated in our previous study on pre-speech auditory modulation (Daliri & Max, 2015).

2.2.Procedure and Instrumentation

Participants were seated approximately 1.5 m from a 23-inch monitor inside a sound-attenuated room. A microphone (WL185, Shure Incorporated, Niles, IL) 15 cm from the participant's mouth transduced the speech signal. After amplification (microphone amplifier: DPS II, ART ProAudio, Niagara Falls, NY; headphone amplifier: S-phone, Samson Technologies Corp., Syosset, NY), this speech signal was 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, 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 8 blocks was randomized for each participant. In other words, with the exception of the first block always being a block from the *delayed-speaking* condition, the blocks from the various conditions were mixed in randomized order.

Each trial started with the presentation of a word in white characters on a black background (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, 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 (i.e., the color change from white to green characters constituted the *go*-signal). 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 of the characters on the screen) as produced in the first block of the session (i.e., the block that was always part of the *delayed-speaking* condition). In the *silent reading* condition (Fig. 1C), participants read the word silently without making any movements and without hearing a word played back.

On 40% of the trials in each block (referred to as tone trials), probe tones were delivered 400 ms after presentation of the word in white characters (Fig. 1D). The probe tones were 1 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 (referred to as no-tone trials).

2.3. Electroencephalographic Recordings

Electroencephalographic (EEG), electromyographic (EMG), and electrooculographic (EOG) activity was recorded continually 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 10-20 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 below, Data Analysis). EMG signals were recorded from four electrodes placed on the skin overlying upper and lower lip muscles (i.e., region of the 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 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 or 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.

2.4. 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. The EEG signals were low-pass filtered off-line (cut-off frequency 50 Hz), and these 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 activity (as a result of the probe tones) as well as non-auditory activity (related to motor, linguistic, cognitive, and visual processes), whereas the averaged response for the no-tone trials reflected only the non-auditory activity. As the experiment's goal was to investigate modulation within the auditory system, 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 in order to isolate, as much as possible, auditory activity evoked by the probe tones (Fig. 1E) (see Luck, 2014, for discussion of evoked potential difference waveforms). 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, 2015; Luck, 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 [equivalent to C3], D20, D27, D28), the central region (Central ROI: electrodes A1 [equivalent to Cz], A2, B1, C1, D1, D15), and the right hemisphere (Right ROI: electrodes B17, B18, B22 [equivalent to C4], B23, B30, B31). The three ROIs, and the individual electrodes included in each ROI, are illustrated in the top row of Fig. 2. The ROI-based AEPs were used to extract 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.

The rationale for including three separate ROIs was based directly on the results from our previous study on pre-speech auditory modulation: a ROI-by-condition interaction indicated that, across all participants, the effect was larger at electrodes located over the central region and over the left hemisphere than over the right hemisphere (Daliri & Max, 2015). We therefore decided to define the same ROIs for the present study as this allows an examination of the consistency of such ROI effects, if any, on the auditory modulation phenomenon. It is important to note, however, that there is no simple and straightforward relationship between the intracranial location of neural generators of the EEG signal and the scalp location of surface electrodes where maximally strong signals (or maximum signal amplitude changes across conditions) are detected (see, among many other sources, Michel et al., 2004). That is, the detection of a maximally strong signal (or a maximum extent of change in a signal) at specific electrode locations cannot be interpreted as indicating that the neural sources of the detected activity are located in brain areas directly underneath these electrodes.

2.5. Statistical analyses

Given that between-subjects comparisons of EEG data reflect not only neuronal activation differences between the groups but also confounding by normal inter-subject variation in properties of the human head volume conductor (Nunez & Srinivasan, 2006), it is generally recommended that EEG studies use experimental designs that test the hypotheses of interest by means of within-subject comparisons and group-by-condition interactions (Coulson, 2007; Luck, 2014). Our study follows this same line of reasoning as pre-speech auditory modulation is defined as a within-subjects effect across conditions (by definition, modulation requires a reference point, here the *silent-reading* control condition), and the primary hypotheses of the study (i.e., whether the extent of such modulation effects differs between the stuttering and nonstuttering groups) is tested by means of group-by-condition interaction analyses.

Hence, N1 and P2 amplitude and latency were used as the dependent variables in 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, the degrees of freedom for within-subject effects 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).

For the stuttering group, Pearson correlation coefficients were also calculated to investigate possible relationships 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 magnitude of modulation in N1 or P2 amplitude.

3. 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.

3.1.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 the Condition $\times$ 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 a larger N1 amplitude in the central ROI than in both the left ROI, $t(19) = -8.502, p < .001$, and the right ROI, $t(19) = -9.983, p < .001$. All other two-way and three-way interactions (ROI $\times$ Group, Condition $\times$ ROI, and Condition $\times$ ROI $\times$ Group) were not statically significant ($p > .106$ in all cases). Thus, unlike the situation in our prior study (Daliri & Max, 2015), the overall auditory modulation effect across all subjects did not differ among the three ROIs.

Most relevant for the purpose of the present study is the statistically significant Condition $\times$ 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*.

In light of the observed Condition $\times$ Group interaction, we also calculated for each subject the *amount* of N1 modulation (i.e., reduction in N1 amplitude) for the *delayed-speaking* and *delayed-hearing* conditions relative to the *silent reading* control condition. The boxplots in

Fig. 3B illustrate the stuttering and nonstuttering individual subject distributions for these data.

3.2.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 $\times$ Group, ROI $\times$ Group, Condition $\times$ ROI, Condition $\times$ ROI $\times$ Group) were statistically significant ($p > .323$ in all cases).

3.3.P2 amplitude

P2 amplitude showed significant main effects of Condition, $F(2, 36) = 4.496$, $p = .018$, and ROI, $F(2, 36) = 31.217$, $p < .001$, and—in contrast to the N1 amplitude data—these effects did interact, $F(4, 72) = 4.264$, $p = .004$. In the central ROI, P2 amplitude was smaller in the *delayed-speaking* condition than in both the *silent reading* condition, $t(19) = 6.396$, $p < .001$, and in 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 $\times$ Group, ROI $\times$ Group, and Condition $\times$ ROI $\times$ Group) were not statistically significant ($p > .115$). Fig. 3C shows the average P2 amplitude in all conditions for both the stuttering and nonstuttering groups. The boxplots in Fig. 3D show the individual subject distributions for amount of P2 amplitude modulation during the *delayed-speaking* and *delayed-hearing* conditions relative to the *silent reading* condition.

3.4.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 $\times$ Group interaction, $F(2, 36) = 4.073$, $p = .025$: the nonstuttering group's P2 latency did not differ between any pairs of ROIs ($p > .170$) whereas the stuttering group's P2 latency was shorter in the central ROI than in the left ROI, $t(9) = 2.834$, $p = .020$, but not in the right ROI ($p = .070$). None of the other two-way and three-way interactions (Condition $\times$ Group, Condition $\times$ ROI, and Condition $\times$ ROI $\times$ Group) were statistically significant ($p > .370$).

3.5.Correlational analyses

There were no statistically significant correlations between stuttering frequency and N1 amplitude modulation in either the *delayed-speaking* condition ($r = -.540$, $p = .107$) or the *delayed-hearing* condition ($r = -.432$, $p = .212$). There was, however, a statistically significant positive correlation between stuttering frequency and P2 amplitude modulation in the *delayed-speaking* condition, $r = .800$, $p = .005$, and a near-significant correlation between these same variables in the *delayed-listening* condition, $r = .621$, $p = .055$.

4. Discussion

Previously, we reported that the typically observed modulation of auditory processing during speech movement planning (Max et al., 2008) is limited in adults who stutter as compared with adults who do not stutter (Daliri & Max, 2015). This finding (and a compatible result for the somatosensory domain in McClean, 1996) suggested 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 the underlying 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 affecting the use of *any* available information to generate predictions about upcoming afferent inputs regardless of whether those inputs are self-generated or externally generated.

To address this question, we recorded AEPs of stuttering and nonstuttering adults in response to unpredictable probe tones delivered while anticipating either speaking (*delayed-speaking* condition) or hearing one's own speech played back (*delayed-hearing* condition) versus a reference condition without predictable auditory input (*silent reading* condition). The *delayed-speaking* condition involved auditory predictions based on the speaker's own movement planning (in each trial, subjects knew that they would hear their own auditory feedback); the *delayed-hearing* condition involved auditory predictions in the absence of movement planning (in each trial, subjects knew that they would hear their own pre-recorded production of the displayed word); and the *silent reading* condition involved no predictions of auditory input.

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 based on external events), the stuttering group failed to show such 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 was observed only in the central ROI—was similar for the stuttering and nonstuttering groups. This absence of a between-group difference in P2 modulation is consistent with our previous results (Daliri & Max, 2015).

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 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 displayed word would be played back to them. In the *delayed-speaking* condition – and during speech production in general – subjects' own planning of movement sequences allows them to predict when the resulting auditory consequences will be heard. Hence, the overall results suggest that stuttering adults' previously demonstrated failure to modulate the auditory system prior to speech onset (Daliri & Max, 2015) 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. In other words, during speech production, stuttering speakers appear to have 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 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, for discussion of feedback contributions in speech production) leads to unnecessary, and disruptive, attempts at correcting ongoing movements. Such “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 only at time points close to movement onset and not in the earlier planning stages (Daliri & Max, 2015). All probe tones were presented 400 ms after initial presentation of the target word and 200 ms before the visual change indicating that the word had to be spoken aloud 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, 2015), 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. 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 individuals who stutter show abnormal structural and functional connectivity in areas involving motor/premotor and auditory regions (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).

Further studies are also needed to clarify the obtained correlational results. For N1 amplitude modulation (which differed between stuttering and nonstuttering groups in both the present study and our prior Daliri & Max, 2015, study), the stuttering subjects' amount of modulation in the *delayed-speaking* or *delayed-hearing* conditions was not statistically significantly correlated with a clinical measure of stuttering frequency. Unexpectedly, for P2 amplitude modulation (which did not differ between the stuttering and nonstuttering groups in either the present study or the prior study), stuttering subjects' amount of modulation in the *delayed-speaking* condition did correlate significantly with stuttering frequency ($r = .800, p = .005$), and this correlation was also relatively high in the *delayed-listening* condition ($r = .621, p = .055$). As we have pointed out elsewhere (e.g., Daliri et al., 2014), interpreting such correlations is not necessarily straightforward because the relationship between physiological variables (here AEP component amplitudes) measured in the context of fluent single-word productions and a count of stuttering moments during oral reading and conversation is likely affected by a wide range of uncontrolled variables. Nevertheless, these results may yield new hypotheses that can be tested in future studies. For example, based on the data available to date, it may be speculated that stuttering subjects' modulator signals to auditory cortex neurons that generate the N1 response reach their targets late within the movement planning phase, essentially arriving after the time when N1 has peaked (resulting in little or no N1 modulation and random variation in this modulation relative to stuttering frequency) and adding to a separate, intact modulation of P2 (resulting in P2 modulation that is subtly, but not necessarily statistically, greater in stuttering vs. nonstuttering subjects, and that is enhanced the most in cases of severe stuttering where the modulator signals may be the most delayed). Thus, important new insights may be gained from studies that investigate stuttering and nonstuttering subjects' *time course* of auditory modulation during speech movement planning.

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 condition with no auditory input. Nonstuttering speakers modulated N1 amplitude prior to both speaking (i.e., auditory predictions based on movement planning) and hearing one's own speech played back (i.e., auditory predictions based on contextual cues in the absence of movement planning) as compared with silent reading (i.e., no auditory predictions). Stuttering speakers showed no N1 modulation in either of the two conditions with predictable auditory input. 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. Selective suppression of self-initiated sounds in an auditory stream: An ERP study. *Psychophysiology*. 2011; 48(9):1276–1283. doi:10.1111/j.1469-8986.2011.01196.x. [PubMed: 21449953]
- Baess P, Jacobsen T, Schroger E. 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*. 2008; 70(2):137–143. [PubMed: 18627782]
- Bubic A, von Cramon DY, Schubotz RI. Prediction, cognition and the brain. *Frontiers in Human Neuroscience*. 2010; 4:25. doi:10.3389/fnhum.2010.00025. [PubMed: 20631856]
- Chang S, Erickson KI, Ambrose NG, Hasegawa-Johnson MA, Ludlow CL. Brain anatomy differences in childhood stuttering. *Neuroimage*. 2008; 39(3):1333–1344. doi:10.1016/j.neuroimage.2007.09.067. [PubMed: 18023366]
- Chang S, Horwitz B, Ostuni J, Reynolds R, Ludlow CL. Evidence of left inferior frontal-premotor structural and functional connectivity deficits in adults who stutter. *Cerebral Cortex*. 2011; 21(11):2507–2518. doi:10.1093/cercor/bhr028. [PubMed: 21471556]
- Chang S, Kenney MK, Loucks TMJ, Ludlow CL. Brain activation abnormalities during speech and non-speech in stuttering speakers. *Neuroimage*. 2009; 46(1):201–212. doi:10.1016/j.neuroimage.2009.01.066. [PubMed: 19401143]
- Cornelisse LE, Gagné J, Seewald RC. Ear level recordings of the long-term average spectrum of speech. *Ear and Hearing*. 1991; 12(1):47–54. [PubMed: 2026288]
- Creutzfeldt O, Ojemann G, Lettich E. Neuronal-activity in the human lateral temporal-lobe .2. responses to the subjects own voice. *Experimental Brain Research*. 1989; 77(3):476–489. doi:10.1007/BF00249601. [PubMed: 2806442]
- Crone NE, Hao L, Hart J, Boatman D, Lesser RP, Irizarry R, Gordon B. Electrographic gamma activity during word production in spoken and sign language. *Neurology*. 2001; 57(11):2045–2053. [PubMed: 11739824]
- Cykowski MD, Kochunov PV, Ingham RJ, Ingham JC, Mangin J, Riviere D, Fox PT. Perisylvian sulcal morphology and cerebral asymmetry patterns in adults who stutter. *Cerebral Cortex*. 2008; 18(3):571–583. doi:10.1093/cercor/bhm093. [PubMed: 17584852]
- Daliri A, Max L. Modulation of auditory processing during speech movement planning is limited in adults who stutter. *Brain and Language*. 2015; 143:59–68. [PubMed: 25796060]
- Daliri A, Prokopenko R, Flanagan JR, Max L. Control and prediction components of movement planning in stuttering vs. nonstuttering adults. *Journal of Speech, Language, and Hearing Research*. 2014; 57(6):2131–2141. doi:10.1044/2014_JSLHR-S-13-0333.
- Delorme A, Makeig S. EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*. 2004; 134(1):9–21. [PubMed: 15102499]
- Flanagan JR, Lolley S. The inertial anisotropy of the arm is accurately predicted during movement planning. *The Journal of Neuroscience*. 2001; 21(4):1361–1369. [PubMed: 11160407]
- Gilmore RL. American-electroencephalographic-society guidelines in electroencephalography, evoked-potentials, and polysomnography. *Journal of Clinical Neurophysiology*. 1994; 11(1):1–142. [PubMed: 8195412]

- Godey B, Schwartz D, De Graaf J, Chauvel P, Liegeois-Chauvel C. Neuromagnetic source localization of auditory evoked fields and intracerebral evoked potentials: A comparison of data in the same patients. *Clinical Neurophysiology*. 2001; 112(10):1850–1859. [PubMed: 11595143]
- Hughes G, Desantis A, Waszak F. Mechanisms of intentional binding and sensory attenuation: The role of temporal prediction, temporal control, identity prediction, and motor prediction. *Psychological Bulletin*. 2013; 139(1):133–151. [PubMed: 22612280]
- Kell CA, Neumann K, von Kriegstein K, Posenenske C, von Gudenberg AW, Euler H, Giraud AL. How the brain repairs stuttering. *Brain*. 2009; 132(10):2747–2760. doi:10.1093/brain/awp185. [PubMed: 19710179]
- Kim KS, Max L. Estimating feedforward vs. feedback control of speech production through kinematic analyses of unperturbed articulatory movements. *Frontiers in Human Neuroscience*. 2014; 8:911. [PubMed: 25426056]
- Luck, SJ. An introduction to the event-related potential technique. 2nd ed.. The MIT Press; Cambridge, MA: 2014.
- Martikainen MH, Kaneko K, Hari R. Suppressed responses to self-triggered sounds in the human auditory cortex. *Cerebral Cortex*. 2005; 15(3):299–302. doi:10.1093/cercor/bhh131. [PubMed: 15238430]
- Max, L. Stuttering and internal models for sensorimotor control: A theoretical perspective to generate testable hypotheses. In: Maassen, B.; Kent, R.; Peters, HFM.; van Lieshout, P.; Hulstijn, W., editors. *Speech motor control in normal and disordered speech*. Oxford University Press; Oxford, UK: 2004. p. 357-387.
- Max L, Onghena P. 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*. 1999; 42(2):261–270.
- Max, L.; Daniels, J.; Curet, K.; Cronin, K. 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. 2008; p. 41-44.
- McClean MD. Lip-muscle reflexes during speech movement preparation in stutterers. *Journal of Fluency Disorders*. 1996; 21(1):49–60.
- Michel CM, Murray MM, Lantz G, Gonzalez S, Spinelli L, de Peralta RG. EEG source imaging. *Clinical neurophysiology*. 2004; 115(10):2195–2222. [PubMed: 15351361]
- Mock JR, Foundas AL, Golob EJ. Modulation of sensory and motor cortex activity during speech preparation. *European Journal of Neuroscience*. 2011; 33(5):1001–1011. doi:10.1111/j.1460-9568.2010.07585.x. [PubMed: 21261757]
- Näätänen R, Picton T. The N1 wave of the human electric and magnetic response to sound: A review and an analysis of the component structure. *Psychophysiology*. 1987; 24(4):375–425. [PubMed: 3615753]
- Nunez, PL.; Srinivasan, R. *Electric fields of the brain: the neurophysics of EEG*. 2nd ed.. Oxford University Press; New York, NY: 2006.
- Oostenveld R, Praamstra P. The five percent electrode system for high-resolution EEG and ERP measurements. *Clinical Neurophysiology*. 2001; 112(4):713–719. [PubMed: 11275545]
- O'Reilly JX, Mesulam MM, Nobre AC. The cerebellum predicts the timing of perceptual events. *The Journal of Neuroscience*. 2008; 28(9):2252–2260. doi:10.1523/JNEUROSCI.2742-07.2008. [PubMed: 18305258]
- Papanicolaou AC, Baumann S, Rogers RL, Saydjari C, Amparo EG, Eisenberg HM. Localization of auditory response sources using magnetoencephalography and magnetic resonance imaging. *Archives of Neurology*. 1990; 47(1):33–37. [PubMed: 2294891]
- Riley, GD. *SSI-4, Stuttering severity instrument for children and adults*. 4th ed.. Pro Ed; Austin, TX: 2008.
- Ross B, Tremblay K. Stimulus experience modifies auditory neuromagnetic responses in young and older listeners. *Hearing Research*. 2009; 248(1):48–59. [PubMed: 19110047]
- Schubotz RI. Prediction of external events with our motor system: Towards a new framework. *Trends in Cognitive Sciences*. 2007; 11(5):211–218. [PubMed: 17383218]

- Schubotz RI, Von Cramon DY. Functional–anatomical concepts of human premotor cortex: Evidence from fMRI and PET studies. *NeuroImage*. 2003; 20:120–131.
- Shadmehr R, Smith MA, Krakauer JW. Error correction, sensory prediction, and adaptation in motor control. *Annual Review of Neuroscience*. 2010; 33:89–108. doi:10.1146/annurev-neuro-060909-153135.
- Small, S.; Dimitrijevic, A. Physiological Mechanisms Underlying ASSRs. In: Tremblay, K.; Burkard, R., editors. *Translational Perspectives in Auditory Neuroscience Normal Aspects of Hearing*. Plural Publishing; San Diego, CA: 2012. p. 201-242.
- Sommer M, Koch M, Paulus W, Weiller C, Buchel C. Disconnection of speech-relevant brain areas in persistent developmental stuttering. *Lancet*. 2002; 360(9330):380–383. doi:10.1016/S0140-6736(02)09610-1. [PubMed: 12241779]
- Sperry RW. Neural basis of the spontaneous optokinetic response produced by visual inversion. *Journal of Comparative and Physiological Psychology*. 1950; 43(6):482–489. doi:10.1037/h0055479. [PubMed: 14794830]
- Wasaka T, Hoshiyama M, Nakata H, Nishihira Y, Kakigi R. Gating of somatosensory evoked magnetic fields during the preparatory period of self-initiated finger movement. *Neuroimage*. 2003; 20(3):1830–1838. doi:10.1016/S1053-8119(03)00442-7. [PubMed: 14642492]
- Waszak F, Cardoso-Leite P, Hughes G. Action effect anticipation: Neurophysiological basis and functional consequences. *Neuroscience and Biobehavioral Reviews*. 2012; 36(2):943–959. doi: 10.1016/j.neubiorev.2011.11.004. [PubMed: 22108008]
- Watkins KE, Smith SM, Davis S, Howell P. Structural and functional abnormalities of the motor system in developmental stuttering. *Brain*. 2008; 131:50–59. doi:10.1093/brain/awm241. [PubMed: 17928317]
- Wolfensteller U, Schubotz RI, Von Cramon DY. Understanding non-biological dynamics with your own premotor system. *NeuroImage*. 2007; 36:33–43.
- Wolpert DM, Diedrichsen J, Flanagan JR. Principles of sensorimotor learning. *Nature Reviews Neuroscience*. 2011; 12(12):739–751. [PubMed: 22033537]
- Wolpert DM, Miall RC, Kawato M. Internal models in the cerebellum. *Trends in Cognitive Sciences*. 1998; 2(9):338–347. doi:10.1016/S1364-6613(98)01221-2. [PubMed: 21227230]

Highlights

1. Fluent speakers' N1 auditory response to pure tones was reduced prior to both speaking and hearing one's own speech.
2. Stuttering speakers showed no auditory N1 modulation in either the speaking or hearing condition.
3. Findings suggest that stuttering speakers have a general auditory prediction deficit.

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 presented in 40% of trials (tone trials) of each condition (D). No auditory probe was presented in the remaining trials (no-tone trials). For each subject and in each condition, the average evoked potential for no-tone trials was subtracted from the average evoked potential for tone trials (E). This procedure provided the best estimate of the auditory cortex's response to the probe tones.

Fig. 2.

Top row: Three regions of interests (ROIs), each consisting of six electrodes (AEPs from the six electrodes in each ROI were averaged). Middle row: 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) conditions. Bottom row: Grand average AEPs for the stuttering group in the same conditions as above for the nonstuttering group. AEPs of the nonstuttering group show a distinct attenuation of N1 amplitude in the *delayed-speaking* and *delayed-hearing* conditions relative to the *silent reading* control condition.

Fig. 3. Group average N1 amplitude (A) and P2 amplitude (C) for stuttering and nonstuttering groups during the *delayed-speaking* condition (green), *delayed-hearing* condition (orange), and *silent reading* condition (black). Subject distribution with regard to the amount of amplitude modulation in the *delayed-speaking* and *delayed-hearing* conditions vs. the *silent reading* condition for N1 and P2 are shown as boxplots (B and D, respectively; vertical axes are adjusted so that greater reductions in amplitude are toward the top of the plots). There was a statistically significant Group $\times$ Condition interaction for N1 amplitude, indicating significant modulation in *delayed-speaking* and *delayed-hearing* conditions for the nonstuttering group but not the stuttering group. The Group $\times$ Condition interaction for P2 was not statistically significant. All data averaged across three ROIs. Error bars indicate standard errors. Asterisks indicate $p < .001$.

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). The participants are rank ordered by stuttering frequency. SSI scores 10–17 correspond to very mild, 18–24 mild, 25–31 moderate, 32–36 severe, 37–46 very severe.

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