Published in final edited form as:

Brain Lang. 2010 November ; 115(2): 141–147. doi:10.1016/j.bandl.2010.07.007.

Spatiotemporal dynamics of speech sound perception in chronic developmental stuttering

Mario Liotti, M.D., Ph.D.^{1,*}, Janis C. Ingham, Ph.D.^{2,3}, Osamu Takai, B.A.¹, Delia Kothmann, Ph.D.⁵, Ricardo Perez, M.S.⁴, and Roger J. Ingham, Ph.D.^{2,3}

¹Department of Psychology, Simon Fraser University, Burnaby, BC

²Department of Speech and Hearing Sciences, University of California, Santa Barbara

³Research Imaging Institute, University of Texas Health Science Center, San Antonio, TX

⁴Department of Radiology, University of Texas Health Science Center at San Antonio, TX

⁵Department of Psychology, St. Edward's University, Austin, TX.

Abstract

High-density ERPs were recorded in eight adults with persistent developmental stuttering (PERS) and eight matched normally fluent (CONT) control volunteers while participants either repeatedly uttered the vowel 'ah' or listened to their own previously recorded vocalizations.

The frontocentral N1 auditory wave was reduced in response to spoken vowels relative to heard vowels (auditory-vocal gating), but no difference in the extent of such modulation was found in the PERS group. Abnormalities in the PERS group were restricted to the LISTEN condition, in the form of early N1 and late N3 amplitude changes. Voltage of the N1 wave was significantly reduced over right inferior temporo-occipital scalp in the PERS group. A laterality index derived from N1 voltage moderately correlated with the PERS group's assessed pre-experiment stuttering frequency. Source localization with sLORETA (Pascual-Marqui, 2002) revealed that at the peak of the N1 the PERS group displayed significantly greater current density in right primary motor cortex than the CONT group, suggesting abnormal early speech motor activation. Finally, the late N3 wave was reduced in amplitude over inferior temporo-occipital scalp, more so over the right hemisphere. SLORETA revealed that in the time window of the N3 the PERS group showed significantly less current density in right secondary auditory cortex than the CONT group, suggesting abnormal speech sound perception. These results point to a deficit in auditory processing of speech sounds in persistent developmental stuttering, stemming from early increased activation of right rolandic area and late reduced activation in right auditory cortex.

© 2010 Elsevier Inc. All rights reserved.

*Corresponding author. Address: Dr. Mario Liotti, Department of Psychology, Simon Fraser University, 8888 University Drive, Burnaby, V5A 1S6, Canada voice: (778) 782-4561 fax: (778) 782-3427 mliotti@sfu.ca.

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.

Keywords

chronic developmental stuttering; ERPs; speech perception; primary motor cortex; cortex; sLORETA; auditory-vocal gating

INTRODUCTION

Adults with persistent developmental stuttering (PERS) have been shown to display both structural (Cykowski, et al., 2008; Sommer, Koch, Paulus, Weiller, & Büchel, 2002) and functional brain abnormalities during speech tasks when compared to fluent control speakers (CONT) (Brown, et al., 2005). Functional neuroimaging studies during fluent and nonfluent speech tasks have revealed increased right hemisphere or bilateral activation of speech motor areas and cerebellum (DeNil, Kroll, Kapur, & Houle, 2000; Braun, et al., 1997; Fox, et al., 1996; Fox, et al., 2000; Ingham, Fox, Ingham & Zamarripa, 2000; Ingham, et al., 2004) and reduced activation in auditory areas (Braun, et al., 1997; Fox, et al., 1996; Ingham, et al., 2000). It has been proposed that persistent stuttering may be characterized by an aberrant interplay between motor/premotor and auditory neural regions, particularly in the right hemisphere (Ingham, 2001).

Functional neuroimaging studies based on the haemodynamic response have been essential in identifying the cortical regions associated with stuttering, but are limited by their coarse temporal resolution. In contrast, electroencephalography (EEG) and magnetoecephalography (MEG) possess exquisite time resolution (1 msec) and make it possible to characterize the timing, order of activation, and dynamic orchestration of brain regions during normal and abnormal speech processing in fluent and nonstuttering speakers. Relatively few electromagnetic studies have been conducted on auditory processing in adult who stutter. Using the EEG technique, Finitzo, Pool, Freeman, Devous and Watson (1991) measured the event-related potentials (ERPs) to the onset of pure tones presented to PERS and CONT speakers. They reported reduced amplitudes of the early N1 and P2 waves in the PERS group. In a second EEG study, Hampton and Weber-Fox (2008) recorded ERPs during the listening of standard and target tones using an oddball paradigm with adult PERS and CONT speakers. As a group, PERS performed within the normal range, but a small subset of PERS speakers showed reduced N1 and P2 waves. Contrary findings, however, were obtained in a MEG study by Biermann-Ruben, Salmelin and Schnitzler (2005), who found no differences in early processing of listened pure tones between adult PERS and CONT speakers. Several studies have explored the possibility that processing abnormalities in PERS speakers may emerge at a *later* stage of planning or execution of an overt speech-motor response. Salmelin et al. (1998) recorded MEG while PERS and CONT speakers were presented with auditory tones while engaged in overt or silent reading. Slight differences in N1 responses were found for the PERS group, with the greatest difference for left hemisphere/ left ear tones, and no difference for the equivalent right hemisphere/right ear responses, suggesting an altered inter-hemispheric balance in the PERS group.

A later MEG study by Salmelin, et al. (2000) investigated the possibility that *linguistic* content of auditory stimuli may be required to show consistent processing abnormalities in

adult PERS speakers. The experimenters presented visual words necessitating a delayed overt reading response. PERS speakers showed an abnormal order of activation of speech planning and execution areas, with the left rolandic area active *before* the left frontal operculum, a sequence opposite to that observed in fluent subjects.

More recently Biermann-Ruben, et al. (2005) hypothesized that the complexity of linguistic processing may also contribute to processing abnormalities in adult PERS speakers. While passive listening of pure tones produced no differences in activation patterns between groups, abnormalities emerged when presentation of speech stimuli (single words or short sentences) required an overt speech response. The PERS group showed a unique activation of the right rolandic region 300-1000 ms from stimulus onset (more pronounced for sentences than words), as well as an *early* activation of the left frontal operculum in the sentence task only. It was concluded that PERS speakers abnormally activate speech planning and execution areas during speech perception in anticipation of an overt speech response, with increasing linguistic complexity resulting in greater abnormalities. It should be noted that both MEG studies mentioned above did not include conditions in which linguistic stimuli were presented, yet an overt speech response was *not* required. Therefore, neither study excludes the possibility that the same abnormalities may also be elicited while *simply listening* to speech stimuli.

An important foundation for the present study is a MEG study of auditory speech perception in healthy volunteers showing reduced magnetic N1 amplitude (and reduced activity in primary auditory cortex) in response to uttered syllables, relative to when the same syllables were replayed to them (auditory-vocal gating, Curio et al., 2000). Similar effects were reported by Salmelin et al. (1998), with attenuated N1 magnetic responses to auditory tones during overt relative to silent reading. Noteworthy is the fact that the study employed auditory tones, leaving untested the possibility that speech sounds may give rise to different results.

Among adults who stutter there is compelling evidence that speech self-monitoring may be abnormally deficient during stuttered speech (see Ingham, 2001b; Ingham et al., 2004). The procedure used by Curio et al. may be especially useful for investigating this issue because it makes it possible to test whether chronic stuttering is associated with exaggerated auditory-vocal gating to speech sounds.

Aims of the present study were to test whether adults with persistent stuttering 1) Would display exaggerated auditory-vocal gating (suppression) of the auditory N1 response to spoken versus heard sounds; (2) Would exhibit abnormalities in speech *perception* while passively *listening* to simple speech sounds. It was predicted that linguistic complexity would not be a necessary precondition for the appearance of processing abnormalities; (3) Would display spatio-temporal abnormalities in the electrical response to speech sounds that can be source-localized to motor cortex and auditory cortex.

To address these aims, high density auditory ERPs to voice onset were recorded while adult PERS and CONT subjects either uttered (SPEAK condition) or passively listened to (LISTEN condition) simple speech sounds with very limited linguistic content (the vowel

‘ah’), and sLORETA (Pascual-Marqui, 2002) was employed to localize the brain source(s) of putative group differences.

RESULTS

Early N1 window (20-80 msec)

Fronto-central N1 analysis—There was a main effect of Condition, $F(1,14)=8.44$, $p=0.012$, with N1 amplitude being markedly reduced in the SPEAK relative to the LISTEN condition (Listen: $-0.60 \pm 0.62 \mu\text{V}$; Speak: $-0.28 \pm 0.40 \mu\text{V}$). Critically, there was no hint of a significant interaction of Condition and Group, $F(1,14)=0.52$, $p=.49$, ns, nor of a main effect of Group, $F(1,14)=0.22$, $p=.88$, ns. N1 voltage reduction for SPEAK relative to LISTEN sounds (audiovocal gating) was comparable in the PERS and CONT groups. See an illustration of such results in Fig. 1, panels B-C.

Posterior N1 analysis—This analysis revealed a Hemisphere main effect, $F(1,14)=5.11$, $p=.04$, with greater N1 voltage for the left than right hemisphere (LH: $0.64 \pm 0.70 \mu\text{V}$; RH: $0.53 \pm 0.78 \mu\text{V}$). Factor condition and the Condition $\times$ Hemisphere interactions were not significant (for both, $p>0.1$), and the Group main effect was far from significance, $F(1,14)=0.12$.

However, there was a highly significant Condition $\times$ Hemisphere $\times$ Group interaction, $F(1,14)=9.37$, $\text{MSE}=3.51$, $p=0.008$. In order to elucidate this interaction, ANOVAs were carried out for each group separately.

In the PERS group, there was a main effect of hemisphere, $F(1,7)=6.0$, $p=.044$, with smaller N1 voltage over the *right* than left hemisphere (LH: $0.61 \pm 0.75 \mu\text{V}$; RH: $0.46 \pm 0.81 \mu\text{V}$), which was qualified by the significant Condition $\times$ Hemisphere interaction, $F(1,14)=8.85$, $\text{MSE}=.234$, $p=.021$, effect size ω^2 partial = .56. Bonferroni corrected ($p<.025$) post-hoc contrasts revealed a significant hemispheric difference for the Listen condition, $t(1,7)=4.70$, $p=.004$ (LH: $0.80 \pm 0.86 \mu\text{V}$; RH: $0.48 \pm 0.78 \mu\text{V}$). This effect was entirely absent in the Speak condition, $t(1,7)=-.22$, n.s. (LH: $0.42 \pm 0.64 \mu\text{V}$; RH: $0.44 \pm 0.84 \mu\text{V}$).

In contrast, in the Control group, there were no main effects of hemisphere, $F(1,7)=0.73$, n.s., nor Condition, $F(1,7)=2.36$, $p=0.17$, and no significant Condition $\times$ Hemisphere interaction, $F(1,7)=1.27$, $p>.25$, n.s. In particular, in the Listen condition, N1 voltage was identical over left ($.88 \pm .93 \mu\text{V}$) and right hemispheres ($.87 \pm .96 \mu\text{V}$) (see Fig. 1, panels A and C).

A laterality index based on the mean amplitude of the N1 wave at temporo-occipital sites (calculated as $[(R-L)/(R+L)] \times 100$) was negatively correlated, in the PERS group, with a measure of speech performance (% syllables stuttered) obtained during reading and speaking tasks in a preliminary session ($r = -.72$, $p=.025$, one tailed). The greater the severity of stuttering, the more negative the laterality index.

sLORETA N1 Source Analysis (50-60 ms)—The mean current density power of the PERS group was significantly greater than that for the CONT group in a region in right

primary motor cortex centered in BA 4 ($x = +30$, $y = -20$, $z = +50$), ($t = 5.65$, $p < 0.05$, two-tailed, corrected for whole-brain comparisons). In contrast, no clusters of significantly greater current density in CONT group relative to the PERS group were found in auditory cortex or elsewhere in the brain within this time window (maximum t value in auditory cortex BA22 = 1.8, ns). These results are illustrated in Figure 1, panel F.

Late N3 Window-LISTEN condition (225-375 ms)

The frontocentral N3 analysis revealed that the main effect of Group was far from significant, $F(1,14) = 1.77$, $p = .20$. In contrast, the posterior N3 analysis revealed a significant main effect of Hemisphere, $F(1,14) = 11.45$, $MSE = .064$, $p < .004$, due to greater amplitude over the left hemisphere, a lack of Group main effect, $F(1,14) = 2.06$, $MSE = .34$, $p = .17$, and a non-significant Group $\times$ Hemisphere interaction, $F(1,14) = 0.88$, $p = .37$. However, using the a-priori hypothesis of a right hemisphere reduction in the PERS group only, we carried out separate paired t -tests on each group (Bonferroni correction: $p < .025$). The hemisphere contrast was significant only in the PERS group, $t(7) = 3.79$, $p = .007$ (LH = $.44 \pm .12 \mu V$; RH = $.05 \pm .09 \mu V$); the CONT group result being $t(7) = 1.49$, $p = .18$ (LH = $.65 \pm .16 \mu V$; RH = $.43 \pm .22 \mu V$). These effects are illustrated in Figure 1, Panels D-E. Furthermore, a laterality score measured with $(R-L)/(R+L) \times 100$ was significantly different between groups, $t(14) = 2.48$, $p = .027$ (PERS = -246.5 ± 90 ; CONT = -20.3 ± 18.2), being more negative (reduced R N3 relative to L N3) in the PERS group.

sLORETA N3 Source analysis—In the time window of the N3 GFP peak (225-375 ms) there was a significantly reduced mean current density power in the PERS group relative to the CONT group in a region of right secondary auditory cortex centered in BA22 ($x = 50$, $y = -15$, $z = 5$) ($t = 4.20$; one-tailed, $p < .05$, corrected for whole-brain comparisons). These effects are shown in Fig 1, Panel G.

DISCUSSION

Auditory-vocal Gating in Developmental Stuttering

This study employed a paradigm adopted from MEG, comparing auditory responses to SPEAK and LISTEN vowels, and therefore testing audiovocal gating. It has been hypothesized that the latter may be operating through the generation of an efference copy -- an inhibitory feed-forward projection from areas involved in motor plan initiation onto the sensory system(s), with resulting attenuation of the perceptual response (Brown et al, 2005). The finding of a reduction of the magnetic N1 response to spoken vs. heard vowels was replicated, for the first time, using high-density ERPs rather than the costly MEG (Curio et al., 2000). Critically however, no difference in N1 suppression was found between PERS and fluent controls. Similar conclusions were reached by Salmelin et al. (1998), who reported no significant group differences in the N1 responses to auditory tones comparing spoken reading conditions to silent reading conditions. These combined findings run against the hypothesis that persistent stuttering speakers exhibit an exaggerated suppression of evoked auditory response to spoken versus heard speech sounds (Brown et al, 2005). However, it is important to recognize that evidence for the efference copy model derives largely from neuroimaging data showing reduced or absent auditory activity in PERS groups

during reading tasks in which they stuttered, while in stutter-free speech (such as the /ah/ response) this was not the case (Brown et al, 2005). Given that the task in the present study did not evoke stuttering, then it would be expected that there would be little or no suppression accompanying this task. Since brain electrical activity was not recorded in similar conditions as in Brown et al's (2005), and that the latter meta-analysis did not compare auditory responses in speaking and listening tasks, the present results do not address or directly challenge their efference copy model. Furthermore, it is necessary to caution against drawing direct parallels between findings based on PET/fMRI measurements of cerebral blood flow and those obtained with measures of brain electrical activity because they do not measure the same brain signal, and may not necessarily have a counterpart in the other modality.

Speech Perception in Developmental Stuttering

While the present study found no evidence of auditory gating deficits in adults with chronic stuttering relative to fluent speakers, it did identify for the first time two abnormalities in the spatiotemporal dynamics of the brain electrical response of adults who stutter evoked by *passive listening* to speech sounds. The early auditory N1 wave to heard sounds was bilaterally symmetric in the normally fluent CONT group, while in the PERS group, the N1 topographical distribution was distorted over the right hemisphere, with a significant reduction in amplitude that peaked over right inferior posterior temporal scalp. Importantly, this effect was uniquely present in the LISTEN condition, with no differences detected in the SPEAK condition. A hemispheric asymmetry index based on the posterior N1 (LISTEN condition) correlated moderately with a measure of stuttering frequency. At later stages of the evoked response to speech sounds, PERS displayed bilateral amplitude reduction of the N3 wave over temporal scalp regions, greater over the right hemisphere.

Early right motor overactivity in chronic stuttering

While the scalp distribution of the N1 effects for the PERS and CONT group did suggest a focal abnormality in right temporal cortex, source modelling at the peak of the N1 wave showed no differences from CONT speakers in current density in auditory cortex. Instead, PERS speakers were found to abnormally recruit right motor cortex while listening passively to simple vocalized sounds. The present result of abnormal recruitment of the right rolandic area during the early stages of auditory speech perception replicate and extend the recent findings of the Biermann-Ruben, et al. (2005) MEG study. They found selective activity in the same region between 300-1000 ms from onset of a word or a brief sentence in their adult PERS group, which they attributed to abnormal preparation of a speech motor response and the complexity of the linguistic processing task. In the present study a similar recruitment was demonstrated much *earlier* in response to hearing simple vocalizations – and without an overt speech-motor response.

Furthermore, this result replicates within the EEG modality the consistent neuroimaging finding of abnormal overactivity in right motor/premotor regions in PERS speakers during stuttered and nonstuttered speech (Braun, et al., 1997; Fox, et al., 1996, 2000; Ingham, et al., 2000, 2004). Additional confirmatory evidence comes from a recent fMRI study in PERS and fluent CONT which similarly found greater activity in right primary motor cortex

and right insula during passive listening of bisyllabic words contrasting the two groups (DeNil, et al., 2008). They suggested that persons who stutter may adopt more articulatory oriented strategies than fluent individuals. This may have also occurred in the present study (although the sounds had minimal articulatory load), despite instructions to the participants to simply listen and not intentionally subvocalize or imagine the speech sounds. It is worth noting that, if indeed the PERS speakers were subvocalizing while listening to speech sounds, then this situation would be similar to what is observed during chorus reading, a condition known to eliminate stuttering as well as the reported PET abnormalities (Fox, et al., 1996). The simple task in the present study also had minimal working memory requirements (possibly relying on a phonological/ articulatory loop).

Finally, it is possible that speech sounds may implicitly/automatically activate motor speech regions. Support for this comes from a recent fMRI study in fluent speakers showing overlapping activation of *bilateral* primary motor cortex in response to uttering or listening passively to meaningless monosyllables (Wilson, et al., 2004). The authors concluded that their findings support the idea that speech perception implicates the motor system “in a process of auditory-to-articulatory mapping to access a phonetic code with motor properties” (Wilson et al., 2004). We propose here that such implicit involvement of the motor system in speech perception is abnormally accentuated, particularly in the right motor cortex, in PERS individuals. Such interpretation would account for DeNil et al.'s (2008) fMRI results of similarly greater right motor cortex activation in PERS than CONT speakers when passively listening to speech sounds, without implicating articulatory strategies in the PERS group.

Correlation with stuttering severity

A laterality index based on N1 amplitude over inferior temporal sites in the PERS group showed a significant, albeit moderate, correlation with stuttering frequency, with more interhemispheric unbalance (right N1 smaller than left N1) predicting worse stuttering. Further evidence that the N1 wave abnormality in PERS speakers may represent a ‘*state*’ marker (waxing and waning with stuttering frequency), rather than a disease or ‘*trait*’ marker of developmental stuttering, is provided by an exploratory analysis of a small independent group (n=4) of adult *recovered* PERS carried out in our laboratory using the same speech sound listening task. In this subgroup there was no hint of right N1 wave abnormality, with individual amplitudes overlapping those of the current CONT group. Further indirect support for the ‘*state*’ abnormality account comes from other studies from our group. The Fox, et al., (1996) PET study showed a normalization of premotor-motor cortical abnormalities in PERS speakers during fluent paragraph reading under chorus reading conditions.

Late right auditory cortex hypoactivity in chronic stuttering

The N3 wave in the ERP to heard voice sounds was reduced in amplitude over bilateral inferior temporal scalp in the PERS speakers, more so over the right hemisphere. Source localization with sLORETA confirmed a significant reduction of current density with a maximum in right secondary auditory cortex BA22. Unfortunately, it was not possible to perform a similar N3 wave analysis for the SPEAK data, due to the high pass settings of the

band-pass filter (4 Hz) effectively eliminating slow waves (including brain signals). Therefore, unlike the early N1, we cannot rule out similar effects for spoken sounds.

The present findings supplement and extend the findings of PET studies that have shown abnormally reduced blood flow in right secondary auditory cortex in PERS speakers during stuttered speech (Braun et al., 1997; Fox et al., 2000; Ingham, et al., 2004). In these PET studies it is noteworthy that the auditory cortex abnormalities did not occur during the real or imagined production of nonstuttered utterances. Likewise, in the present study the early N1 abnormal effects did not appear during overt (nonstuttered) utterances, and behaviorally PERS and the CONT groups performed quite similarly in the SPEAK condition. It is a well known fact that people who stutter are less likely to stutter on shorter utterances, such as saying “ah”. So, it was expected that they would not stutter on this task. Interestingly, Ingham, et al., (2000) also demonstrated deactivation of right secondary auditory cortex in a group of adult PERS speakers during *imagined* stuttering, suggesting that abnormal speech motor *programming* without overt speech output may be sufficient to generate suppression of auditory activity. Similarly, the ERP abnormalities reported here were elicited in the absence of overt speech output.

Conclusions and Caveats

The idea of a central deficit in auditory processing in developmental stuttering has been proposed long ago and tested repeatedly, yielding inconsistent results (e.g., Hall & Jerger, 1978; Rosenfield & Goodglass, 1980). Our findings support the notion that developmental stuttering involves a deficit in auditory processing which is not generalized, but rather restricted to speech sounds. We found no evidence for auditory ERP abnormalities to spoken (nonstuttered) sounds in adults who stutter, and no support for exaggerated N1 wave suppression relative to heard speech sounds (Brown et al, 2005). Rather, the ERP abnormalities were confined to the passive listening of speech sounds.

There are currently numerous theories of stuttering that could accommodate the present findings. That the neural processing of a phonated sound may be linked to an abnormal sound perception system began to be elucidated by Stromsta in the 1970s (Stromsta, 1972). That system, as has been shown (i.e., Curio et al., 2000) responds very differently during speech and when listening to self-produced speech. The findings from this study are certainly compatible with that difference, but because there is no accepted general theory of stuttering, there remains a need for additional research to show precisely how listening to a particular sound connects to an abnormal neural system – and of course to then elucidate whether that system is a by-product of stuttering or a causal factor.

The present results suggest that: 1) abnormalities in auditory processing of speech sounds coexist with abnormal responses in speech-motor areas, both being integral components of a speech-dedicated auditory-motor neural network; and 2) the dynamic interplay between auditory and motor areas is disturbed by developmental stuttering, with excessive activity in right rolandic cortex at an early stage of perceptual processing of the speech sounds. The results of the present study should be considered preliminary, and need to be replicated by future studies with larger groups, as well as fine-grained investigations in single subjects. A future combined event-related EEG-fMRI study of audiovocal gating in the same individuals

who stutter and in fluent speakers may be required to clarify how audiovocal gating findings with EEG and fMRI techniques complement or relate to each other. Functional connectivity analyses of EEG and fMRI data and structural connectivity techniques could be integrated to help understanding if auditory abnormalities in speakers who stutter depend on abnormal task-induced functional connectivity or on an underlying aberrant structural connectivity in audiomotor projections (Cykowski et al., 2010). These limitations notwithstanding, our findings provide convincing and novel evidence that abnormalities in activation and timing of right motor and right auditory areas in developmental stuttering can be shown during *passive listening* of speech sounds in the absence of overt (fluent or disfluent) speech, and that they partially predict current symptom severity.

METHODS

Participants

Eight adults diagnosed with developmental stuttering (all male, mean age 44 years; range 27-56 years, all right-handed) took part in the study. They began to stutter during early childhood, and currently displayed mild to severe stuttering. Eight males (mean age 41 years; range 25-55 years), with no history of speech impediment formed the comparison control group. All participants were native English speakers, with no history of hearing impairment, neurologic or psychiatric disorders, or drug abuse. None was currently medicated. Written informed consent was obtained from all participants.

Equipment

Participants sat in front of a table housing a microphone and a loudspeaker. The microphone was placed midway between the participant's mouth and the loudspeaker, which were 60 cm apart. The microphone was routed to a Macintosh LCII computer equipped with SuperLab 1.7 (Cedrus, San Pedro, CA), making it possible to record voice onset time for speech sounds produced by the participants and replay their voiced sounds from the loudspeaker. The participant wore a headset microphone (AKG C420) allowing the voice signal to be recorded using a DAT recorder (Sony PCM-R500). During each condition, voice onset times were computed and digital codes were transmitted through a serial port to the EEG digitization computer, to allow off-line voice-onset time-locked ERP averaging.

Speech performance measures

All PERS participants were assessed prior to the experiment during 5-min audio-visually recorded oral reading and spontaneous speaking tasks (Costello & Ingham, 1984). The average percent syllables stuttered scores across both tasks for the PERS participants was recorded. An independent judge re-assessed all recordings with a discordance rate of less than 1.0%.

Tasks

These were closely adapted after Curio et al. (2000). For the 'SPEAK' condition, participants were trained to intermittently utter the vowel "ah" for 1 s at 3 s intervals, at a loudness level of 70-80 db for about 90 s. During the 'LISTEN' condition, the previous SPEAK block was replayed to the subject through the loudspeaker. Loudness level was

individually adjusted to match the loudness of the previous SPEAK block (70-80 db SPL at 30 cm). Subjects were instructed to hold their mouth slightly open and the jaw relaxed, so as to minimize motion and muscle artifacts, and to blink as little as possible. In the listen condition, subjects were instructed to simply listen to the vowel sounds.

Participants alternately completed eight SPEAK and eight LISTEN blocks, separated by short resting periods. An average of 150 “ahs” was spoken or heard by each participant in each condition. All tasks were completed by a participant within approximately 30 min. This study employed time-locking to onset of *voicing* rather than the more traditional onset of *sound* for the following reasons. Curio's audiovocal gating task (2010) compares auditory responses elicited by the participant's spoken syllables to the auditory responses evoked in response to the same syllables replayed to the participant in the next block. Since own-speech sounds were generated ‘online’ and not digitized in advance, and they were expected to slightly differ in terms of intensity raising time and duration, sound onset triggers could not be programmed and transmitted to the EEG digitizing computer as in traditional auditory ERP studies. Second, while ‘online’ speech sounds were recorded, the EEG digitizing software did not include an audio channel to allow synchronization of the recorded sounds with the ongoing EEG signals. We had much better control of voice onset time, since adjustments of loudness and careful calibrations of the threshold for vocal responses were repeated for each block of the study.

Vocal onset triggers were found to lag the real vowel sound onset by about 50 ms. This delay was estimated by the timing of the auditory N1 wave in two subjects where similar vocalizations were recorded in advance and ERPs were additionally computed time-locked to the onset of the speech sounds. The delay was ascribable to the gradual onset of the ‘aah’ sounds, and the time required to reach the sound pressure level to trigger a vocal response. This systematic delay did not affect the analysis presented below because the triggering speech sounds were the same for the SPEAK and LISTEN conditions, and sounds pressure levels were carefully matched for each SPEAK/LISTEN pair of each participant. Importantly, mean sound pressure levels for the LISTEN blocks did not differ between the CONT (74.16 ± 5.4 dB) and PERS (72.44 ± 7.2 dB) groups as confirmed by an independent samples t-test ($p > .20$).

EEG recording

High-density brain electrical activity was continuously recorded using a customized 64 electrode cap (Electrocap Inc®, Eaton, OH) including four eye movement electrodes (two at the external canthi and two infraorbital) and a left mastoid electrode, all referenced to the right mastoid (amplifier settings: bandpass = 0.01-100 Hz, gain = 10K, sampling rate = 400 Hz, impedances $< 5k\Omega$). Eye movement artefacts (blinks and eye movements) were rejected off-line. ERPs time-locked to the onset of the vocal response were selectively averaged for each participant and group for the SPEAK and LISTEN conditions. For the N1 wave measurements, EEG off-line processing for the N1 wave (see below) included band-pass filtering at 4-28 Hz to remove slow artifacts arising from head, breathing or articulator movements and from slow motor-related brain activities during the SPEAK recordings (see Curio et al, 2000 for a similar approach to N1 analysis with MEG). For the slower

component N3 (see below), only a low-band pass filter was applied (<28 Hz). Following that, ERPs were algebraically re-referenced to average reference (Tucker et al., 1994). ERP participant averages for each group were then grand-averaged across participants. ERP amplitudes were aligned to a 200 ms pre-stimulus baseline period. Final ERP subject averages included the following mean number of trials. For CONT, SPEAK: 158 (range: 120-220), LISTEN: 148 (range: 110-215); for PERS, SPEAK: 145 (range: 110-165), LISTEN: 145 (range: 115-165). A repeated Measures ANOVA confirmed the lack of differences in number of trials as a function of trial type, Group, or trial type $\times$ Group (ANOVA test, for all, $p > .15$).

EEG Analysis

Grand-average ERPs elicited during the SPEAK and the LISTEN conditions revealed an ordered succession of auditory components with typical scalp topography (N1, P2 and N3, see Fig. 1). Note that the use of the average reference rather than a conventional average mastoid or linked ears reference gave rise to a polarity inversion of the typical frontocentrally distributed negative N1 and N3 components over bilateral ventral temporal scalp (with relative positivities there), reflecting the bipolar nature of EEG source dipoles (see waveforms and scalp topographies in Fig 1) (Tucker et al, 1994). Visual inspection of such grand-average waveforms and topographical maps revealed a substantial reduction of amplitude of the fronto-central N1 waveform the SPEAK relative to the LISTEN condition (Fig.1, panels B-C), and a striking reduction of amplitude of the N1 wave in the LISTEN condition for the PERS group over right inferior temporo-occipital scalp (See Fig 1, panel A). To analyze such effects, mean voltage amplitudes around the N1 peak (20-80 msec post-voice onset) were extracted for both conditions, hemispheres and groups over frontocentral and posterior inferior scalp (corresponding to extrema in N1 voltage). For the former, a region of interest (ROI) was employed, collapsing sites F3i/F4i, C3a/C4a, C5a/C6a); for the latter, two symmetric ROIs were computed, collapsing sites T11/TI2, TO1/TO2, O1'/O2', I1/I2). Later on, for the LISTEN condition, the N3 wave also appeared reduced in amplitude for the PERS group relative to the CONT group (See Fig. 1, panel D-E). Mean voltage amplitudes around the N3 peak (225-375 ms time window) were computed for both hemispheres and groups. Given the similar scalp topography as the prior N1 wave, the same ROIs were employed over frontocentral and posterior inferior scalp. Due to the polarity inversion of the N1 and N3 waves over frontocentral and posterior inferior scalp, separate Repeated Measures ANOVAs were carried out over anterior and posterior scalp regions. The frontocentral N1 analysis included Condition as within and Group as between factor. The posterior inferior N1 analysis additionally included hemisphere as within factor. The N3 wave analyses omitted the within factor condition, being limited to the LISTEN condition. For all ANOVAs, p value for significance was set at .05. Degrees of freedom were corrected using the Greenhouse-Geisser epsilon method to correct for sphericity.

Source Localization

Standardized low-resolution electromagnetic tomography (sLORETA, Pascual-Marqui, 2002) was used to localize the cortical sources of the early and late scalp ERP group differences between PERS and CONT. During each time window sLORETA performs a linear computation of the current density in 6239 cortical gray matter voxels (sized 5 mm³),

and based on the assumption that cortical activity is smoothly distributed, it localizes the maximum activity within the blurred-activity image providing its location in stereotactic MNI space (Pascual-Marqui, 2002). For each subject and group, sLORETA was computed around the times of maximal global field power (GFP) within the early N1 and late N3 time windows. The GFP is computed as the average standard deviation within the potential field; the GFP peaks are hypothesized to index time points associated with maximal neuronal activity and thus offer optimal signal-to-noise ratio (Lehmann & Skrandies, 1984). Individual subject average data were transformed into sLORETA images, which contain the 3-D distribution of electrical cortical activity. The power of the estimated electrical current density (in ampere per meter squared, A/m²) underwent logarithmic transformation to limit the effect of outliers. Differences in activity between groups were statistically analyzed on a voxel-by-voxel basis using the independent samples Student t test, and the threshold was set to $P < 0.05$ through a nonparametric permutation test with 5000 randomizations accounting for multiple whole-brain comparisons.

For the N1 group contrast, to increase signal to noise, a more restricted time window centered around the GFP peak for the group comparison was employed (50-60 ms rather than 20-80 ms). For the N3 group contrast, the GFP peak was more sustained and coincided with the 225-375 ms N3 scalp window. Based on the direction of the N1 and N3 group voltage differences on the scalp (voltage reductions in both cases in the PERS group), a-priori predictions were made of correspondent current density reductions for the s-LORETA group contrasts. Therefore, and to maximize signal to noise, one-tailed $p < 0.05$ significance thresholds were initially chosen for both time windows. However, the N1 current density group comparison image revealed prominent current density *increases* rather than the expected decreases. Therefore, a two-tailed $p < 0.05$ significance threshold was employed for the N1 group contrast to test for *both* regional increases and decreases.

Acknowledgments

Supported by the Research Imaging Institute and a NIH grant (RO1 DC007893) awarded to RJI, and by funds from the Canada Foundation for Innovation (CFI) and National Science and Engineering Council of Canada (NSERC) awarded to ML.

REFERENCES

- Biermann-Ruben K, Salmelin R, Schnitzler A. Right rolandic activation during speech perception in stutterers: a MEG study. *Neuroimage*. 2005; 25(3):793–801. [PubMed: 15808980]
- Brown S, Ingham RJ, Ingham JC, Laird AR, Fox PT. Stuttered and fluent speech production: an ALE meta-analysis of functional neuroimaging studies. *Human Brain Mapping*. 2005; 25:105–117. [PubMed: 15846815]
- Braun AR, Varga M, Stager S, Schulz G, Selbie S, Maisog, et al. Altered patterns of cerebral activity during speech and language production in developmental stuttering. An H2(15)O positron emission tomography study. *Brain*. 1997; 120:761–784. [PubMed: 9183248]
- Costello, JM.; Ingham, RJ. Assessment strategies for children and adult stutterers.. In: Curlee, R.; Perkins, WH., editors. *Nature and Treatment of Stuttering: New Directions*. College-Hill; San Diego CA: 1984. p. 303-333.
- Curio G, Neuloh G, Numminen J, Jousmaki V, Hari R. Speaking modifies voice-evoked activity in the human auditory cortex. *Human Brain Mapping*. 2000; 9:183–91. [PubMed: 10770228]

- Cykowski M, Kochunov P, Ingham RJ, Ingham JC, Mangin JF, Rivière D, et al. Perisylvian sulcal morphology and cerebral asymmetry patterns in adults who stutter. *Cerebral Cortex*. 2008; 18:571–583. [PubMed: 17584852]
- De Nil LF, Kroll RM, Kapur S, Houle S. A positron emission tomography study of silent and oral single word reading in stuttering and nonstuttering adults. *Journal of Speech Language & Hearing Research*. 2000; 43:1038–1053.
- De Nil LF, Beal DS, Lafaille SJ, Kroll RM, Crawley AP, Gracco VL. The effects of simulated stuttering and prolonged speech on the neural activation patterns of stuttering and nonstuttering adults. *Brain and Language*. 2008; 107:114–123. [PubMed: 18822455]
- Finitzo, T.; Pool, KD.; Freeman, FJ.; Devous, MD.; Watson, BC. Cortical dysfunction in developmental stutterers.. In: Peters, HFM.; Hutstijn, W.; Starkweather, CW., editors. *Speech Motor Control and Stuttering*. Elsevier; Amsterdam: 1991. p. 251–261.
- Fox PT. Brain imaging in stuttering: Where next? *Journal of Fluency Disorders*. 2003; 28:265–272. [PubMed: 14643065]
- Fox PT, Ingham RJ, Ingham JC, Hirsch TB, Downs JH, Martin C, et al. A PET study of the neural systems of stuttering. *Nature*. 1996; 382:158–161. [PubMed: 8700204]
- Fox PT, Ingham RJ, Ingham JC, Zamarripa F, Xiong JH, Lancaster JL. Brain correlates of stuttering and syllable production. A PET performance-correlation analysis. *Brain*. 2000; 123:1985–2004. [PubMed: 11004117]
- Hall J, Jerger J. Central auditory function in stutterersf. *Journal of Speech and Hearing Research*. 1978; 21:324–337. [PubMed: 703279]
- Hampton A, Weber-Fox C. Non-linguistic auditory processing in stuttering: evidence from behavior and event-related brain potentials. *Journal of Fluency Disorders*. 2008; 33(4):253–73. [PubMed: 19328979]
- Ingham RJ. Brain imaging studies of developmental stuttering. *Journal of Communication Disorders*. 2001; 34:493–516. [PubMed: 11725861]
- Ingham RJ, Fox PT, Ingham JC, Zamarripa F. Is overt stuttered speech a prerequisite for the neural activations associated with chronic developmental stuttering? *Brain & Language*. 2000; 75:163–194. [PubMed: 11049665]
- Ingham RJ, Fox PT, Ingham JC, Xiong JH, Zamarripa F, Hardies LJ, Lancaster JL. Brain correlates of stuttering and syllable production: Gender comparison and replication. *Journal of Speech Language Hearing Research*. 2004; 47:321–341.
- Lehmann D, Skrandies W. Spatial analysis of evoked potentials in man--a review. *Progress in Neurobiology*. 1984; 23(3):227–50. [PubMed: 6395186]
- Max L, Guenther FH, Gracco VL, Ghash SS, Wallace ME. Unstable or insufficiently activated internal models and feedback biased motor control as sources of dysfluency: A theoretical model of stuttering. *Contemporary Issues in Communication Science and Disorders*. 2004; 31:105–122.
- Pascual-Marqui RD. Standardized low-resolution brain electromagnetic tomography (sLORETA): technical details. *Methods & Findings in Experimental & Clinical Pharmacology*. 2002; 24:5–12. [PubMed: 12575463]
- Rosenfield DB, Goodglass H. Dichotic testing of cerebral dominance in stutterers. *Brain & Language*. 1980; 11:170–180. [PubMed: 7427716]
- Rosenfield, DB.; Jerger, J. Stuttering and auditory function.. In: Curlee, RF.; Perkins, WH., editors. *Nature and treatment of stuttering: New directions*. College-Hill Press; San Diego, CA: 1984. p. 73–87.
- Salmelin R, Schnitzler A, Schmitz F, Jancke L, Witte OW, Freund HJ. Functional organization of the auditory cortex is different in stutterers and fluent speakers. *Neuroreport*. 1998; 9:2225–2229. [PubMed: 9694204]
- Salmelin R, Schnitzler A, Schmitz F, Freund HJ. Single word reading in developmental stutterers and fluent speakers. *Brain*. 2000; 123:1184–1202. [PubMed: 10825357]
- Sommer M, Koch MA, Paulus W, Weiller C, Büchel C. Disconnection of speech-relevant brain areas in persistent developmental stuttering. *Lancet*. 2002; 360(9330):380–3. [PubMed: 12241779]
- Stromsta, C. *Elements of stuttering*. Atsmorts; Oshtemo, Michigan: 1986.

- Tucker DM, Liotti M, Potts GF, Russell GS, Posner MI. Spatiotemporal analysis of brain electrical fields. *Human Brain Mapping*. 1994; 1:134–152.
- Weber-Fox C. Neural systems for sentence processing in stuttering. *Journal of Speech Language Hearing Research*. 2001; 44:814–825.
- Wilson SM, Saygin AP, Sereno MI, Iacoboni M. Listening to speech activates motor areas involved in speech production. *Nature Neuroscience*. 2004; 7:701–702.
- Woldorff MG, Hillyard SA, Gallen CC, Hampson SR, Bloom FE. Magnetoencephalographic recordings demonstrate attentional modulation of mismatch-related neural activity in human auditory cortex. *Psychophysiology*. 1998; 35:283–92. [PubMed: 9564748]

Figure 1.

Panel A: Grand average ERP waveforms in the LISTEN condition over 5 representative sites for CONTR (blue) and PERS (purple) groups for the first 200 ms following voice onset, illustrating group differences in amplitude and spatial topography of the early group response to the heard sound.

Panel B: Grand average ERP waveforms to speech sounds in the LISTEN condition (in blue) and the SPEAK condition (in purple) over 5 representative sites for the CONTR group (left) and the PERS group (right) for the first 200 ms following voice onset, illustrating

fronto-central differences in the amplitude of the N1 component response to spoken and heard sounds independent of group.

Panel C: Scalp topography voltage maps of the early ERP effect (20-80 ms) for the LISTEN and SPEAK conditions in the CONT group and the PERS group.

Panel D: Grand average ERP waveforms in the LISTEN condition, time-locked to the voice trigger onset over 3 representative sites for CONT (blue) and PERS (purple) groups for 700 ms following voice onset, illustrating differences in amplitude and spatial topography of the late group response to the heard sound.

Panel E: Scalp topography voltage maps of the late ERP effect (225-375 ms) for the CONT (left) and PERS (right) speakers.

Panel F: s-LORETA results showing areas where current density was greater for PERS than CONT at the peak of the early group effect (40-50 ms). Note maximum in right rolandic Cortex BA 4.

Panel G: s-LORETA results (225-375 ms) showing areas where current density was greater for CONT than for PERS. Note maximum in right Auditory Cortex BA 22.