

2015

Neural mechanisms of speech motor learning in persons who stutter

<https://hdl.handle.net/2144/16041>

Downloaded from DSpace Repository, DSpace Institution's institutional repository

BOSTON UNIVERSITY
SARGENT COLLEGE OF HEALTH AND REHABILITATION SCIENCES

Thesis

**NEURAL MECHANISMS OF SPEECH MOTOR LEARNING
IN PERSONS WHO STUTTER**

by

ANNA OH

B.S., University of Toronto, 2011

Submitted in partial fulfillment of the
requirements for the degree of
Master of Science

2015

Approved by

First Reader

Jason Tourville, Ph.D.
Research Assistant Professor of Speech, Language and Hearing Sciences

Second Reader

Frank Guenther, Ph.D.
Professor of Speech, Language and Hearing Sciences
Professor of Biomedical Engineering

Third Reader

Swathi Kiran, Ph.D., CCC-SLP
Professor of Speech, Language and Hearing Sciences

NEURAL MECHANISMS OF SPEECH MOTOR LEARNING

IN PERSONS WHO STUTTER

ANNA OH

ABSTRACT

Fluent speech production requires rapid coordination among respiratory, laryngeal, and articulatory processes and is mediated by multiple neural systems (Bohland & Guenther, 2006). Stuttering is a fluency disorder characterized by core deficits in speech motor planning. Previous research indicates people who stutter (PWS) exhibit deficits in speech motor sequence learning and are slower and less accurate over practice relative to fluent speakers (Ludlow, Siren, & Zikira, 2004; Namasivayam & VanLieshout, 2004; Smits-Bandstra & De Nil, 2007; Smits-Bandstra, De Nil, & Saint-Cyr, 2006). Furthermore, the neural bases of impaired speech motor sequence learning in PWS are not well understood. We present a study in which PWS (n=18) and persons with fluent speech (PFS) (n=17) were taught phonotactically illegal (e.g. gbesb) and phonotactically legal (e.g. blerk) speech motor sequences over two practice sessions. Functional magnetic resonance imaging (fMRI) was used to investigate brain regions underlying the production of learned illegal syllables and novel illegal syllables. With practice, subjects produced syllables more accurately, which is indicative of motor sequence learning. Our findings suggest a speech motor performance deficit in PWS. Furthermore, these findings indicate speech motor sequence learning relies on a speech motor sequence learning network.

TABLE OF CONTENTS

INTRODUCTION	1
METHODOLOGY	8
Speech Stimuli	9
Behavioral Learning Paradigm	10
fMRI Paradigm	11
Behavioral Data Analysis	12
fMRI Data Analysis	16
RESULTS	19
Practice Session Behavioral Measures of Learning.....	19
fMRI Behavioral Measures of Learning.....	21
fMRI Syllable Segment Duration Measures	24
fMRI Analysis.....	25
Neural-Behavioral Correlation Analysis.....	27
DISCUSSION.....	28
REFERENCES	35

List of Tables

Table 1. Orthographic representations of syllables with phonotactically legal consonant clusters in English (left) and phonotactically illegal consonant clusters in English (right). 9

List of Figures

- Fig 1.** Segmented spectrogram templates for legal syllable BRALK (top) and illegal syllable PTASCHT (bottom). The black vertical bars mark phoneme onsets. Orthographic labeling of each phoneme is represented at the top of each spectrogram. 14
- Fig 2.** Percent accuracy for PWS (white) compared to PFS (grey) in practice session 1 (left) and practice session 2 (right). Error bars represent standard error. 20
- Fig 3.** Changes in behavioural learning indices between practice session 1 and practice session 2. Both PWS (white) and PFS (grey) produced fewer errors for learned illegal syllables in practice session 2 compared to practice session 1 (left). There were no significant differences between practice sessions for syllabic utterance duration (middle) and reaction time (right) across Groups or Conditions. 21
- Fig 4.** Percentage of errors for PWS (white) compared to PFS (grey) in the fMRI session. Error bars represent standard error. Subjects produced more errors for the learned illegal syllables compared to the legal syllables and for the novel illegal syllables compared to the legal syllables. 22
- Fig 5.** Reaction Time for PWS (white) compared to PFS (grey) in the fMRI session. Error bars represent standard error. 23
- Fig 6.** Percent Dysfluency of PWS (white) compared to PFS (grey) in the fMRI session. Error bars represent standard error. PWS produced significantly more dysfluencies compared to PFS. 24
- Fig 7.** Segment duration results from syllable productions in the fMRI session. 25
- Fig 8.** BOLD responses in the novel illegal-learned illegal contrast for all subjects ($CWP < 0.01667$) Abbreviations: AI = anterior insula, PMC = premotor cortex, pSTG = posterior superior temporal gyrus, ITO = inferior temporal-occipital cortex, SPL = superior parietal lobule, preSMA = presupplementary motor cortex. 26

INTRODUCTION

Developmental stuttering is a motor speech disorder that has a significant impact on communication. Stuttering is present in approximately 1% of the population and is characterized by speech sound repetitions, prolongations, and blocks. Stuttering is also often accompanied by secondary behaviors (e.g. facial grimace, eye blinking) that have been acquired as an escape or avoidance strategy. It is a multidimensional disorder that involves underlying negative thoughts and feelings (Craig, 2002).

The etiology of stuttering is unknown; however, the symptoms suggest that it involves a deficit in neuromotor planning and/or execution. Several studies have attributed stuttering to limited speech motor skills that are strained during highly complex motor tasks, as well as differences in the timing and execution of motor plans (Buchel & Sommer, 2004; Navisavayam & Lieshout, 2011; Peters, Hulstijn, & Van Lieshout, 2000; Ludlow & Loucks, 2003; Van Lieshout, Hulstijn, & Peters, 1996).

Numerous studies have suggested stuttering involves a limitation in nonspeech and speech motor skill learning (Namasivaya & Lieshout, 2011; Smits-Bandstra, DeNil, & Saint-Cyr, 2005). Motor skill learning refers to the acquisition of temporal and spatial accuracy of movements. With practice, movements become more stable and accurate and learners exhibit more automaticity and less reliance on sensory feedback (Willingham, 1995). Compared to fluent speakers (PFS), persons who stutter (PWS) exhibit limitations

in speech motor sequence learning in both speech and non-speech tasks (Namasivayam & Lieshout, 2004; Smits-Bandstra, DeNil, Rochon, & 2006).

Studies of non-speech motor learning, e.g., novel finger tap sequences, indicate that PWS do not improve performance with practice to the same degree as PFS. PWS exhibit longer and more variable reaction times when initiating practiced motor sequences and longer duration times (Smits-Bandstra & De Nil, 2009). Furthermore, PWS transfer and retain skills to a lesser degree compared to fluent speakers (Namasivayam & VanLieshout, 2004; Smits-Bandstra, DeNil, & Rochon, 2006).

These differences were also exhibited when speech motor sequences were learned. Speech production, like other motor skills, requires the learning of smaller sequential movements for articulation to become stable and efficient (Rosenbaum, 1983). In a study by Bauerly and DeNil (2011), PWS were taught nonsense speech sequences over two days of practice. The PWS group demonstrated significantly longer and more variable durations of sequences relative to fluent speakers following a 24-hr consolidation period. Smits-Bandstra, DeNil and Saint-Cyr (2006) noted a similar trend in longer durations in PWS following syllable sequence learning compared to PFS. More recently, Sasisekaran & Weisberg (2014) manipulated syllable length, phonotactic constraint, and complexity to characterize the nature of motor skill learning deficits in PWS. Participants were instructed to repeat nonwords across two sessions separated by a 1-hour consolidation period. Findings showed that PWS exhibited more errors for complex and phonotactically

illegal nonwords vs. simple words. Kinematic analysis indicated that PWS benefited less from practice on measures of inter-articulatory coordination for both simple and complex nonwords. In summary, findings from numerous studies suggest that PWS benefit less from practice in nonspeech and speech motor learning tasks when compared to PFS. Motor learning limitations are particularly evident when syllable sequences increase in length, complexity, and phonotactic constraint. Thus, there appear to be general deficits in motor learning in PWS extending to the speech motor system.

A wide range of studies, including pharmacological treatments for stuttering, the effect of deep brain stimulation on fluency, neurotransmitter mapping, and functional neuroimaging support behavioral findings that indicate a disruption in motor sequence learning mechanisms in PWS. These studies demonstrate differences in the brains of PWS and fluent speakers in brain regions associated with the planning and initiation of motor plans (see Alm, 2004, for review). Specifically, anomalies in the basal-ganglia-thalamo-cortical motor circuit and cortical areas such as SMA have been implicated in stuttering, as these brain areas are critical for motor learning and execution (Alm, 2004; Bandstra & DeNil, 2007; Civier et al., 2013; Ingham et al., 2013; Lu et al., 2010). If PWS do exhibit an impaired ability to learn novel speech motor sequences, it is likely that the impairment is associated with disruptions within the speech motor sequence network.

Evidence from various neuroimaging studies indicates that brain regions associated with learning novel speech motor sequences are disrupted in PWS. In a previous study of

speech motor sequence learning, Rauschecker, Pringle and Watkins (2008) found that covert repetition of novel pseudowords was associated with changes in neural activity in language and motor areas of the brain. Activity increased bilaterally in inferior frontal cortex, superior temporal cortex, putamen, dorsal premotor cortex, anterior cerebellum, medial SMA/preSMA, cingulate motor area, and the cerebellar vermis. These findings are consistent with regions that have been implicated in non-speech motor sequence learning. Studies have attributed motor learning skills to changes in premotor cortex, preSMA/SMA, basal ganglia, and the cerebellum (Cunnington, Windischberger, Deecke, & Moser, 2002; Hikosaka, Sakai, Miyauchi et al., 1996; Hikosaka, Nakamura, Sakai et al., Toni, Kramus, Turner, & Passingham, 1998). In a more recent study of speech motor sequence learning in healthy adults, Segawa, Beal, Tourville and Guenther (2015), found that motor learning success correlated with decreased activation in left frontal operculum. Results showed that speech learning was mediated by less activation in frontal and posterior superior temporal cortex, preSMA, and basal ganglia, whereas the production of novel speech motor sequences increased activation in left hemisphere frontal operculum, adjacent anterior insula cortex, posterior superior temporal sulcus, planum temporale, inferior temporal-occipital cortex, globus pallidus, bilateral lateral superior parietal lobule, and right hemisphere pre SMA.

In the current study, we aimed to increase our understanding of the neural mechanisms that underlie learning novel speech sequences and how these processes are impaired in PWS. PWS and PFS practiced phonotactically illegal and legal speech motor sequences

over two days and the effects of practice were compared between groups. To assess learning, we used standard motor sequence learning indices (e.g., reaction time, duration, error rate) and measures of acoustic variables (formant and pitch). Structural and functional neuroimaging was used to characterize the neural correlates of speech motor sequence learning in PWS and PFS. The duration of individual segments within each syllable sequence was assessed using dynamic time warping to determine the effect of specific segments or segment transitions on learning (Holmes & Holmes, 2001).

Based on previous behavioral findings, we expected that both groups would exhibit learning effects, as measured by changes in motor sequence learning indices. However, we hypothesized that PWS would show learning deficits and less automatization of sequence production relative to PFS. Specifically, we expected practice would result in a greater reduction in rate of phonemic errors, reaction time, total duration, and the duration of sub-syllabic segments in PFS.

Sequence duration is a traditional measure of motor sequence learning and performance is characterized by shorter sequence duration over practice (Schmidt, 2004). In both non-speech (fingertapping) and speech (nonword) motor sequence learning tasks, both PFS and PWS have demonstrated shorter sequence durations after practice (Bauerly & DeNil, 2011; Smits-Bandstra, Rochon, & DeNil, 2006). We expected that PWS would show less reduction in utterance due to practice compared to PFS. Studies have shown reduced practice effects in PWS in both speech and non-speech tasks. PWS exhibit longer

sequence durations, one measure of learning success, when compared to PFS, indicating limited motor sequence learning abilities (Bauerly & DeNil, 2011; Smits-Bandstra, Rochon, & DeNil, 2006).

Based on previous neuroimaging findings, we expected that learning would be mediated by a speech motor sequence network. We expected decreased activation in frontal operculum, posterior superior temporal cortex, premotor cortex, preSMA/SMA, basal ganglia, and cerebellum in the learned illegal condition compared to the novel illegal condition, as these areas are expected to be involved in motor sequence learning (Segawa, Beal, Tourville & Guenther, 2015). We also expected to find significant between-group differences in activity within the speech motor sequence network in the novel illegal learning condition compared to the learned illegal condition.

Furthermore, we expected to find a correlation between learning success and activation in the brain regions underlying speech motor sequence learning. A positive correlation between measures of learning success (defined here as *a reduction* in error rate, reaction time, total duration, and segmental duration due to practice) and activity within the speech motor sequence network was expected. In a previous study of speech motor sequence learning, indices of learning success were correlated positively with activation in the left frontal operculum in the novel illegal vs. learned illegal contrast (Segawa, Tourville, Beal & Guenther, 2015). Specifically, the reduction in total duration between novel and learned syllables was positively correlated with activity in left frontal

operculum-anterior insula and larger FA values under the pSTG (Segawa, Tourville, Beal, & Guenther, 2015). As we expected PWS to benefit less from practice and demonstrate less reduction in utterance and segment duration for learned illegal syllables, a greater difference in duration was expected between novel illegal syllables and learned illegal syllables. Therefore, increased activity was expected to be noted in the brain regions mediating this behavioral measure of learning success.

Lastly, we expected that PWS would demonstrate differences in neural activity within the speech motor sequence network in the learned illegal condition compared to the novel illegal condition. Previous studies have shown that the brain regions critical for motor learning and execution are impacted in PWS (Alm, 2004; Civier et. al., 2013). Overall, we expected that PWS would show an impaired ability to learn novel speech motor sequences, which would be associated with disruptions in the speech motor sequence network.

METHODOLOGY

Data were collected from eighteen persons who stutter (14 male, mean age 25.9 ± 6.4 years) and seventeen fluent speakers (15 male, mean age 26.1 ± 6.4 years). Inclusion criteria were English as a first language, no history of neurological, psychiatric, and speech, language, and hearing disorders (other than stuttering in PWS). All participants reported normal or corrected-to-normal vision and were right-handed, as measured by the Edinburgh handedness inventory. Informed consent was obtained according to the Boston University Institutional Review Board and the Massachusetts General Hospital Human Research Committee. Stuttering severity ranged from mild to severe as measured by the Stuttering Severity Instrument (4th Edition) ($M= 24, 13-43$). Assessment of the affective and cognitive reactions to stuttering also ranged from mild to severe as shown by scores on the Overall Assessment of the Speaker's Experience with Stuttering (OASES) ($M= 2.48, 1.57-3.83$). The Peabody Picture Vocabulary Test, 4th Edition (PPVT-4) was administered to both groups and showed normal receptive language skills in both PWS ($M=114.50 \pm 11.99$) and PFS ($M=113.31 \pm 8.36$). An independent t-test comparing the groups in PPVT-4 scores found no significant difference. 1 PFS subject and 2 PWS were excluded from the study due to errors in the acquisition of fMRI data. Overall, 17 PWS and 16 PFS were included in measures of the overall duration, errors, RT, and dysfluencies. 12 PWS and 10 PFS were included in the subsyllabic segment duration measures. Subjects were excluded due to poor performance (schwa insertions, prolonged utterances) and/or poor recording quality (distortions and low SNR ratio). 14 PWS and

14 PFS were included in the neuroimaging analysis. Subjects were excluded due to a large amount of errors or unusable trials.

Speech Stimuli

Subjects produced single syllables composed of bi- or tri- consonantal onset and coda clusters (Table 1). Syllables were either *legal*, those that contain consonant clusters that are phonotactically legal in English (e.g., dralf), or *illegal*, those that contain consonant clusters that are phonotactically illegal in English (e.g., gvazf). All illegal consonant clusters exist in another natural language; however, participants with prior experience with these languages were excluded from the study. Each syllable had a unique

Table 1.

Orthographic representations of syllables with phonotactically legal consonant clusters in English (left) and phonotactically illegal consonant clusters in English (right).

<i>Legal Syllables</i>			<i>Illegal Syllables</i>			
blerk	twerve	shridth	fsefk	bdangt	zbapk	vbimk
fremf	swarf	kwanst	vthasp	tbastf	kvachk	bzinsch
krenth	dralf	splerst	shtazg	fthamch	gbesb	fzichb
tralp	gralve	stips	bvimpf	fpesch	tpipf	ftebsch
gwefth	prenge	plarth	tvitp	vgamsh	vsepsh	kpeshch
spridth	dwilm	thrimf	ptachst	gvazf	shkevt	zgekf
bralk	thwilb	flisk	fshikp	dkedv	zdebg	
glanch	skeln	klelth	zvekch	tgitk	tfipshch	

consonant cluster, and the number of phonemes per syllable was balanced across conditions. Syllables did not form either an orthographic or phonological word according to the MRC Psycholinguistic Database (Coltheart, 1981).

Behavioral Learning Paradigm

The behavioral learning paradigm described in Segawa, Tourville, Beal, and Guenther (2015) was implemented. Each participant was tested over two consecutive practice sessions. Participants were instructed to repeatedly produce 15 randomly presented illegal syllables and 15 randomly presented legal syllables. Each syllable was produced 30 times in each session. Subjects were randomly assigned to one of four groups; each group produced the same subset of legal and illegal syllables. To facilitate memory consolidation, practice sessions took place one or two days before scanning. Participants were presented with auditory models of each syllable, as well as orthographic representations. Stimulus duration and amplitude were normalized using Praat (<http://www.praat.org>). Onset of the visual and auditory stimuli was aligned. After presentation of the stimulus and a jittered pause of 500 to 1000 ms, a tone cued subjects to produce the target syllable. Subject utterances were recorded with a Samson C01U USB studio condenser microphone. Participants were instructed to produce the syllables as quickly and as accurately as possible. They were asked to replicate the auditory stimulus and produce all the sounds displayed in the orthographic representation of the syllables.

fMRI Paradigm

Subjects produced 15 *legal* and 15 *illegal* syllables that they learned over the two practice sessions and the 15 *novel illegal* syllables that they had not learned, resulting in three syllable production conditions: *learned legal*, *learned illegal*, and *novel illegal*. A baseline condition in which subjects viewed asterisks on the screen and rested quietly was also included.

A sparse sampling paradigm, which allowed subjects to hear auditory cues and produce target syllables in the absence of scanner noise, was used to acquire fMRI data (Hall et al., 1999). Subjects were instructed to follow the same behavioral paradigm used during the practice sessions, however a pause was added after the syllable production to temporally align the image acquisition to the expected peak of the hemodynamic response (Belin, Zatorre, Hoge, Evans, & Pike, 1999). A single trial lasted 10 s. Each run consisted of 40 trials and lasted 7 minutes. Subjects completed 8 runs, 80 trials per condition, and 5 or 6 productions of each syllable. Conditions were pseudo-randomly distributed across the 8 runs with at least 8 instances of each condition appearing in each run.

Instructions and visual stimuli were projected onto a screen viewed from within the scanner via a mirror attached to the head coil. Auditory stimuli were played over Sensimetrics model S-14 MRI-compatible insert headphones. Subjects' responses were transduced by a Fibersound model FOM1-MR-30m fiber optic microphone and sent to a Lenovo ThinkPad X61s, where they were recorded using MATLAB at 44.1 kHz.

MRI data were acquired using a 3 Tesla Siemens Trio Tim scanner with a 32-channel

head coil. For each subject, a high-resolution T1-weighted volume was acquired (MPRAGE, voxel size: 1 mm³, 256 sagittal images, TR: 2530 ms, TE: 3.44 ms, flip angle: 7°). Blood oxygen level dependent (BOLD) activity during the behavioral task was assessed with gradient echo EPI scans (41 horizontal slices, in plane resolution: 3.1 mm, slice thickness: 3 mm, gap: 25%, TR: 10 s, TA: 2.5 s, TE: 20 ms) automatically registered to the AC-PC line and was collected sparsely with 10 s between scan onsets.

Behavioral Data Analysis

To evaluate speech motor sequence learning, we assessed the following learning success measures that indicate how easily a speaker produces a speech sequence: error rate, reaction time, and total utterance duration (Sternberg, Monsell, Knoll, & Wright, 1978). Non-fluent productions were labeled by type of stutter: repetition, prolongation, mid-word block or clustered. Errors were defined as phoneme additions, deletions, and substitutions, and utterance repetitions, omissions, and restarts. A trained rater judged stutter and error types for all trials. Reaction time was defined as the time from the GO signal to onset of voicing and total duration was defined as the time of voicing onset to offset.

To assess the effects of learning from practice session 1 to practice session 2, we compared the error rate, total duration, and reaction time for the first five productions of each syllable during each practice session. Each behavioral learning index was averaged first within each syllable, then within each condition, and then within each subject. We used analysis of variance (ANOVA) and post-hoc paired t-tests to compare the mean error rate, duration, and reaction time in the illegal and legal conditions.

In addition to measuring the total duration of each utterance, the durations of phonemes within each syllable produced in the fMRI session were measured. Due to the phonotactic constraints of the utterances, a pre-existing toolbox for the automatic phonemic segmentation of speech was not used to segment boundaries. Segment boundaries were determined by using a dynamic time warping method (DTW) that compared each recorded syllable utterance to a target template for that utterance (e.g., Holmes & Holmes, 2001).

Templates were created by manually identifying phoneme boundaries in the acoustic model of each target syllable. Each acoustic waveform was visually and aurally inspected and segmented in software implemented in MATLAB. Acoustic cues were identified using a set of criteria based on guidelines for acoustic phonetic analysis (Ladefoged, 2012). Examples of illegal and illegal syllable template are shown in Fig. 1. The templates were inspected by multiple raters and consensus of phoneme boundaries was achieved. Segmentation criteria for consonants by manner of articulation and for vowels were as follows:

Plosive: The onsets of plosives were labeled at the beginning of a period of silence, visualized as a ‘gap’ in spectral energy corresponding to a vocal tract occlusion. The offset of plosives were labeled at the end of a burst of spectral energy representing a release phase before the beginning of the following formant structure. The onsets of plosives that occurred at the beginning of syllables (e.g. BLERK) were labeled at the beginning of the energy burst.

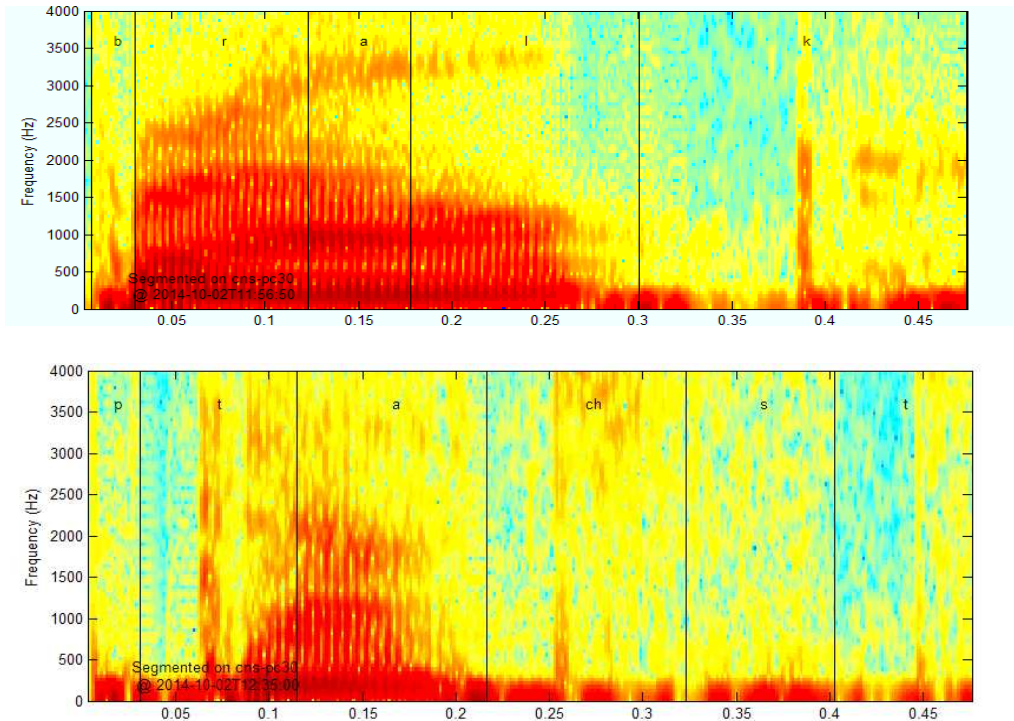


Fig 1. Segmented spectrogram templates for legal syllable BRALK (top) and illegal syllable PTASCHT (bottom). The black vertical bars mark phoneme onsets. Orthographic labeling of each phoneme is represented at the top of each spectrogram.

Nasal: The onsets of nasals were labeled at the beginning and ending of areas of little or no spectral energy characterized by faint formant structures and an overall reduction in amplitude. A distinguishing acoustic feature of nasals included blank regions between formants. Nasal offsets were marked by changes in amplitude in the spectrogram.

Fricative: Fricative onsets were labeled at the beginning of high amplitude bursts of scattered spectral energy corresponding to air traveling through a constricted opening in the vocal tract. The most distinguishing features of fricatives were random spectral energy distributed over higher frequencies. Fricatives were identified by their average frequencies, with sibilant (e.g. /s/) and nonsibilant phonemes (e.g. /f/) visualized at higher

and lower frequencies, respectively. Fricative offsets were marked by a change in high amplitude, randomly scattered spectral energy. Onsets of affricate phonemes, plosive followed by fricative combinations, were labeled at the beginning period of silence (‘gap’) in spectral energy. The offset of affricate phonemes were labeled when there was a change in change in high amplitude, randomly scattered spectral energy.

Approximant: Approximant transitions were distinguished by clear formant patterns lower in amplitude relative to vowel formant patterns. Approximant onsets were identified by the movement or narrowing of formants. The onset and offset were labeled at the rising movement or the second formant. The onset of /r/ was labeled at the rising movement of the third formant. In vowel-/r/ transitions, the onset was labeled at the lowering of the third formant.

Vowel: Vowel transitions were labeled at the beginning and end of clearly distinguishable F1, F2, and F3 formant patterns.

Phoneme boundaries were identified by using the DTW algorithm, implemented in MATLAB, to optimally align that appropriate template spectrogram with the spectrogram of each recorded utterance. DTW is well-suited for this application because it is able to determine the similarity of two time-series that vary in both total and inter-segmental duration. Segmented spectrograms manually inspected by a trained rater and edited as needed. A subset (approximately 15%) of boundaries were adjusted following the same criteria established for labeling the target templates. The most common boundary error included delayed onset of consonant boundaries. The accuracy of the DTW alignment

was also affected by subject performance and recording quality. Errors were noted on schwa epenthesized clusters and prolonged phonemes. Low signal to noise ratio and high distortion also resulted in errors in alignment. Results of the DTW algorithm were used to obtain onset, coda, and rime durations of the *learned illegal* and *novel illegal* syllables in the fMRI session.

fMRI Data Analysis

The Nipype (Gorgolewski et al., 2011) neuroimaging software interface, which permits the use of preferred processing routines from various neuroimaging analysis packages, was used to analyze imaging data. Using SPM8 image processing tools (<http://www.fil.ion.ucl.ac.uk/spm/software/spm8>), functional images were motion-corrected and realigned to the subject's anatomical volume and high-pass filtered with a standard 128 s cutoff frequency.

We estimated BOLD responses using a general linear model (GLM). The response of each stimulus event was modeled as a finite impulse response. The model included 4 condition-specific variables – *learned illegal*, *novel illegal*, *learned legal*, and *baseline* - and additional covariates: utterance duration measures, linear detrending covariates, and motion parameters. The model was estimated within each subject and estimates for the *novel illegal* and *learned illegal* conditions were contrasted (*novel illegal - learned illegal*) at each voxel. Group statistics were calculated separately for cortical and subcortical regions. Surface-based analysis was used to assess group BOLD response differences in the novel illegal and learned illegal conditions in the cerebral cortex. T1 volume

segmentation and cortical surface reconstruction for each subject were performed with the FreeSurfer image analysis suite (Dale, Fischl, & Sereno, 1999; Fischl et al., 2002; Fischl, Sereno, & Dale, 1999). The activity of cortical voxels in the *novel illegal - learned illegal* contrast volume for each subject were mapped to that subject's cortical surface. Subject data were aligned by inflating each individual surface to a sphere and registering it to a template representing the average surface curvature of a set of neurologically normal adult brains (Fischl, Sereno, Tootell, & Dale, 1999). The surface-based contrast data were smoothed with a 6 mm full-width half-maximum (FWHM) kernel and then averaged across subjects and group-level t-statistics were calculated at each vertex. Vertex-wise statistics were thresholded at $p < 0.001$ (uncorrected) and cluster-level significance thresholds were then estimated separately for each hemisphere using a Monte Carlo simulation over 10,000 iterations in which each iteration measured the maximum cluster size in smoothed random noise data (Hayasaka & Nichols, 2003). Results were cluster-thresholded in each hemisphere at cluster-wise probability (CWP) < 0.0167 to correct for both surface-based tests in each hemisphere and one subcortical volume-based test.

Neural activity underlying speech motor sequence learning was examined by comparing BOLD responses in the *novel illegal - learned illegal* contrast across all subjects using a vertex-wise paired two sample t-test. Between-groups differences in learning effects were also examined by comparing the mean *novel illegal - learned illegal* contrast of the two groups with a two-sample t-test. To increase power, an additional region-of-interest (ROI) based comparison of the group responses was also performed. For each subject, the

novel illegal - learned illegal response was averaged within each cluster of significant activity determined by vertex-wise contrast of all subjects. Differences in the mean response for each cluster (or ROI) in the PWS and PFS groups were tested with a two-sample t-test; significance thresholds were corrected to ensure a false discovery rate < 5%.

To assess brain regions associated with learning success, each learning measures was correlated with the *novel illegal - learned illegal* BOLD responses within each ROI using Pearson's r.

RESULTS

Practice Session Behavioral Measures of Learning

Measures of error, syllabic utterance duration, and reaction time between practice sessions are shown in Fig 2. For both PWS and PFS practice resulted in a significant increase in accuracy for the *learned illegal* syllables ($p < 0.01$, $t(16) = 6.09$, and $p < 0.01$, $t(13) = 5.52$, respectively) but not for the *legal* syllables, ($p > 0.12$, $p > 0.07$). This is reflected in a Day x Condition interaction effect for accuracy that was just above significant level ($F(1,117) = 3.79$, $p > 0.05$). In PFS, the increase in accuracy for the *illegal* condition was significantly higher than that of the *legal* condition ($p < 0.01$, $t(32) = 4.51$), but this difference did not reach significance in PWS ($p > 0.07$). No main effect of group on accuracy was found ($F(1,58) = 3.34$, $p > .073$) nor was there Condition x Group interaction ($F(1, 57) = 0.87$, $p > 0.35$). A Day x Group interaction for accuracy was also not significant ($F(1,57) = 0.31$, $p > 0.37$). So there was no evidence for a group difference in learning as indicated by a change in accuracy due practice. Group differences in accuracy rates, however, were found. PWS exhibited a higher error rate in the *illegal* condition in both practice session 1 ($p < 0.05$, $t(28) = 2.39$) and practice session 2 ($p < 0.05$, $t(29) = 2.41$) compared to PFS.

Reaction time between practice sessions did not differ significantly between conditions ($F(1,57) = 1.73$, $p = 0.19$, n.s.) nor group ($F(1,57) = 0.5$, $p = 0.48$, n.s.). There was no significant Group x Condition interaction effect for reaction time ($F(1,57) = 0.05$, $p = 0.08$, n.s.). Similarly, syllabic utterance duration did not significantly change from

practice session 1 to practice session 2 for Condition ($F(1,57) = 0.24, p = 0.62, n.s.$) nor Group ($F(1,57) = 0.08 (p=0.77, n.s.)$). The syllabic utterance duration Group x Condition interaction was also not significant ($F(1,57) = 0.08 p=0.77, n. s.$).

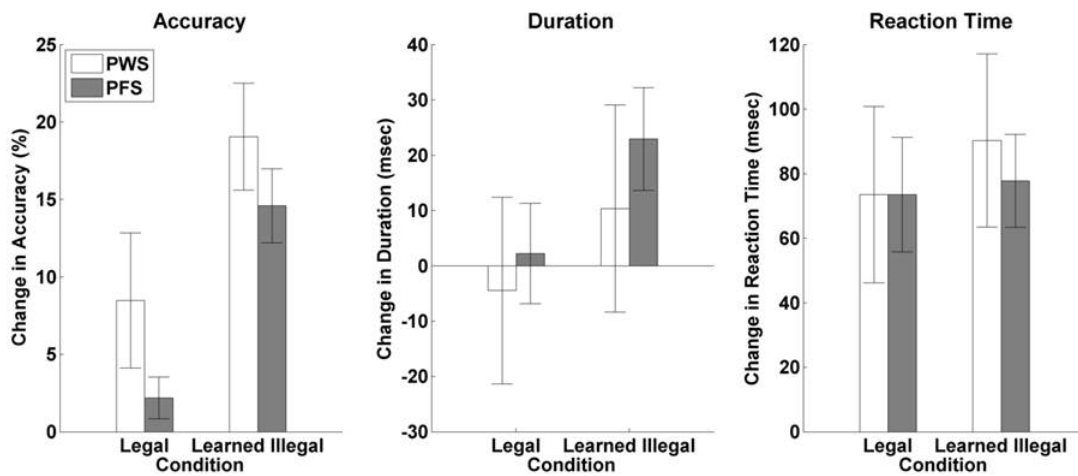


Fig. 2. Changes in behavioural learning indices between practice session 1 and practice session 2. Both PWS (white) and PFS (grey) produced fewer errors for learned illegal syllables in practice session 2 compared to practice session 1 (left). There were no significant differences between practice sessions for syllabic utterance duration (middle) and reaction time (right) across Groups or Conditions.

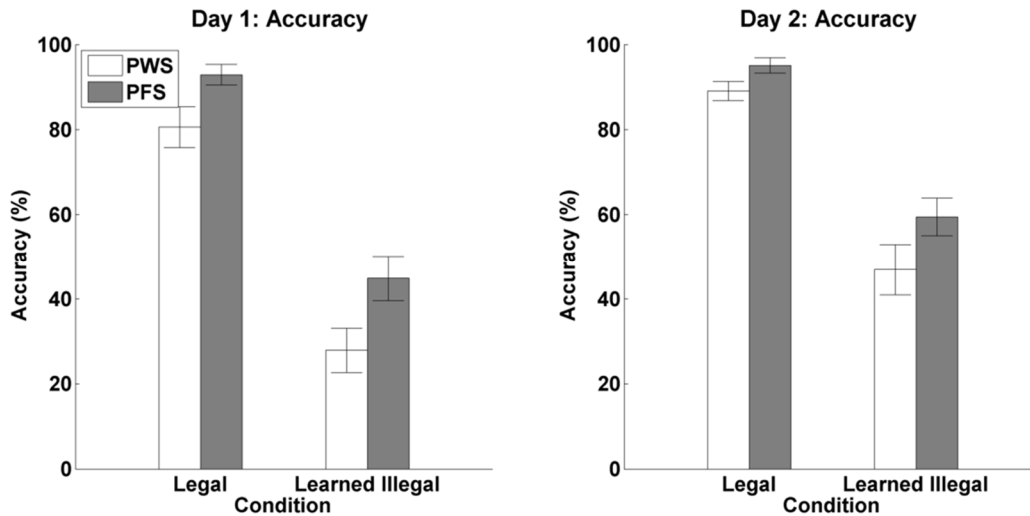


Fig 3. Percent accuracy for PWS (white) compared to PFS (grey) in practice session 1 (left) and practice session 2 (right). Error bars represent standard error.

fMRI Behavioral Measures of Learning

In the fMRI session, significant differences in error across conditions were found ($F(2, 93) = 49.73, p < 0.001$) (Fig 4.) Post-hoc paired t-tests compared pairs of conditions. Subjects demonstrated significantly more errors for the *learned illegal* syllables compared to the *legal* syllables ($t(64) = 8.4, p < 0.01$) and for *novel illegal* syllables compared to *legal* syllables ($t(64) = -10.7, p < 0.01$). The difference in percentage of errors noted between the *learned illegal* vs. *novel illegal* condition ($t(22) = 0.5, p = 0.6, n.s.$), the number of errors exhibited by PWS and PFS were not significantly different ($F(1, 93) = 2.2 (p = 0.14, n.s.)$) and a Group x Condition interaction for accuracy did not reach significance ($F(2, 93) = 0.33, p = 0.72, n.s.$).

Reaction time did not differ significantly between conditions ($F(2, 93) = 0.17, p = 0.842, n.s.$) but did differ significantly between the groups ($F(1, 93) = 17.77, p < 0.001$) with PWS

demonstrating longer reaction times than PFS (Fig. 5). The Group x Condition interaction effect was not significant ($F(2,93) = 0.35, p = 0.70, n.s.$). Syllabic utterance duration did not differ significantly between groups ($F(1,93) = 0.5, p = 0.48, n.s.$) nor conditions ($F(2,93) = 1.57, p = 0.08, n.s.$) Similarly, the Group x Condition interaction was not significant, $F(2, 93) = 0.06, p = 0.94, n.s.$ As expected, PWS demonstrated more dysfluencies compared to the PFS group, $F(1, 93) = 25.93, p < 0.001$ (Fig. 6) but no main effect of Condition ($F(2,93) = 1.32, p = 0.27, n.s.$) nor a Group x Condition interaction was noted ($F(2,98) = 1.02, p = 0.34, n.s.$).

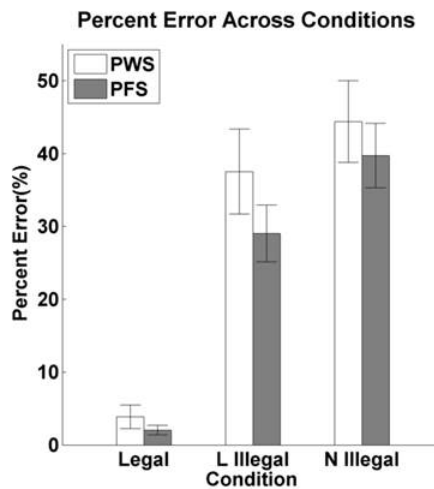


Fig. 4. Percentage of errors for PWS (white) compared to PFS (grey) in the fMRI session. Error bars represent standard error. Subjects produced more errors for the learned illegal syllables compared to the legal syllables and for the novel illegal syllables compared to the legal syllables.

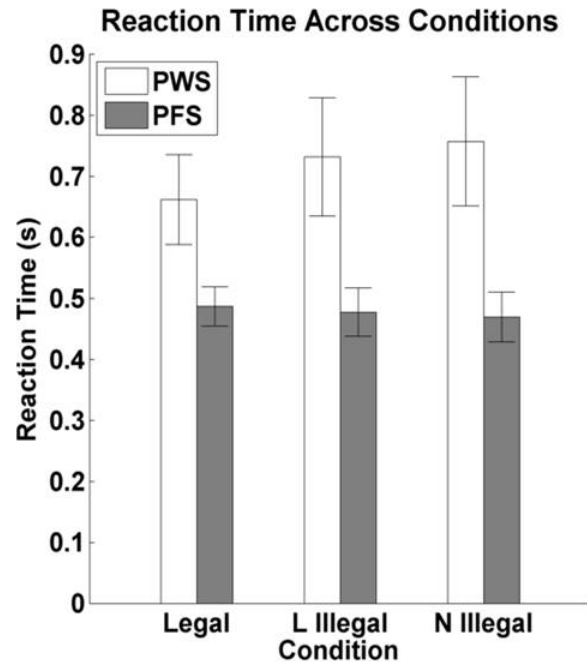


Fig 5. Reaction Time for PWS (white) compared to PFS (grey) in the fMRI session. Error bars represent standard error.

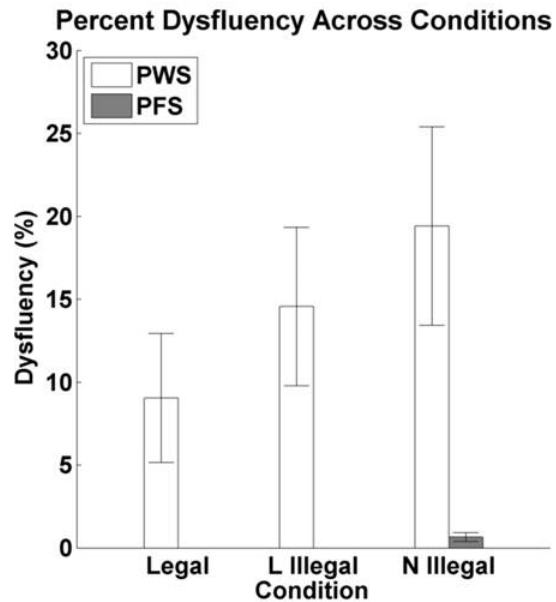


Fig 6. Percent Dysfluency of PWS (white) compared to PFS (grey) in the fMRI session. Error bars represent standard error. PWS produced significantly more dysfluencies compared to PFS.

fMRI Syllable Segment Duration Measures

Mean durations of syllable onset and coda productions in the fMRI session, expressed relative to total syllabic duration, are shown in Fig. 7. The PWS group demonstrated significantly longer onset duration than PFS ($F(1, 60) = 15.12, p < 0.001$). No effect of condition ($F(2,60) = 1.34, p = 0.27, n.s.$) on syllable onset was noted and no significant Group x Condition interaction was found ($F(2,60) = 0.3, p = 0.7, n.s.$). PFS demonstrated significantly longer codas than PWS ($F(1, 60) = 7.19, p < 0.01$), however, across conditions, $F(2, 60) = 5.94, p < 0.01$. Post-hoc t-tests indicated significantly longer coda durations for *legal* syllables compared to the *learned illegal* syllables ($t(22) = 3.11, p < 0.01$) and the *legal* syllables compared to the *novel illegal* syllables ($t(22) = 2.86, p < 0.001$) but coda

duration did not differ between the *learned illegal* syllables and *novel illegal* syllables ($t(22) = 0.52, p=0.60, n.s.$). No Group x Condition interaction for coda duration was found ($F(2,60) = 0.04, p = 0.96, n.s.$).

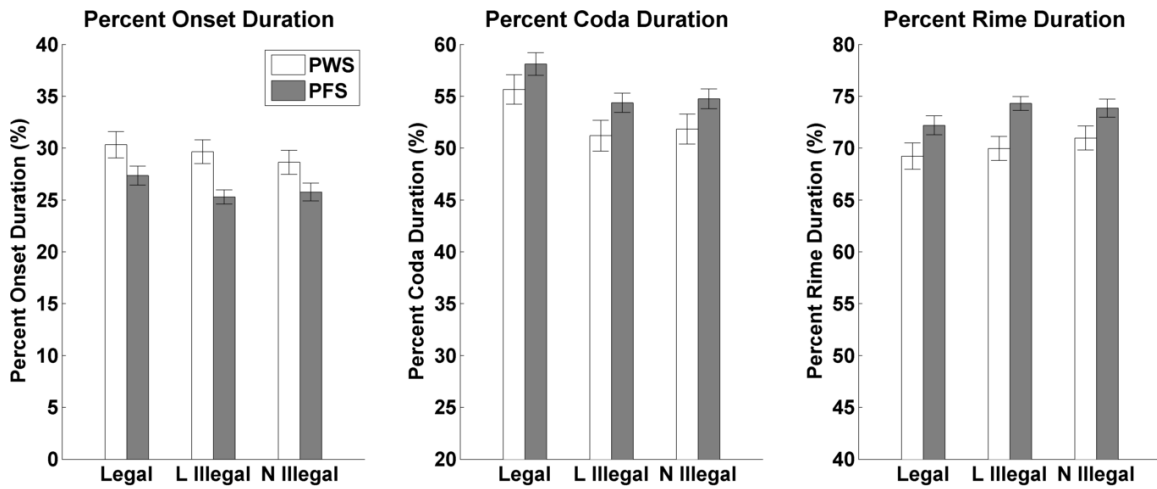


Fig. 7. Segment duration results from syllable productions in the fMRI session.

fMRI Analysis

Fig. 8 shows the cortical regions that were significantly more active during production of *novel illegal* than *learned illegal* syllables for all subjects (voxel-level $p < 0.001$, CWP < 0.01667). The production of *novel illegal* syllables resulted in greater BOLD response in left hemisphere anterior insula, premotor cortex (2 clusters), posterior superior temporal gyrus, inferior temporal-occipital cortex, superior parietal lobule, and pre-supplementary motor cortex. Activation in right hemisphere anterior insula, and superior parietal lobule was also found.

A vertex-wise comparison of the *novel – illegal – learned illegal* contrast in the two groups revealed no significant differences. However, ROI analyses revealed significantly greater activation in the posterior dorsal superior temporal sulcus in PWS compared to PFS in the production of *novel illegal syllables* compared to *illegal syllables* ($t= 2.42$, $p<0.05$).

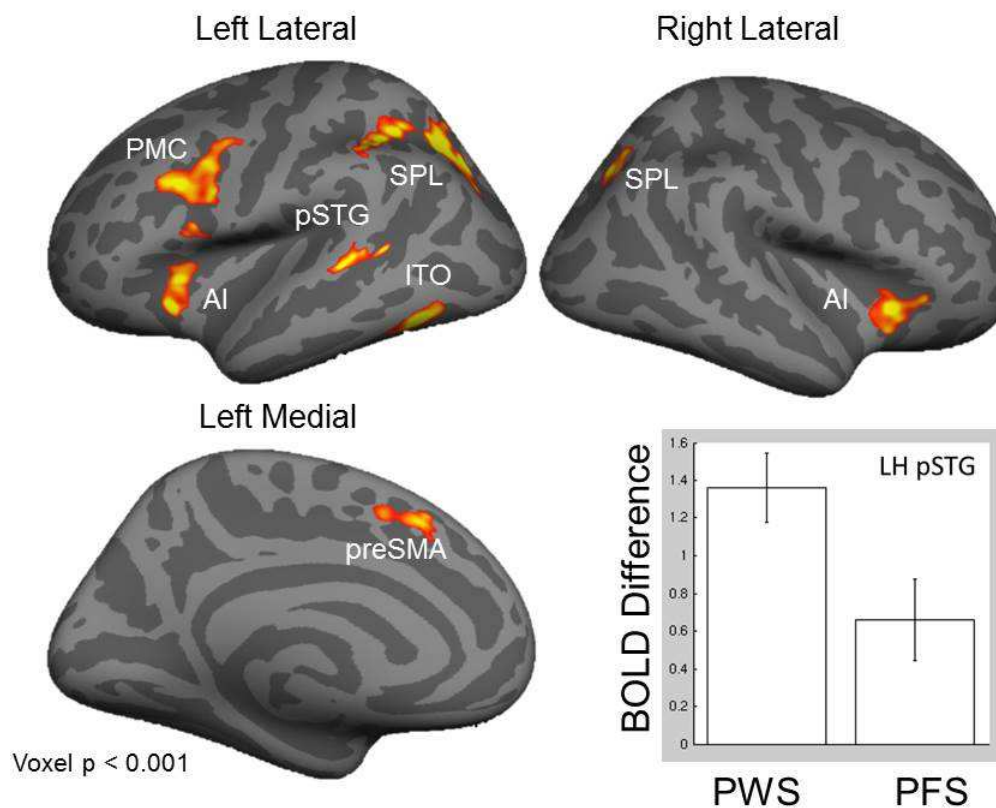


Fig. 8. BOLD responses in the novel illegal-learned illegal contrast for all subjects (CWP < 0.01667) Abbreviations: AI = anterior insula, PMC = premotor cortex, pSTG = posterior superior temporal gyrus, ITO = inferior temporal-occipital cortex, SPL = superior parietal lobule, preSMA = presupplementary motor cortex.

Neural-Behavioral Correlation Analysis

Learning success, as measured by the subject-normalized difference in percent error between *novel illegal* and *learned illegal* syllables, was positively correlated with the mean BOLD response difference in left ventral premotor cortex (dorsal) ($r = 0.55$, $p < 0.05$) in the same two conditions. No other measures of learning were found to correlate significantly with this BOLD contrast in the regions tested.

DISCUSSION

In summary, we found evidence of speech motor sequence learning over two days of practice in both PWS and PNS and identified the neural correlates of learning new speech motor sequences. Both groups demonstrated significant behavioral gains from practice session 1 to practice session 2, which is indicative of learning. Learning was associated with reduced activity in a network of regions that was similar to one found in a previous study of speech motor sequence learning (Segawa, Tourville, Beal, Guenther, 2015) in fluent speakers. In contrast to our hypothesis, we did not find evidence of speech motor sequence learning deficits in PWS; however a significant speech motor performance deficit in PWS was observed.

The significant behavioral gains in accuracy over the practice sessions that were noted is indicative of speech motor sequence learning. Subjects were able to produce legal syllables more accurately than illegal syllables during both practice sessions. Over practice days, error rates decreased for the illegal syllables; however, only PFS demonstrated a significantly greater decrease in error rate for illegal syllables (e.g. gbesb) versus legal syllables (e.g. bralk). Executing practiced movements with higher accuracy over repeated trials is a hallmark of motor learning (Schmidt & Wrisberg, 2004). In early stages of speech motor learning, performance is more variable and prone to error, requiring conscious attentional and sensory demands. With repeated practice, performance becomes less variable and more accurate (Ackerman, 2007; Willingham, 1998). Through repeated practice of phonotactically illegal syllables, subjects were able to produce sequences more effortlessly and accurately. Unexpectedly, PWS showed a

greater mean increase in accuracy in both the legal and learned illegal conditions compared to PFS. Though insignificant, this would suggest that PWS were better learners of novel speech sequences than PFS. It is most likely that PWS exhibited a higher degree of learning due to ceiling effects and the improved accuracy rate was due to a lower accuracy in both conditions at the onset of practice.

Our RT findings are consistent with some previous studies of motor speech sequence learning that found that PWS are slower to execute and exhibit longer newly learned nonspeech and speech motor sequences (Smits-Bandstra, DeNil, & Saint-Cyr, 2006; Bauerly & DeNil, 2011). However, it has been pointed out that behavioral learning indices such total duration may not be sensitive enough to capture group differences in motor sequence learning and/or motor control strategies (Peters, Hulstijn, & Van Lieshout, 2000; Smits-Bandstra, DeNil, & Saint-Cyr, 2006). Therefore, subtle group differences in speech motor learning may not have been captured by RT and duration measures.

Unexpectedly, the behavioral gains established in the practice sessions were not observed in the imaging session. We expected learning to remain stable across practice days and a consolidation period of 1–2 days before the imaging session. While subjects did produce learned illegal syllables more accurately than similar novel illegal syllables during the fMRI sessions, this difference was not significant. The lack of a significant difference in accuracy for the novel and learned illegal syllables could be the result of a lack of power due to the removal of trials due to image or audio recording quality or dysfluent or otherwise erroneous productions. Furthermore, we may have not seen a significant

difference in accuracy between novel and learned illegal syllables in the fMRI session, as we did not differentiate high proficiency learners from low proficiency learners based on the practice session data. Therefore, subjects were included in the imaging session regardless of how well or poorly they learned in the practice sessions. The lack of a difference in RT or and total duration in the novel and illegal syllable conditions for either group is consistent with the practice session results which found no effect of practice on these measures. .

PWS did exhibit significantly longer reaction times across conditions in the neuroimaging session. This finding is consistent with previous studies of motor sequence learning in PWS that have found differences in nonspeech and speech sequence initiation (Van Lieshout, Hulstijn, & Peters, 1996; Smits-Bandstra, DeNil & Rochon, 2006; Smits-Bandstra, DeNil, & St-Cyr, 2006). Potentially, PWS were slower to access the production of the initial motor programs in the syllable sequence, resulting in longer RTs. Another explanation for a significant group difference in RT could be differences in the preferred strategic approach to the task in the imaging session. Longer RTs across all conditions suggests that PWS were more cautious in their approach. Recall that PWS exhibited higher rates of dysfluency in the neuroimaging session compared to PFS. It is plausible that PWS were more careful in their approach to the task to avoid dysfluencies, resulting in overall longer reaction times.

Contrary to our hypothesis, the two groups did not differ significantly on total utterance duration measures. This finding is consistent with previous studies that found no differences in the syllable duration of learned speech sequences between PFS and PWS

(Bauerly & DeNil, 2011; Smits-Bandstra, DeNil, & St-Cyr, 2006; Navisavayam & Van Lieshout, 2008). However, measures of sub-syllabic duration did show significant group differences in the timing of syllable onsets and codas across conditions. PWS exhibited longer onset durations and shorter coda compared to PFS.

As expected, we found a speech motor sequence learning network that was less active in the production of novel illegal syllables vs. learned illegal syllables in both groups of speakers. The network identified in our study included left hemisphere anterior insula, premotor cortex, posterior superior temporal gyrus), inferior temporal-occipital cortex, superior parietal lobule, and pre-supplementary motor cortex. Activation in right hemisphere anterior insula, and superior parietal lobule was also found. This network is nearly identical to a speech motor sequence learning network that was identified in a previous study examining learning novel speech sequences in PFS (Segawa, Tourville, Beal, & Guenther, 2015). The authors found that speech motor learning was associated with decreased activation in frontal operculum-anterior insula, posterior superior temporal cortex, premotor cortex, preSMA/SMA, basal ganglia, and cerebellum. These areas are not only related to speech production, but also involved in motor sequencing learning processes. In the present study, only cortical differences were investigated, therefore, it is unknown if there were differences in activity in basal ganglia, cerebellum, or other subcortical areas.

We expected to find significant between-group differences in activity within the speech motor sequence network in the novel illegal condition compared to the learned illegal condition due to differences in learning. As described above, we did not see significant

differences in learning between the two groups and, consequently, there was little difference in the *novel illegal – learned illegal* BOLD contrast for the two groups. ROI analyses did reveal greater activity within the poster superior temporal gyrus in PWS, however. This increase in activation could be related to the higher error rate and greater dependence on auditory feedback in PWS. During speech production, the posterior superior temporal gyrus plays an important role in monitoring and refining motor programs for speech movements based on auditory feedback (Guenther et al., 2006; see also Hashimoto & Sakai, 2003; Hirano et al., 1997; Indefrey, 2011; Indefrey & Levelt, 2004). It is plausible that PWS relied on increased monitoring of auditory feedback during the production of novel illegal syllables, resulting in significantly greater activity in the posterior superior temporal gyrus.

The difference in BOLD response in dorsal PMC during the novel and learned illegal conditions was significantly correlated with learning success measured as the normalized difference in accuracy in the novel and learned illegal conditions. These findings are inconsistent with the study conducted by Segawa et al. (2015) that found learning success, as measured by utterance duration, was correlated with a larger difference in activity in FO in the novel illegal-learned illegal contrast. The different learning indicator noted in the two studies may be due to how participants were trained to complete the task. In Segawa et al.'s (2015) study, participants underwent brief pre-practice training to eliminate schwa insertions in the production of illegal consonant clusters, reducing the likelihood of this type of error. Segawa et al. also instructed subjects to produce the syllables as quickly as possible. These instructions were not given in the current study

resulting in a larger total error rate and a de-emphasis on movement speed. Another difference includes learning success being positively correlated with activity in dorsal PMC. The dorsal PMC is connected to the posterior superior temporal gyrus and surrounding regions which play a role in error monitoring. Potentially, this region was activated while receiving error signals from posterior superior temporal gyrus.

Clinically, focus has been placed on remediating speech motor patterns in persons who stutter (PWS) to facilitate fluency. Intervention approaches focus on the establishment of new speech motor patterns and include techniques such as rate modification, prolonged speech, stretched syllables, light articulatory contacts, and slowed transitions between vowels and consonants. PWS are also taught how to coordinate their breath, voice, and articulation to achieve fluency (Blomgren, 2010; Carey, 2010; Cocomazzo, 2012, O'Brian, 2003; Onslow, Costa, Andrews, Harrison, & Packman, 1996). Overall, effective fluency intervention is largely dependent on motor skill learning abilities. Our findings suggest that PWS that the underlying deficit in PWS is associated with the speech motor production system rather than the speech motor learning system. Deficits in the speech motor production system would have implications for stuttering intervention methods. If the speech motor learning system is intact, repeated practice of new motor skills may not be as facilitative of fluency as modifying sensory feedback during speech production in PWS. Fluency intervention could emphasize motor performance versus speech motor sequence learning.

In summary, our findings showed behavioral gains as a result of speech motor sequence learning and identified a network of cortical areas involved in learning in PWS and PFS.

For both groups, learning was associated with reduced activity within a motor sequence learning network that included frontal and posterior superior temporal cortex. ROI analyses did reveal greater activity within the poster superior temporal gyrus in PWS, which is suggestive of greater reliance on auditory feedback in this population. For both groups, learning success was significantly correlated with activity in dorsal premotor cortex, suggesting that this region receives error feedback from the posterior superior temporal gyrus during speech.

REFERENCES

- Alm, P. (2004). Stuttering and the basal ganglia circuits: a critical review of possible relations. *Journal of communication disorders*, 37(4), 325–69. doi:10.1016/j.jcomdis.2004.03.001
- Bauerly, K. R., & De Nil, L. F. (2011). Speech sequence skill learning in adults who stutter. *Journal of fluency disorders*, 36(4), 349–60. doi:10.1016/j.jfludis.2011.05.002
- Belin, P., Zatorre, R. J., Hoge, R., Evans, A. C., & Pike, B. (1999). Event-related fMRI of the auditory cortex. *NeuroImage*, 10(4), 417–429.
- Bigi, B. (2012). SPPAS: a tool for the phonetic segmentations of Speech. In Proceedings of the eighth international conference on Language Resources and Evaluation (pp. 1748–1755).
- Blomgren, M. (2010, November). Stuttering treatment for adults: An update on contemporary approaches. In *Seminars in speech and language* (Vol. 31, No. 4, p. 272).
- Bohland, J. W., Bullock, D., & Guenther, F. H. (2010). Neural representations and mechanisms for the performance of simple speech sequences. *Journal of cognitive neuroscience*, 22(7), 1504–29. doi:10.1162/jocn.2009.21306
- Bohland, J. W., & Guenther, F. H. (2006). An fMRI investigation of syllable sequence production. *NeuroImage*, 32(2), 821–41. doi:10.1016/j.neuroimage.2006.04.173
- Büchel, C., & Sommer, M. (2004). What causes stuttering? *PLoS biology*, 2(2), e46.
- Carey, B., O'Brian, S., Onslow, M., Block, S., Jones, M., & Packman, A. (2010). Randomized PFSled non-inferiority trial of a telehealth treatment for chronic stuttering: the Camperdown Program. *International journal of language & communication disorders*, 45(1), 108–20. doi:10.3109/13682820902763944
- Civier, O., Bullock, D., Max, L., & Guenther, F. H. (2013). Computational modeling of stuttering caused by impairments in a basal ganglia thalamo-cortical circuit involved in syllable selection and initiation. *Brain and language*, 126(3), 263–78. doi:10.1016/j.bandl.2013.05.016
- Coltheart, M. (1981). The MRC psycholinguistic database. *Quarterly journal of experimental psychology*, 33(4), 497–505.

- Cocomazzo, N., Block, S., Carey, B., Brian, S. O., Onslow, M., & Packman, A. (2012). Camperdown Program for adults who stutter : a student training clinic Phase I trial, 365–372. doi:10.1111/j.1460-6984.2012.00150.x
- Cunnington, R., Windischberger, C., Deecke, L., & Moser, E. (2002). The preparation and execution of self-initiated and externally-triggered movement: a study of event related fMRI. *NeuroImage*, 15(2), 373–85. doi:10.1006/nimg.2001.0976
- Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical surface-based analysis - I. Segmentation and surface reconstruction. *NeuroImage*, 9(2), 179–194.
- Doyon, J., Bellec, P., Amsel, R., Penhune, V., Monchi, O., Carrier, J., ... Benali, H. (2009). Contributions of the basal ganglia and functionally related brainstructures to motor learning. *Behavioural brain research*, 199(1), 61–75. doi:10.1016/j.bbr.2008.11.012
- Fischl, B., Salat, D. H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., et al. (2002). Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain. *Neuron*, 33(3), 341–355.
- Fischl, B., Sereno, M. I., & Dale, A. M. (1999). Cortical surface-based analysis - II: Inflation, flattening, and a surface-based coordinate system. *NeuroImage*, 9(2), 195–207.
- Fischl, B., Sereno, M. I., Tootell, R. B. H., & Dale, A. M. (1999). High-resolution intersubject averaging and a coordinate system for the cortical surface. *Human brain mapping*, 8(4), 272–284.
- Gorgolewski, K., Burns, C. D., Madison, C., Clark, D., Halchenko, Y. O., Waskom, M. L., et al. (2011). Nipype: a flexible, lightweight and extensible neuroimaging data processing framework in python. *Frontiers in neuroinformatics*, 5, 13.
- Guenther, F. H. (2006). Cortical interactions underlying the production of speech sounds. *Journal of communication disorders*, 39(5), 350–65. doi:10.1016/j.jcomdis.2006.06.013
- Hall, D. A., Haggard, M. P., Akeroyd, M. A., Palmer, A. R., Summerfield, A. Q., Elliott, M. R., et al. (1999). "Sparse" temporal sampling in auditory fMRI. *Human brain mapping*, 7(3), 213–223.
- Hayasaka, S., & Nichols, T. E. (2003). Validating cluster size inference: random field and permutation methods. *NeuroImage*, 20(4), 2343–2356.

- Hikosaka, O., Sakai, K., Miyauchi, S., Takino, R., Sasaki, Y., & Pütz, B. (1996). Activation of human presupplementary motor area in learning of sequential procedures: a functional MRI study. *Journal of neurophysiology*, 76(1), 617–21. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8836248>
- Hikosaka, Okihide, Nakamura, K., Sakai, K., & Nakahara, H. (2002). Central mechanisms of motor skill learning. *Current opinion in neurobiology*, 12(2), 217–222. doi:10.1016/S0959-4388(02)00307-0
- Holmes, W. (2001). Speech synthesis and recognition. CRC Press.
- Indefrey P., Levelt W. (2004). The spatial and temporal signatures of word production components. *Cognition* 92, 101–144. doi:10.1016/j.cognition.2002.06.001
- Ingham, R. J., Wang, Y., Ingham, J. C., Bothe, A. K., & Grafton, S. T. (2013). Regional brain activity change predicts responsiveness to treatment for stuttering in adults. *Brain and language*, 127(3), 510–9. doi:10.1016/j.bandl.2013.10.007
- Jouen, A.-L., Verwey, W. B., van der Helden, J., Scheiber, C., Neveu, R., Dominey, P. F., & Ventre-Dominey, J. (2013). Discrete sequence production with and without a pause: the role of cortex, basal ganglia, and cerebellum. *Frontiers in human neuroscience*, 7(September), 492. doi:10.3389/fnhum.2013.00492
- Ladefoged, P. (2012). Vowels and Consonants. John Wiley & Sons.
- Lu, C., Chen, C., Ning, N., Ding, G., Guo, T., Peng, D., ... & Lin, C. (2010). The neural substrates for atypical planning and execution of word production in stuttering. *Experimental neurology*, 221(1), 146–156.
- Ludlow, C. L., & Loucks, T. (2004). Stuttering: a dynamic motor PFS disorder. *Journal of fluency disorders*, 28(4), 273–295.
- Max, Ludo, Anthony J. Caruso, and Vincent L. Gracco. Kinematic analyses of speech, orofacial nonspeech, and finger movements in stuttering and nonstuttering adults. *Journal of speech, language, and hearing research*, 46.1 (2003): 215–32..
- Nakamura, K., Sakai, K., & Hikosaka, O. (1998). Neuronal activity in medial frontal cortex during learning of sequential procedures. *Journal of neurophysiology*, 80(5), 2671–2687.
- Namasivayam, A. K., & van Lieshout, P. (2008). Investigating speech motor practice and learning in people who stutter. *Journal of fluency disorders*, 33(1), 32–51. doi:10.1016/j.jfludis.2007.11.005

- O'Brian, S., Onslow, M., Cream, A., & Packman, A. (2003). The Camperdown Program: outcomes of a new prolonged-speech treatment model. *Journal of speech, language & hearing research*, 46(4).
- Onslow, M., Costa, L., Andrews, C., Harrison, E., & Packman, A. (1996). Speech outcomes of a prolonged-speech treatment for stuttering. *Journal of speech and hearing research*, 39(4), 734.
- Peeva, M. G., Guenther, F. H., Tourville, J. a, Nieto-Castanon, A., Anton, J.-L., Nazarian, B., & Alario, F.-X. (2010). Distinct representations of phonemes, syllables, and supra-syllabic sequences in the speech production network. *NeuroImage*, 50(2), 626–38. doi:10.1016/j.neuroimage.2009.12.065
- Peters, H. M., Hulstijn, W., & Van Lieshout, P. H. (1999). Recent developments in speech motor research into stuttering. *Folia phoniatrica et logopaedica*, 52(1–3), 103–119.
- Rauschecker, A. M., Pringle, A., & Watkins, K. E. (2008). Changes in neural activity associated with learning to articulate novel auditory pseudowords by covert repetition. *Human brain mapping*, 29(11), 1231–42. doi:10.1002/hbm.20460
- Robb, M. & Blomgren, M. (1997). Analysis of F2 transitions in the speech of stutters and nonstutters. *Journal of fluency disorders*, 22, 1–16.
- Robb, M., Blomgren, M., & Chen Y. (1998). Formant frequency fluctuation in stuttering and nonstuttering adults. *Journal of fluency disorders*, 23 (1), 73–84.
- Sasisekaran, J. & Wesiberg, S. (2014). Practice and retention of nonwords in adults who stutter. *Journal of fluency disorders*, 41, 55–71.
- Shuster, L. I., & Lemieux, S. K. (2005). An fMRI investigation of covertly and overtly produced mono- and multisyllabic words. *Brain and language*, 93(1), 20–31. doi:10.1016/j.bandl.2004.07.007
- Smith, A., Sadagopan, N., Walsh, B., & Weber-Fox, C. (2010). Increasing phonological complexity reveals heightened instability in inter-articulatory coordination in adults who stutter. *Journal of fluency disorders*, 35(1), 1–18.
- Smits-Bandstra, S., & De Nil, L. (2009). Speech skill learning of persons who stutter and fluent speakers under single and dual task conditions. *Clinical linguistics & phonetics*, 23(1), 38–57. doi:10.1080/02699200802394914

- Smits-Bandstra, S., & De Nil, L. F. (2007). Sequence skill learning in persons who stutter: implications for cortico-striato-thalamo-cortical dysfunction. *Journal of fluency disorders*, 32(4), 251–278. doi:10.1016/j.jfludis.2007.06.001
- Smits-Bandstra, S., De Nil, L. F., & Saint-Cyr, J. a. (2006). Speech and nonspeech sequence skill learning in adults who stutter. *Journal of fluency disorders*, 31(2), 116–36. doi:10.1016/j.jfludis.2006.04.003
- Smits-Bandstra, S., De Nil, L., & Rochon, E. (2006). The transition to increased automaticity during finger sequence learning in adult males who stutter. *Journal of fluency disorders*, 31(1), 22–42; quiz 39–40. doi:10.1016/j.jfludis.2005.11.004
- Sternberg, S., Monsell, S., Knoll, R. L., & Wright, C. E. (1978). The latency and duration of rapid movement sequences: Comparisons of speech and typing. In G. E. Stelmach (Ed.), *Information Processing in Motor PFS and Learning* (pp. 117–152). New York: Academic Press.
- Toni, I., Krams, M., Turner, R., & Passingham, R. E. (1998). The time course of changes during motor sequence learning: a whole-brain fMRI study. *NeuroImage*, 8(1), 50–61. doi:10.1006/nimg.1998.0349
- Van Lieshout, P.H.H.M., Hulstijn, W., & peters H.F.M (1996a). From planning to articulation in speech production: What differentiates a person who stutters from a person who does not stutter. *Journal of speech and hearing research*, 39, 546–564.

VITA

ANNA OH

Education

B.S. University of Toronto
M.S. Boston University

2011
2015