Brain & Language 149 (2015) 97–105

Contents lists available at ScienceDirect

Brain & Language

journal homepage: www.elsevier.com/locate/b&l

Speech preparation in adults with persistent developmental stuttering

Jeffrey R. Mock [a,*], Anne L. Foundas [c], Edward J. Golob [a,b]

[a] Department of Psychology, Tulane University, United States
[b] Program in Neuroscience, Tulane University, United States
[c] Department of Neurology at University of Missouri, Kansas City School of Medicine, United States

ARTICLE INFO

Article history:
Received 4 August 2014
Accepted 16 May 2015

Keywords:
Speech preparation
Stuttering
Motor efference copy
Auditory event-related potentials

ABSTRACT

Motor efference copy conveys movement information to sensory areas before and during vocalization. We hypothesized speech preparation would modulate auditory processing, via motor efference copy, differently in men who stutter (MWS) vs. fluent adults. Participants ($n = 12$/group) had EEG recorded during a cue-target paradigm with two conditions: speech which allowed for speech preparation, while a control condition did not. Acoustic stimuli probed auditory responsiveness between the cue and target. MWS had longer vocal reaction times ($p < 0.01$) when the cue-target differed (10% of trials), suggesting a difficulty of rapidly updating their speech plans. Acoustic probes elicited a negative slow wave indexing motor efference copy that was smaller in MWS vs. fluent adults ($p < 0.03$). Current density responses in MWS showed smaller left prefrontal responses and auditory responses that were delayed and correlated to stuttering rate. Taken together, the results provide insight into the cortical mechanisms underlying atypical speech planning and dysfluencies in MWS.

1. Introduction

Ordinary speech is the product of extraordinary cognitive processing. Speech preparation translates an abstract message into lexical and phonological representations, which in turn are transformed into motoric codes that guide the movement of vocal articulators (Indefrey & Levelt, 2004). A coalition of neural networks support speech planning and execution, with prominent interplay between perisylvian regions associated with speech motor control and auditory processing (Hickok, 2012; Indefrey, 2011).

Given the rapidity and complexity of speech behavior it is not surprising that there are many opportunities for speech planning and production systems to become unstable, leading to moments of stuttering. Stuttering is characterized by disruptions in speech output, typically from repetitions, prolongations, and blocking (Bloodstein & Bernstein Ranter, 2008). The average age of stuttering onset is between ages 2 and 4 years, with the vast majority of children naturally recovering by age six (Bloodstein & Bernstein Ranter, 2008; Yairi & Ambrose, 2013). However, when stuttering continues through the teenage years and into adulthood the person is diagnosed with persistent developmental stuttering (Bloodstein & Bernstein Ranter, 2008).

Computational modeling suggests that precise temporal coordination between feedforward and feedback systems mediated by perisylvian speech regions are vital for speech fluency (Civier, Tasko, & Guenther, 2010; Guenther, 2006), and may contribute to developmental stuttering (Brown, Ingham, Ingham, Laird, & Fox, 2005; Max, Guenther, Ghosh, & Wallace, 2004). The feedforward system utilizes a motor efference copy to provide information about intended vocalizations to sensory areas, and provide online motor control (Crapse & Sommer, 2008; Sperry, 1950; Todorov, 2004; Von Holst, 1954). The feedback system monitors auditory and somatosensory vocal output and identifies discrepancies between the intended and actual vocal output that need correction (Guenther, 2006; Perkell, 2012). Structural and functional MRI studies in people who stutter report differences in frontal premotor/motor and posterior temporal perisylvian speech areas that may, respectively, implement feedforward and feedback operations (Beal, Gracco, & Brettschneider, 2013; Beal, Gracco, Lafaille, & De Nil, 2007; Chang, Erickson, Ambrose, Hasegawa-Johnson, & Ludlow, 2008; Chang, Horwitz, Ostuni, Reynolds, & Ludlow, 2011; Foundas et al., 2004; Fox et al., 1996; Kikuchi, Ogata, Umesaki, & Yoshiura, 2011; Mock et al., 2012; Sommer, Koch, Paulus, Weiller, & Büchel, 2002). Also, a recent resting-state fMRI study suggests even in the absence of speech production demands, adults who stutter show anomalous resting-state networks within and between anterior and posterior perisylvian regions (Xuan et al., 2012).

---

* Corresponding author.
E-mail address: jmock@tulane.edu (J.R. Mock).

http://dx.doi.org/10.1016/j.bandl.2015.05.009

Neurophysiological correlates of the feedforward system have been studied in humans using electrical (event-related potentials, ERPs) and magnetic field measures. It is well-established that ERPs or magnetic fields elicited by speech sounds are attenuated at ~100 ms latency (the N100/M100 component) when speaking relative to passive listening of recorded speech sounds (Curio, Neuloh, Numminen, Jousmäki, & Hari, 2000; Ford & Mathalon, 2004; Houde, Nagarajan, Sekihara, & Merzenich, 2002). Similar motor induced suppression is found in sensory responses to auditory and somatosensory stimuli during manual movements (Cohen & Starr, 1987; Schafer & Marcus, 1973), which suggests that N100 suppression to self-produced stimuli is not specific to vocal behaviors. The feedforward system has been hypothesized to be impaired in those who stutter (Brown et al., 2005; Civier, Bullock, Max, & Guenther, 2013; Civier et al., 2010; Max et al., 2004). However, during speech production in both children and adults who stutter motor induced suppression of the N100/M100 when speaking vs. passive listening is comparable to fluent controls (Beal et al., 2010, 2011; Liotti et al., 2010).

The above studies examined the effects of the feedforward system on auditory cortical activity in people who stutter during speech production. This project will examine feedforward influences during speech preparation in men who stutter (MWS), which could be particularly important given that over 90% of stuttering events occur on the initial sound/syllable of an utterance (Bloodstein & Bernstein Ranter, 2008; Sheehan, 1974). Our previous EEG study found that auditory processing is also subject to feedforward modulation before speaking, during the stage of speech preparation (Mock, Foundas, & Golob, 2011). In Mock et al. (2011) we used a delayed picture naming paradigm to (1) separate speech preparation from speech execution and (2) quantify auditory cortical responses by presenting auditory probes at multiple time points during speech preparation. The purpose of the acoustic probes was to assess auditory processing under conditions that encouraged vs. discouraged feedforward influences. This was done by comparing a speech condition that induced speech motor planning to a control condition where participants could not prepare a specific vocal response. The main behavioral result was that in the speech condition vocal reaction times to pictures were much faster when participants could prepare the appropriate vocal response to name the target. The auditory probes elicited ERPs that showed a long lasting negative shift (termed negative slow wave) that was larger in the speech vs. control condition and became greater as one got closer to speaking. The negative slow wave was suggested to index feedforward influences during speech preparation on auditory cortical processing.

We hypothesized that if feedforward processing is weaker in people who stutter (Brown et al., 2005; Civier et al., 2013; Max et al., 2004), then the auditory ERP negative slow wave would be smaller in MWS compared to fluent adults. Our paradigm can also examine the ability to update a speech motor program by comparing vocal reaction times to expected vs. unexpected targets (for details see Methods section below). Due to the MWS having anomalous networks within and between perisylvian regions (Chang et al., 2011; Xuan et al., 2012), we hypothesized even when speaking fluently, MWS will be slower than fluent adults at quickly updating to a different speech motor program.

2. Methods

2.1. Participants

Twelve MWS (age 34.8 ± 11.4 years, range 23–54 years, education 16.6 ± 2.9 years) were matched to 12 fluent men by age (±2 years, avg. 33.7 ± 12.1, range 23–55) and education (±3 years, avg. 16.7 ± 2.7 years, range 12–20 years). All participants were right-handed, spoke English as their first and primary language, and reported normal neurological and psychiatric health. Hearing threshold tests (0.5–8.0 kHz) were in the normal range for all participants. For each MWS the number of stuttering-like dysfluencies (monosyllabic word repetitions, part-word repetitions, blocks and prolongations) within the first 300 syllables of a conversation and reading passage was quantified from a videotaped session. All stuttering-like dysfluencies were summed and divided by the total number of spoken syllables to obtain a percent of syllables stuttered. The percent of syllables stuttered for the conversation and reading passages were averaged to obtain a stuttering rate (mean = 9.8 ± 1.3%, range = 3.5–19.5%). Stuttering rate was independently coded by the first author and a certified speech-language pathologist, and was reliable (intraclass correlation coefficient = 0.89). Each participant signed a consent form, and all experimental procedures were performed in accordance with a protocol approved by the Tulane University Institutional Review Board that was consistent with the Declaration of Helsinki.

2.2. Experimental design and task

A schematic of the design and trials/condition are shown in Fig. 1A. Participants sat in front of a computer monitor (distance = 100 cm) and viewed cue (letters 3 cm height) and target (picture 9 × 7 cm) stimuli in the center of the screen (500 ms duration). In the speech condition participants were instructed to use the cue word, which was one or two syllables, to get ready to verbally name the picture. Participants were asked to be fast but accurate, and to only use one word responses. The visual cue, in terms of meaning, matched the target picture on 90% of trials (termed "same trials"). For the remaining 10% of the trials the cue did not match the target, termed "different trials". If the cue is used to prepare a speech motor program then vocal reaction times to the target should be faster on same vs. different trials. As expected, Mock et al. (2011) showed that reaction times were much faster on the same trials. This result showed that speech preparation occurred in this task, which is a prerequisite for a speech-related feedforward influence on auditory processing.

In the control condition all cues were "XXXX" so the participants could not prepare a specific vocal response but did have information on the timing of the upcoming target. On 90% of the control trials the participants did not name the target picture but instead passively viewed the stimuli. To control for attentiveness, on 10% of the control trials a visual prompt was given 1000 ms after the target that instructed the participant to name the previous target (catch trials). Thus within the control condition, speech preparation to name a specific target would not occur between the cue and target when auditory responsiveness was being probed.

The impact of speech motor planning on auditory cortical activity was examined by measuring auditory ERPs elicited by sounds that were delivered between the cue and target. The interval between cue and target was 1.5 s, with an inter-trial interval (target to next cue) of 2.5 s. In each trial one auditory probe (~60 dB nHL, ~200 ms duration) was presented between the cue and target, at either an early (600 ms) or late (1200 ms) probe onset relative to cue onset (50% probability for early/late probe onset). The auditory probes were either a consonant–vowel or pure tone (1000 Hz, 5 ms rise/fall time). In the speech condition the consonant–vowel auditory probes were subdivided into two classes (match, mismatch). The only difference between the match and mismatch auditory probes was their relationship to the cue on a given trial. The sound of the match probes corresponded to the first consonant–vowel of the cue word, while the mismatch probes did not. Each sound type (match, mismatch, tone) was randomly presented in 20 trials/block (33% probability). A total of 60 different