SPEECH MOTOR ADAPTATION DURING PERTURBED AUDITORY FEEDBACK IS ENHANCED BY NONINVASIVE BRAIN STIMULATION

Terri L. Scott, Laura Haenchen, Ayoub Daliri, Julia Chartove, Frank H. Guenther, & Tyler K. Perrachione

Department of Speech, Language, and Hearing Sciences, Boston University, USA
tlscott@bu.edu

ABSTRACT

To maintain fluent speech production across time, learned speech motor commands must be continuously updated using auditory feedback to help speakers match their articulations to intended speech acoustics. However, the cortical mechanisms responsible for *sensorimotor adaptation* during speech remain unknown. Here, we investigated the role of speech motor cortex in modifying stored speech motor plans. In a within-subjects design, participants underwent sham and anodal transcranial direct current stimulation (tDCS) over speech motor cortex while speaking and receiving altered auditory feedback of the first formant. Anodal tDCS increased both the rate and magnitude of sensorimotor compensation for feedback perturbation. Computational modeling of our results using the Directions Into Velocities of Articulators (DIVA) framework of speech production revealed that tDCS affected speech motor cortex by increasing learning rate and decreasing sensitivity to somatosensory feedback. This study demonstrates a causal role for speech motor cortex in integrating auditory feedback into speech motor plans.

Keywords: speech production, tDCS, sensorimotor adaptation, auditory feedback perturbation

1. INTRODUCTION

Rapid and fluent speech relies on learned motor commands as well as the ability to adapt to changing conditions. Sensorimotor adaptation—learned adjustments to motor commands due to sensory feedback—serves an important role in preserving the intelligibility of speech. The acoustics of contrastive speech categories depend on minuscule differences in articulatory positions that must be updated continuously during vocal tract ontogeny [12]. Auditory feedback provides the information necessary for modifying feedforward speech motor commands to counteract these physical changes.

Sensorimotor adaptation has been demonstrated experimentally for several different auditory characteristics of speech using artificially altered feedback [16, 17, 22, 24, 27]. Speakers often produce a compensatory response by independently adjusting their production of the perturbed feature [13, 27] to oppose the acoustic perturbation, however, the cortical mechanisms supporting the integration of auditory feedback with motor planning are unknown. Speech motor control models, such as the Directions Into Velocities of Articulators (DIVA) model [9, 10, 11, 14], posit that stored motor programs for common phoneme sequences are supported by the left ventral premotor cortex (vPMC) and are integrated with compensatory responses in left ventral motor cortex (vMC) [25]. Correlational support for this model comes from neuroimaging studies in which neural activation in these regions is found during speech production [1, 8, 26], and is proportional to speakers' compensation for unexpected, intermittent, auditory feedback perturbations [2, 21].

The first aim of our study was to determine the causal role of left vPMC/vMC in sensorimotor adaptation to auditory perturbation of speech. Participants underwent an established speech production task with perturbed auditory feedback while we measured the magnitude and rate of sensorimotor adaptation reflected by changing speech acoustics. To modulate neural function of left vPMC/vMC during the task, participants simultaneously received transcranial direct current stimulation (tDCS)—a noninvasive neurostimulation technique in which a low current is applied over the scalp via electrodes to induce small changes to the electric field in underlying cortex. The direction of current flow in tDCS is believed to determine the effect of stimulation on cortical function, with anodal stimulation increasing neural excitability and cathodal stimulation decreasing excitability [6, 7, 19]. Additionally, tDCS is believed to modulate cortical plasticity, as its neuromodulatory effects can be measured for some time after stimulation has ceased [19, 20, 23].

The second aim of this study was to ascertain how tDCS affected cortical function for speech motor adaptation using computational simulations of cortical functioning during speech production. Within

the DIVA model, we identified several candidate neurocomputational variables that could hypothetically be altered by anodal tDCS in our experiment. First, by increasing cortical excitation under tDCS, the auditory feedback control system could elicit greater automatic feedback-based corrections, termed *reflexive responses* [3, 15]. Alternatively, tDCS could primarily act to increase plasticity within premotor cortices, thereby increasing *adaptive responses* to perturbed auditory feedback. Because the mechanisms through which tDCS modulates cortical function and thereby behavior are largely unknown, computational models help clarify the potential mechanism by which such neurostimulation affects cortical function.

2. METHODS

2.1. Participants

Right-handed, native speakers of American English free from speech, language, or hearing deficits completed this study (N = 18; 4 male, 14 female; age 18-28 years, $M = 20.4 \pm 2.1$). Participants provided written informed consent, approved and overseen by the Institutional Review Board at Boston University, and were paid for their participation.

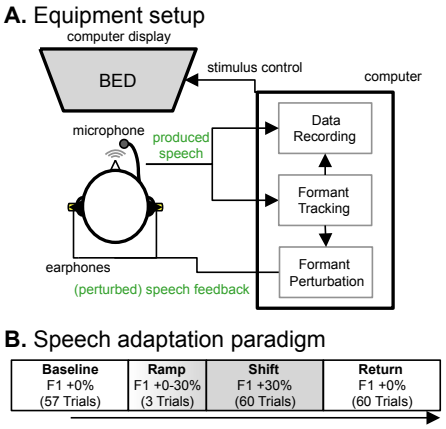
2.2. Experimental Design and Procedures

Participants completed three separate sessions of the experiment. Each session was separated from the previous by at least 7 days to reduce the potential for carry-over of learning across sessions. Because not all speakers adapt to auditory perturbation [22], participants completed an initial screening session without tDCS to confirm the presence of sensorimotor adaptation. Participants were assigned to receive either anodal or sham stimulation during their second session and the other during their third session.

2.2.1. Behavioral paradigm

Each session was conducted in a sound-attenuated chamber. Stimulus delivery, recording, and real-time resynthesis for auditory perturbation were controlled via the Audapter software [4] implemented in MATLAB vR2014b (The Mathworks, Natick, MA). Participants were prompted by the Audapter software to say aloud the words "bed," "dead," and "head," in a pseudorandom order. The paradigm began with a brief training phase, in which participants received feedback to insure they were producing the words at a sufficient loudness (72-88 dB SPL) and duration (400-600 ms). Participants continued to receive feedback about the intensity and duration of their speech during the experiment. The Audapter software performed real-time analysis, replay, resynthesis, and recording of participants' speech acoustics (F1 and F2) (Fig.1A).

Figure 1: Schematic of the (A) equipment setup and (B) behavioral paradigm.



The behavioral paradigm consisted of four phases (Fig.1B). The *baseline* phase consisted of 57 trials in which participants spoke the target words and heard their own, unperturbed speech as auditory feedback. Next, during the *ramp* phase, real-time perturbation of participants' F1 was increased linearly from 10% to 30% across three trials. During the *shift* phase, which lasted for 60 trials, participants' heard real-time auditory feedback perturbation of their own speech in which F1 was increased by 30%. Finally, during the 60 trials of the *return* phase, participants again heard their own unperturbed speech.

2.2.2. tDCS stimulation

Neurostimulation was controlled and delivered using a Soterix MxN high-definition (HD) tDCS system. Stimulating electrodes (2 mA) were placed at FC5 and C5, and return electrodes were placed at AF7, FC1, C1, and P5, in a roughly center-surround configuration [5, 18]. This montage was selected to optimize field intensity and current flow over left vPMC and vMC, though it is possible that adjacent regions may have been stimulated due to variability in subject anatomy and current spreading.

After insuring the resistance of each channel was < 10 kΩ, anodal stimulation began with a 30-s linear ramp from 0 to 2 mA, with tonic 2 mA stimulation continuing for the remainder of the session (~20 min). The procedure for sham stimulation was the same, but after the 30-s ramp to 2 mA, stimulation was linearly decreased over 30 s back to 0

mA, where it remained throughout the behavioral paradigm. This procedure effectively blinded participants to whether they were receiving anodal or sham stimulation during the behavioral task, which was begun 60 s after the onset of stimulation.

2.3. Statistical analysis

Vowel formant frequencies were isolated in recordings by analyzing 60% of the word's duration beginning 10% after the onset of voicing. Outlier trials, in which F1 deviated by more than two standard deviations from the mean value in the respective session and phase, were excluded from the analysis (< 5% of total trials). The number of F1 outliers did not differ as function of condition. Participants' F1 and F2 measurements were then normalized proportionally to the mean F1 and F2 values obtained during the baseline phase of each session.

Frequency measurements from sets of three trials (corresponding to one trial per stimulus word) were averaged to form "blocks" for statistical analysis. We assessed whether participants' motor compensation to auditory F1 perturbation during speech production was affected by tDCS in a series of linear mixed-effects analyses. Corresponding analyses were run on F2 acoustics as a control, as auditory feedback of F2 was not perturbed. Models were comprised of fixed factors for *condition* (anodal vs. sham tDCS), *time* (block number), and *session*, as well as random intercepts for participants. Significance of main effects and interactions was determined by adopting significance criterion of $\alpha = 0.05$, with p-values based on the Satterthwaite approximation of the degrees of freedom.

2.4. DIVA model simulations

Using a simplified version of the DIVA model [12, 14], we performed computer simulations to investigate which aspects of motor learning were responsible for changes in compensatory responses under neurostimulation. Eq.1 defines F1 production on a given trial (n) as:

$$(1) \quad F1_{produced}(n) = F1_{FF}(n) + \Delta F1_{FB}(n)$$

where $F1_{FF}(n)$ is the feedforward, or learned component of the produced sound, and $\Delta F1_{FB}(n)$ is the feedback-based correction. This feedback-based correction is composed of reflexive responses to unexpected auditory or somatosensory feedback, the sizes of which are given by gain factors α_A and α_S respectively:

$$(2) \quad \begin{aligned} \Delta F1_{FB}(n) = {} & \alpha_A \times (F1_{AT} - F1_{perceived}(n)) \\ & + \alpha_S \times (F1_{ST} - F1_{FF}(n)) \end{aligned}$$

Here, $F1_{AT}$ and $F1_{ST}$ are the F1 values specified by previously learned auditory and somatosensory targets, respectively, and $F1_{perceived}$ is the value of F1 heard by the subject before feedback control mechanisms become active. In the simulations, $F1_{AT}$ and $F1_{ST}$ are set to the average F1 of the baseline phase, corresponding to the assumption that the auditory and somatosensory targets will not change substantially over the course of the experiment. Eq.2 shows that the feedback-based correction is proportional to the perceived deviations from learned targets. Even though it is only auditory feedback that is experimentally manipulated, the feedforward command ($F1_{FF}$) is modified as compensatory gestures are learned, thus introducing a deviation from the original somatosensory target. Eq.3 describes the trial-by-trial updating of the feedforward command:

$$(3) \quad F1_{FF}(n+1) = F1_{FF}(n) + \lambda_{FF} \times \Delta F1_{FB}(n)$$

where λ_{FF} is a learning rate parameter for the feedforward command. In words, the feedforward command for the next trial is updated by adding a fraction of the feedback-based corrective command.

To fit the model to the data from the sham and anodal tDCS conditions, a particle swarm optimization procedure was used to find optimized values of the three free parameters of the model (α_A, α_S, and λ_{FF}) to fit the mean data for each block in each condition. The parameter estimates resulting from this procedure were highly robust to initial conditions, indicative of reaching the global minimum of the root mean square error (RMSE) measure.

3. RESULTS

3.1. Sensorimotor compensation under tDCS

In the first analysis, we tested the effect of *condition* (anodal vs. sham) to determine whether overall compensation magnitude differed under tDCS. We observed an effect of *condition*, such that compensation was significantly greater during anodal tDCS than during sham across the entire perturbation phase (Fig.2; $\beta = 0.0122$, $t = -4.23$, $p = 2.7 \times 10^{-5}$). Speakers' F1 during the latter half of the perturbation phase under anodal stimulation was 91.4% $\pm$ 4.5% that of the baseline, whereas under sham stimulation it was 93.1$\pm$ 5.0% of baseline. The corresponding analysis of F2 showed no effect of stimulation on this unperturbed feature ($\beta = -0.0017$, $t = 0.76$, $p = 0.45$; anodal: 100.0% $\pm$ 2.2% of baseline; sham: 99.6% $\pm$ 2.0%).

In the second analysis, we tested the interaction between *time* (block number) and *condition* (anodal

Figure 2: Speech adaptation under perturbation during anodal tDCS vs. sham. Shaded regions indicate standard errors.

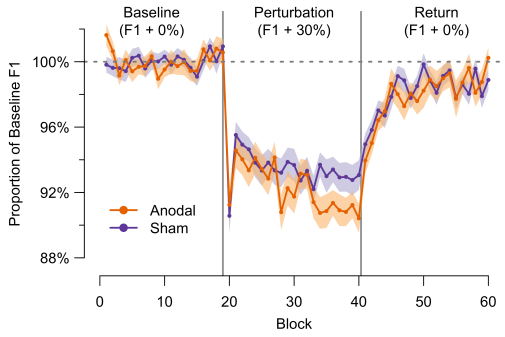
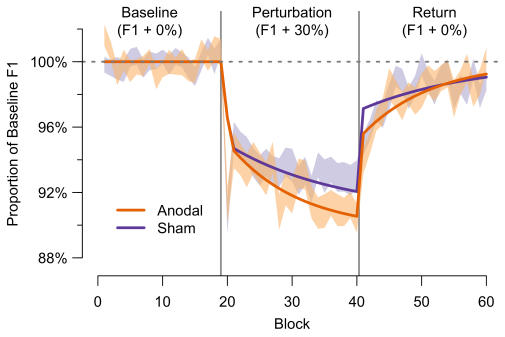


Figure 3: Model fits for anodal tDCS and sham. Shaded regions indicate standard errors around the mean behavioral data.



vs. sham) to determine whether there was a difference in the rate at which F1 compensatory responses increased over the perturbation period as a function of brain stimulation. A significant *time* × *condition* interaction revealed that F1 compensatory responses increased more rapidly under anodal stimulation than sham (Fig.2; $\beta = 0.0012, t = -2.49, p = 0.01$).

We repeated our analyses to test the effect of *session* (2 or 3) on the overall compensation magnitude during the auditory feedback perturbation. In an analysis including only *session*, there was a significant effect ($\beta = 0.012, t = 4.09, p = 4.8 \times 10^{-5}$), showing less overall compensation in the third session, however an analysis including both *condition* and *session* did not reveal any interaction, indicating that repeatedly undergoing the behavioral paradigm had an impact on compensation independent of the stimulation manipulation.

3.2. DIVA model fits

The DIVA model fits to the two experimental conditions are provided in Fig.3. In both cases, the model fit falls within the standard error of the sample mean (sham: RMSE = 7.41 Hz, Pearson's $r = 0.94$; anodal tDCS: RMSE = 7.01 Hz; $r = 0.96$) for all blocks except the ramp block and the block following cessation of feedback perturbation.

Whereas the value for the auditory feedback control gain is nearly the same for the two conditions ($\alpha_A = 0.17$ in both), the somatosensory feedback control gain decreased by 15% from $\alpha_S = 0.39$ during sham to 0.33 during anodal tDCS. The value of the trial-to-trial feedforward command learning rate increased by 13% under anodal tDCS ($\lambda_{FF} = 0.71$) compared to sham ($\lambda_{FF} = 0.63$).

4. DISCUSSION

In the current study, we have demonstrated a causal role for premotor/motor cortex in the integration of sensory feedback and feedforward motor plans during speech production. When applying non-invasive neurostimulation to the left vPMC/vMC, we observed both an increased magnitude and rate of sensorimotor compensation in response to perturbed auditory feedback. Moreover, we found that these effects were specific to F1—the perturbed frequency—and did not generalize to F2, which further indicates a learning-specific effect, and not a global disruption to speech-motor control.

Because the mechanisms by which tDCS affects cortical activity are uncertain, we modeled compensatory responses under anodal tDCS and sham stimulation using the DIVA model. We found that the best-fit model indicated that the rate of feedforward learning was increased under anodal tDCS and the somatosensory feedback gain was decreased. Through computational modeling, we were able to eliminate other possible cortical mechanisms that may have been affected by tDCS, such as increased sensitivity to auditory errors. This is not to say that auditory acuity does not play a role in compensation, as has been shown in previous studies [27], but that the areas we stimulated (left vPMC/vMC) do not mediate auditory acuity and error detection, and instead serve as a hub for the integration of feedforward and feedback speech motor commands. That the feedforward learning rate was modulated during anodal stimulation suggests that these areas must support, at some level, motor representations of speech sounds that can be modified and stored for later use.

5. REFERENCES

[1] Basilakos, A., Smith, K. G., Fillmore, P., Fridriksson, J., Fedorenko, E. 2018. Functional characterization of the human speech articulation network. *Cereb Cortex* 28(5), 1816–1830.

[2] Behroozmand, R., Shebek, R., Hansen, D. R., Oya, H., Robin, D. A., Howard, M. A., Greenlee, J. D. 2015. Sensory-motor networks involved in speech production and motor control: an fMRI study. *Neuroimage* 109, 418–28.

[3] Burnett, T. A., Freedland, M. B., Larson, C. R., Hain, T. C. 1998. Voice F0 responses to manipulations in pitch feedback. *J Acoust Soc Am* 103(6), 3153–3161.

[4] Cai, S., Boucek, M., Ghosh, S. S., Guenther, F. H., Perkell, J. S. 2008. A system for online dynamic perturbation of formant trajectories & results from perturbations of the Mandarin triphthong /iau. *In Proceedings of the 8th ISSP* 65–68.

[5] Datta, A., Bansal, V., Diaz, J., Patel, J., Reato, D., Bikson, M. 2009. Gyri-precise head model of transcranial direct current stimulation: improved spatial focality using a ring electrode versus conventional rectangular pad. *Brain Stimul* 2(4), 201–7, 207.e1.

[6] Dayan, E., Censor, N., Buch, E. R., Sandrini, M., Cohen, L. G. 2013. Noninvasive brain stimulation: from physiology to network dynamics and back. *Nat Neurosci* 16(7), 838–44.

[7] Filmer, H. L., Dux, P. E., Mattingley, J. B. 2014. Applications of transcranial direct current stimulation for understanding brain function. *Trends Neurosci* 37(12), 742–53.

[8] Ghosh, S. S., Tourville, J. A., Guenther, F. H. 2008. A neuroimaging study of premotor lateralization and cerebellar involvement in the production of phonemes and syllables. *J Speech Lang Hear Res* 51(5), 1183–202.

[9] Golfinopoulos, E., Tourville, J. A., Guenther, F. H. 2010. The integration of large-scale neural network modeling and functional brain imaging in speech motor control. *Neuroimage* 52(3), 862–74.

[10] Guenther, F. H. 1994. A neural network model of speech acquisition and motor equivalent speech production. *Biol Cybern* 72(1), 43–53.

[11] Guenther, F. H. 1995. Speech sound acquisition, coarticulation, and rate effects in a neural network model of speech production. *Psychol Rev* 102(3), 594–621.

[12] Guenther, F. H. 2016. *Neural Control of Speech*. Cambridge, MA: The MIT Press.

[13] Guenther, F. H., Espy-Wilson, C. Y., Boyce, S. E., Matthies, M. L., Zandipour, M., Perkell, J. S. 1999. Articulatory tradeoffs reduce acoustic variability during American English /r/ production. *J Acoust Soc Am* 105(5), 2854–2865.

[14] Guenther, F. H., Ghosh, S. S., Tourville, J. A. 2006. Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain Lang* 96(3), 280–301.

[15] Hain, T. C., Burnett, T. A., Kiran, S., Larson, C. R., Singh, S., Kenney, M. K. 2000. Instructing subjects to make a voluntary response reveals the presence of two components to the audio-vocal reflex. *Experimental Brain Research* 130(2), 133–141.

[16] Houde, J. F., Jordan, M. I. 1998. Sensorimotor adaptation in speech production. *Science* 279(5354), 1213–1216.

[17] Jones, J. A., Munhall, K. G. 2000. Perceptual calibration of F0 production: evidence from feedback perturbation. *J Acoust Soc Am* 108(3 Pt 1), 1246–51.

[18] Kuo, H. I., Bikson, M., Datta, A., Minhas, P., Paulus, W., Kuo, M. F., Nitsche, M. A. 2013. Comparing cortical plasticity induced by conventional and high-definition 4×1 ring tDCS: a neurophysiological study. *Brain Stimul* 6(4), 644–8.

[19] Nitsche, M. A., Paulus, W. 2000. Excitability changes induced in the human motor cortex by weak transcranial direct current stimulation. *J Physiol* 527 Pt 3, 633–9.

[20] Nitsche, M. A., Roth, A., Kuo, M. F., Fischer, A. K., Liebetanz, D., Lang, N., Tergau, F., Paulus, W. 2007. Timing-dependent modulation of associative plasticity by general network excitability in the human motor cortex. *J Neurosci* 27(14), 3807–12.

[21] Niziolek, C. A., Guenther, F. H. 2013. Vowel category boundaries enhance cortical and behavioral responses to speech feedback alterations. *J Neurosci* 33(29), 12090–8.

[22] Purcell, D. W., Munhall, K. G. 2006. Adaptive control of vowel formant frequency: evidence from real-time formant manipulation. *J Acoust Soc Am* 120(2), 966–77.

[23] Rroji, O., van Kuyck, K., Nuttin, B., Wenderoth, N. 2015. Anodal tDCS over the primary motor cortex facilitates long-term memory formation reflecting use-dependent plasticity. *PLoS One* 10(5), e0127270.

[24] Shiller, D. M., Sato, M., Gracco, V. L., Baum, S. R. 2009. Perceptual recalibration of speech sounds following speech motor learning. *J Acoust Soc Am* 125(2), 1103–13.

[25] Tourville, J. A., Guenther, F. H. 2011. The DIVA model: A neural theory of speech acquisition and production. *Lang Cogn Process* 26(7), 952–981.

[26] Tourville, J. A., Reilly, K. J., Guenther, F. H. 2008. Neural mechanisms underlying auditory feedback control of speech. *Neuroimage* 39(3), 1429–1443.

[27] Villacorta, V. M., Perkell, J. S., Guenther, F. H. 2007. Sensorimotor adaptation to feedback perturbations of vowel acoustics and its relation to perception. *J Acoust Soc Am* 122(4), 2306–2319.