Published in final edited form as:

Cortex. 2019 February ; 111: 274–285. doi:10.1016/j.cortex.2018.11.014.

Using tDCS to facilitate motor learning in speech production: The role of timing

Adam Buchwald¹, Holly Calhoun¹, Stacey Rimikis¹, Mara Steinberg Lowe¹, Rebecca Wellner¹, and Dylan Edwards²

¹New York University, Department of Communicative Sciences and Disorders, 665 Broadway Suite 900, New York, NY 10012, USA

²Burke Medical Research Institute, Edith Cowan University and Weill Cornell Medicine, 785 Mamaroneck Ave, White Plains, NY 10605, USA

Abstract

There exists debate regarding the extent to which transcranial direct current stimulation (tDCS) can affect or enhance human behavior. Here, we examined a previously unexplored domain: speech motor learning. We investigated whether speech motor learning in unimpaired participants can be enhanced using a single-session tDCS experiment, and investigated whether the timing of tDCS relative to a behavioral task affected performance. Participants (N=80) performed a twenty minute learning task with nonwords containing non-native consonant clusters (e.g., *GDEEVOO*), and were assigned to groups receiving either sham or active tDCS either immediately before or during the task. Both accuracy and properties of errors were examined throughout the course of the practice task, and then practice was compared to a retention period 30 minutes later (R1) and two days later (R2). For cluster and whole-(non)word accuracy measures, acquisition was observed for all groups during the practice session. Compared to the beginning of practice, the tDCS-Before group showed significantly greater improvement than both the sham group and the tDCS-During group at R1. An effect was also observed for vowel duration in errors (/gdivu/ → [gədivu]), with the tDCS-Before group showing significant shortening of vowel errors throughout practice. Overall, the findings suggest that tDCS can improve speech motor learning, and that the improvement may be greater when tDCS is applied immediately before practice, warranting further exploration of this new domain for tDCS research.

Keywords

tDCS; speech; motor learning; motor control

Corresponding author: Adam Buchwald, New York University, Department of Communicative Sciences and Disorders, 665 Broadway, Suite 900, New York, NY 10012, Phone: (212) 998-5260, Fax: (212) 995-4356, buchwald@nyu.edu.

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

Conflict of interest: The authors declare no competing financial interests

1 Introduction

There has been great interest and debate over the past decade about using transcranial direct current stimulation (tDCS) to affect or enhance behavioral outcomes in various domains, including normal and impaired motor control (Kang, Summers, & Cauraugh, 2015), cognition (Flöel, 2014; Katz et al., 2017), language (Baker, Rorden, & Fridriksson, 2010; Meinzer, Lindenberg, et al., 2014; Wortman-Jutt & Edwards, 2017), and speech fluency (Chesters, Möttönen, & Watkins, 2018). While some tDCS studies have reported positive or promising outcomes, particularly in research involving learning (e.g., Meinzer, Jähnigen, et al., 2014) and motor control (e.g., Orban de Xivry & Shadmehr, 2014), the accumulated evidence remains equivocal, and the neurocognitive mechanisms underlying changes from neuromodulation remain incompletely understood (Bestmann & Walsh, 2017; Wiethoff, Hamada, & Rothwell, 2014). This has led researchers to raise concerns about the limitations of tDCS and its effectiveness in various domains (Bestmann & Walsh, 2017; Horvath, Forte, & Carter, 2015a, 2015b; Medina & Cason, 2017; Nilsson, Lebedev, Rydström, & Lövdén, 2017; Westwood, Olson, Miall, Nappo, & Romani, 2017). Despite these concerns, tDCS investigations persist due to the intriguing possibility of enhancing performance and/or rehabilitation using relatively inexpensive and well-tolerated instrumentation (Bikson et al., 2016). This persistence may be necessary to determine whether there are optimal approaches to using tDCS for a given domain, due to the large number of parameters in tDCS studies (e.g., montage, current level, dosage, timing). We report below on a study examining tDCS timing relative to a task in a previously unexamined domain: speech motor learning.

Speech motor learning (operationalized here as learning to produce novel speech sound sequences) involves both motor control and learning, and is an increasingly widely studied domain because of the potential application to individuals with acquired speech deficits such as apraxia of speech (see Maas et al., 2008 for a review). A great deal of attention has been given to understanding the extent to which speech motor learning can be enhanced based on the same approaches to training (i.e., the structure of practice and feedback) that facilitate motor learning in non-speech domains (Maas et al., 2008; Schmidt & Lee, 2005). Following this idea, we investigated whether using tDCS as an adjunct to a speech motor learning task would facilitate learning as it has been reported to do in non-speech tasks (Nitsche et al., 2003; Stagg & Nitsche, 2011), examining both the acquisition period as well as two retention time-points.

To begin addressing the optimal approach, we also examined whether the order of tDCS and a speech learning task would affect learning outcomes. This question builds on previous findings that tDCS increases cortico-motor excitability, and that those increases can last for at least an hour when stimulation is followed by motor practice (Edwards et al., 2009). In this study, we compared whether tDCS administered immediately before a speech task would result in different learning outcomes than administration during the task, as there have been mixed findings regarding the order of tDCS and a learning task in limb motor learning. One study reported that tDCS immediately before a robotic-arm treatment task enhanced acquisition of arm motor learning in stroke rehabilitation (Giacobbe et al., 2013), at least for movement smoothness, whereas tDCS during the task negatively impacted a measure of aim. However, this finding contrasts with two other studies that showed better motor skill

acquisition when tDCS was administered during a motor task rather than before it in both unimpaired learners (Stagg et al., 2011) and stroke patients (Sriraman, Oishi, & Madhavan, 2014; Stagg et al., 2011). We note that of the three, only the Sriraman et al. study included a measure of learning retention. For same day retention, their findings indicated a continued benefit of tDCS during the task, whereas the following day, both active tDCS groups achieved similar performance, which exceeded that of the individuals who received sham stimulation. Stagg and Nitsche (2011) proposed that the interaction of tDCS and motor learning relies on changes in membrane potential, with potentially different changes during stimulation and in the after-effects of stimulation. As the work on timing has yielded somewhat mixed results which focus on motor skill acquisition but not retention, we opted to use this parameter as the first one to test the optimal approach to using tDCS to enhance speech motor learning.

We trained participants to produce difficult non-native consonant sequences (e.g., /gd/ in GDEEVOO; /fn/ in FNEEGDWOP) and compared their learning rates based on the tDCS condition (sham, stimulation immediately before task, stimulation during task). These sequences were composed of legal English sounds, but the sequence of consonants was *phonotactically illegal* (i.e., not appearing in English words) in the word/syllable onset position (see Segawa, Tourville, Beal, & Guenther, 2015; and Steinberg Lowe & Buchwald, 2017 for related tasks). These non-native sequences are known to be difficult to produce (Davidson, 2006a, 2006b; Wilson, Davidson, & Martin, 2014) and perceive (Berent, Steriade, Lennertz, & Vaknin, 2007; Davidson, 2007). Production errors on these sequences most commonly include schwa vowel insertion (e.g., GDEEVOO → [gədivu]), and listeners have trouble distinguishing correct tokens from those with this error. To address perceptual limitations, accuracy was assessed using acoustics as a guide for coding participant errors. In addition, for our most difficult sequence (/gd/ onsets), tokens produced with schwa insertion errors were further analyzed to determine whether the duration of the schwa shortened over the course of the study, which would indicate that participants were actively learning the target motor plan and approaching the target.

2 Materials and Methods

2.1 Participants

Eighty participants (50F, mean: 23.7 years) recruited from advertisements and flyers in the NYU community completed the study. Exclusion criteria included a history of speech or language impairment, familiarity with languages that contained any of the specific consonant clusters being trained (e.g., Russian, Polish, Greek, Hebrew) as well as phonetic training through academic coursework. All participants reported normal or corrected-to-normal vision, and informed consent was obtained according to the NYU Langone Medical Center Institutional Review Board. The power calculation we conducted prior to this work was based on other studies of speech motor learning that did not involve non-native clusters (Sadagopan & Smith, 2013; Sasisekaran, Smith, Sadagopan, & Weber-Fox, 2010). Those calculations indicated that the estimated sample size using an α -level of .01 and $1-\beta$ of .95 was $N=16$ in each group. We recruited a more conservative 20 participants in each group because we were concerned that the task that we used would be harder and lead to overall

smaller effects. An additional 25 (14F, 11M) participants were initially consented, but 13 did not return for day two, a technical problem ruled out the participation of seven participants, and five did fit our exclusion criteria but did not make that clear until after they were consented.

2.2 Speech Stimuli

The speech stimuli were disyllabic nonwords beginning with non-native native consonant clusters (e.g., GDEEVOO; FNEEGDWOP; see Appendix A). Eight non-native clusters were used, with four nonwords for each cluster trained during the practice session and four saved for retention session testing (see Procedure). All nonwords contained legal English sounds and sequences other than the initial consonant cluster, but differed in the Consonant-Vowel (CV) shape of the syllables (e.g., GDEEVOO /gdivu/ has the shape [CCV.CV]; FNEEGDWOP/fnigdwop/ has the shape [CCVC.CCVC]).

The auditory stimuli were recorded by a phonetically-trained Russian-English bilingual speaker using a Shure SM-10 head-mounted microphone attached to a Marantz PMD660 digital recorder. The soundfiles were spliced to leave 10ms of silence at the onset and offset of each item using Praat (Boersma & Weenink, 2017), and the stimulus amplitude was normalized. Orthographic versions of the nonwords were also presented to participants to ensure that errors in cluster production did not arise from misperception.

2.3 Speech motor learning task procedure

All testing took place in a sound-attenuated testing room. Participants sat in front of a computer and their productions were recorded using a Shure BETA 58A microphone in a desktop microphone stand connected to the Marantz PMD660. This setup permitted the participant to have the tDCS setup on their scalp throughout the learning task.

2.3.1 Pre-practice.—To ensure that individuals understood the task, we used two monosyllabic items in the pre-practice (/fnrt/ and /ferk/). These items were selected because English listeners' perception of these clusters is typically accurate (Davidson, 2010), allowing us to provide accurate feedback. Feedback was provided regarding whether the target was produced accurately (*knowledge of results*) and also what specific aspects of the production were wrong in cases where there was an error (*knowledge of performance*). Additionally, all participants were instructed to attend to the consonant cluster at word onset and to try not to produce a vowel before or in between the two consonants. Pre-practice lasted approximately two minutes.

2.3.2 Practice—During practice, participants produced 4 nonwords per cluster (32 nonwords with illegal clusters total) five times each along with 62 filler stimuli not containing clusters. Random stimulus presentation orders were generated and then modified to ensure that no stimulus was presented twice in succession. Items were balanced across participants such that half of the participants were trained on one half of the nonwords and the other half on the other nonwords. For each stimulus, participants were presented with both auditory and orthographic versions simultaneously. The practice session was structured to include a large amount of practice of the clusters presented randomly and in varied

contexts. No online feedback was provided given the difficulty of perceiving the accuracy of these clusters.

2.3.3 Retention.—The shorter-term retention (R1) and longer-term retention (R2) sessions were identical, with R1 beginning approximately 30 minutes after the end of practice, and R2 taking place in a separate session two days later. For each cluster, participants produced the four trained nonwords as well as four novel, untrained nonwords beginning with the cluster. Each stimulus was produced three times (192 total cluster stimuli) during the retention sessions, with an additional 110 filler items not containing clusters. The stimuli were randomized and presented electronically using the E-Prime 3.0 software (Psychology Software Tools Inc., 2016).

2.4 tDCS procedure

The tDCS electrodes were placed on all participants 20 minutes before the beginning of the pre-practice (see Figure 1 for depiction of experiment). These 20 minutes were spent filling out paperwork and performing computerized working memory tasks (e.g., forwards and backwards digit and block span, all with keyboard or button box responses). A 1×1 Soterix battery-driven current stimulator delivered current using rubber-carbon electrodes (35cm²) with surrounding saline-soaked sponges. The anode was placed over the left motor cortex (C3 in the 10/20 EEG system) and the cathode was placed over the right supraorbital area. This electrode montage was selected as we hypothesized that the underlying mechanism of learning in this task would be similar to other motor learning tasks that have used this montage. For participants receiving active tDCS, the device delivered 1mA of current for 20 minutes. For sham tDCS, the electrode montage was identical and current ramped up to 1mA over 30 seconds to simulate the sensation of stimulation, but then was immediately decreased and turned off (Ambrus et al., 2012). Stimulation intensity and duration were selected based on the most commonly used parameters in the literature.

Participants were randomly assigned to receive either sham or active stimulation, and also to receive that stimulation immediately before or during the pre-practice/practice session. Active/sham is set with a single toggle on our device and was done in advance by a separate member of the research team. Participants were blind to stimulation condition but were aware of stimulation timing. Participants were asked which tDCS condition they believed they were in on Day 2 and were at chance in accurately identifying their condition (44/80 correctly identified their condition, with 22/40 sham and 22/40 active accurately identifying the condition). The electrode montage and a simulation of the neural regions affected by the current are presented in Figure 2. This montage was selected largely because of its common use in the literature on tDCS and motor learning, with 20 minutes being the modal duration of stimulation. We will return to a discussion of these parameters in the discussion.

2.5 Data analysis

2.5.1 Cluster and nonword accuracy—All recorded productions containing consonant clusters were transcribed using the acoustic waveform, spectrogram, and perception. The most common error type in producing these clusters is the insertion of a schwa vowel ([ə]) either between the two consonants (/gdivu/ → [gədivu]) or before the

cluster (/gdivu/ → [əgdivu]). Because these insertions are difficult for English speakers to perceive, we used acoustic properties to identify the presence of a vowel. Following Wilson et al. (2014), a vowel was identified by the presence of formant structure in the spectrogram (particularly higher formants F2, F3, etc., as these are not confusable with f0) and a corresponding vocalic (periodic) portion in the acoustic waveform. Figure 3 depicts two /fn/ clusters, one produced accurately (Fig. 3A) and one with an epenthetic vowel (Fig. 3B).

Other errors such as consonant deletion (e.g., /gdivu/ → [divu]) and substitution (e.g., /gdivu/ → [glivu]) were identified perceptually and verified with the acoustic record. The remaining sounds in the nonwords (e.g., the [ivu] portion of /gdivu/) were transcribed perceptually. Cluster accuracy reflects the accuracy of the first two consonants only; whole nonword accuracy reflects the entire stimulus. For each participant, the scoring was completed by a single coder who was blinded to the tDCS condition of the participant as well as the session that each token came from. Interrater reliability was evaluated in two ways. First, 25% of the participants (N=20) were coded completely by two separate individuals (point-to-point interrater agreement: 85.0%). Second, each /gd/ token from each speaker was analyzed to determine if it contained epenthesis by two independent coders (point-to-point interrater agreement: 85.7%)

2.5.2 Epenthesis duration—As a secondary measure of cluster acquisition, we examined whether the inserted vowel in /gd/ cluster errors shortened throughout the experiment as an indication that participants were coming closer to the target articulation. The duration of these vowels was measured based on the criteria described above for determining their presence: vowel onset was marked at the onset of the periodic vocalic portion that coincided with the higher formant structure, and the offset was marked at the end of that portion. Each /gd/ token was scored by two independent coders for accuracy and for vowel duration. For each token where the measurements differed by more than 5 ms, two coders then re-examined the item and agreed upon the measurement. These were the final values used in the analyses.

2.6 Experimental design and statistical analysis

We examined accuracy changes during practice based on tDCS condition as well as changes between overall practice performance and the two retention periods. To evaluate this statistically, we used separate mixed-effects models for each dependent variable: cluster accuracy, whole nonword accuracy, and epenthesis duration. Each statistical analysis was performed in R (R Core Team, 2017). To examine acquisition (changes during practice), the nonwords containing the target clusters were divided into quintiles (32 for each participant). For these practice models analyzing cluster and nonword accuracy, we used logistic mixed effects models with stimulation group (hence Group), Quintile, and the Group × Quintile interaction as fixed effects, and Subject and Item as random effects.

We examined retention of learning in two ways. To determine what was acquired and whether it was retained, the beginning of the practice session was compared with the two retention sessions. Additionally, a comparison between the end of the practice session and the two retention sessions provided a measure of how well the learned knowledge was

retained. For each accuracy measure (word and cluster), we ran models that included Group, Session, and stimulus Type (trained vs. untrained) as fixed effects with the same random effects structure as described above. Our sessions were Q1 (first quintile of practice), Q5 (last quintile of practice), R1, and R2. We used a single mixed model in order to use subject performance as a random factor overall, with the comparisons of interest being Q1 vs. R1 and R2 and Q5 vs. R1 and R2. To simplify the analyses, we first ran a model comparing sham-before to sham-during to determine whether these two groups could be collapsed for the final analyses. In all cases, the models indicated that these groups did not differ in their performance, and they were collapsed into a single group.

To analyze the epenthesis data, we used linear mixed-effects models with inserted schwa duration as the dependent variable, and the same fixed and random effects structure as described above for both evaluating change during acquisition and change from Q1 and Q5 to the retention sessions. For all models, log-likelihood comparisons were used to determine the effects that significantly contributed to the final models selected for each dependent variable (Baayen, Davidson, & Bates, 2008).

3 Results¹

For all comparisons, the sham conditions were collapsed into a single group as the two groups did not differ significantly, yielding three participant groups: Sham, Before, and During.

3.1 Whole nonword and cluster accuracy: Acquisition

During practice, Quintile was a significant predictor of whole nonword accuracy for all groups (Table 1). In addition, there was a marginal Group $\times$ Quintile interaction for Before vs. Sham ($\hat{\beta}=0.07$, $z=1.84$, $p=0.0652$) and a significant interaction for During vs. Sham ($\hat{\beta}=0.09$, $z=2.32$, $p=0.0203$) with no difference between the two active tDCS conditions (Fig 4A). For cluster accuracy, Quintile was a significant predictor of accuracy for the Before and During groups, and marginally significant for the Sham group (see Table 1), indicating that all groups improved during the practice session. There were no Group $\times$ Quintile interactions (Fig 4B).

3.2 Whole nonword and cluster accuracy: Retention

The session models were used to determine whether the change in cluster accuracy from Practice to Retention differed across stimulation groups by examining the interaction of Group and Session. Log likelihood tests indicated that removing stimulus Type (trained vs. untrained) did not significantly alter the fit of the model for word accuracy ($\chi^2(1)=1.02$, $p=0.3129$) or cluster accuracy ($\chi^2(1)=0.01$, $p=0.9214$). For whole word accuracy (Figure 4A), the Before condition exhibited significantly greater improvement from Q1 than both During ($\hat{\beta}=0.31$, $z=2.09$, $p=0.03489$) and sham ($\hat{\beta}=0.34$, $z=2.59$, $p=0.00876$) at R1, and a significantly greater improvement than During from Q5 to R1 ($\hat{\beta}=0.50$, $z=3.35$, $p=0.0008$). Additionally, Sham exhibited greater improvement than During from Q5 to R1 ($\hat{\beta}=0.40$,

¹All data are archived in csv format: https://osf.io/dpsyzy/?view_only=a256f1d6e5fb4d97adac2147f58811ae

$z=3.08$, $p=0.0020$) and R2 ($\hat{\beta} = 0.28$, $z=2.19$, $p=0.0288$), with no other significant interactions. The cluster accuracy model revealed that the Before condition exhibited significantly greater improvement in cluster accuracy (Fig 4B) than During at R1 ($\hat{\beta} = 0.38$, $z=2.44$, $p=0.0107$), and Sham exhibited greater improvement than During at both R1 ($\hat{\beta} = 0.36$, $z=2.68$, $p=.0074$) and R2 ($\hat{\beta} = 0.27$, $z=2.00$, $p=.04516$) at R1, with no other significant interactions).

3.3 Epenthesis duration

During practice, Quintile was a significant predictor of vowel duration for the Before epenthesis errors ($N=259$, $b=1.58$, $t=3.56$, $p=0.0004$), but not for the During errors ($N=258$, $p=.4997$) or Sham errors ($N=485$, $p=.8558$). This indicates that only Before participants' error vowels decreased in duration during practice (Fig 5a). There was a significant Group $\times$ Quintile interaction for Before vs. During ($b=1.88$, $t=3.00$, $p=0.0028$) and for Before vs. Sham ($b=1.52$, $t=2.75$, $p=0.0060$), with no difference between the During and Sham conditions (Fig 5A).

To determine whether the change in vowel duration from Q1 and Q5 to the Retention sessions differed across groups, we examined the interaction of Group and Session in the session model. For Q1, participants in the Before group exhibited a significantly different change from the During group at both R1 ($b=5.16$, $t=2.34$, $p=.0194$) and R2 ($b=6.35$, $t=2.89$, $p=.0039$), with the duration of epenthetic vowels decreasing for the Before group and increasing for the During group at both retention sessions. The Sham group also exhibited a bigger change than During from Q1 to R2 ($b=5.58$, $t=2.87$, $p=0.0042$). These differences reflect an increase in the duration of error vowels for the During group, and a decrease for the Sham group. For Q5, Sham showed a significantly different change than Before at R2 ($b=3.41$, $t=1.99$, $p=0.0457$) and a marginal difference from During at R2 ($b=3.36$, $t=1.94$, $p=0.0531$). These changes reflect that the error vowels produced by participants in the Sham group were shorter at R2 than at Q5, whereas the two other groups exhibited increases (see Fig 5). No other interaction terms were significant.

4 Discussion

The results of this study suggest that the order of tDCS relative to a speech motor learning task may affect an individual's performance changes. Participants performed a difficult speech motor learning task involving non-native word-initial consonant sequences, and received either sham or active tDCS immediately before or during a speech motor learning task. All participant groups exhibited significant improvement in accuracy during acquisition, as well as from the beginning of acquisition to at least one retention session. This indicates that the short learning task involving non-native consonant clusters (that did not include experimenter feedback) led to measurable gains overall. The finding of improvement by all groups allows us to test whether different approaches to using tDCS affected the learning effects, consistent with other approaches to evaluating tDCS (Allman et al., 2016; Holland & Crinion, 2011).

4.1 tDCS timing interacts with speech motor acquisition

With respect to performance during the practice session, there were no significant differences between the two active stimulation groups; both groups showed increases in accuracy reflecting acquisition of the target clusters during practice. It is worth noting that the group receiving stimulation during the task improved more during practice than sham while there were no differences between the Before group and the Sham group, indicating that stimulation during the learning task may be particularly helpful during acquisition (as seen in Stagg et al., 2011 and Sriraman et al., 2014 discussed in the introduction). However, the timing of stimulation did not lead to differences among participants receiving active stimulation. In terms of retention, participants receiving stimulation immediately before the task showed significantly greater improvement than those receiving stimulation during the task when comparing a short-term retention session 30 minutes later to both the beginning of the practice and the end of the practice. These findings are novel, but reflect our comparison of the retention period to the beginning and end of acquisition. Thus, while some reports have indicated increased acquisition associated with stimulation during a motor learning task, previous studies did not compare the practice performance to the retention performance.

It is worth noting that the group effects were obtained only in the shorter-term retention session performed 30 minutes and not two days later. It is possible that this distinction is an indication that none of the stimulation conditions would truly affect longer-term measures of learning, although we hasten to note that the participants received a very small amount of training to learn complex speech motor plans that they had never previously produced. In considering the significance of this result, it is worth considering the short-term retention finding carefully. The shorter-term retention began 30 minutes after the stimulation ended for the During group and 50 minutes after it ended for the Before group, indicating a clear difference between the two active tDCS groups. However, each of those time frames are within the window where 20 minutes of tDCS with this montage is expected to affect cortical activity (Edwards et al., 2009; Nitsche & Paulus, 2001), and if anything, the continued effects of stimulation in the During group should be more likely to persist. Thus, it is unclear how this difference in time after stimulation ended would lead to the observed pattern with better retention for the Before group. It has also been reported that the current level shown here may have very small effects on cortical behavior, and that much more current is required particularly for instantaneous changes (Vöröslakos et al., 2018). It remains possible that the effects seen here occur because participants in the Before group received a full 20 minutes of stimulation (albeit at 1mA) prior to the task whereas participants receiving stimulation during the task did not have any buildup of possible cortical changes prior to beginning the learning task. Addressing this type of mechanistic question is outside of the scope of this work, but should be a clear goal of future work examining timing-related changes in how tDCS can affect behavior.

An additional issue that could have impacted the findings is the nature of the tasks performed before the speech learning task, as these tasks were performed during stimulation for the Before group. Our design was intended to minimize speech production during the twenty minute period before the task so participants performed computerized tasks with

non-verbal responses (e.g., block span) as well as digit span tasks with spoken responses. The goal was to include tasks that would not engage the speech motor control network too strongly, as working memory has been largely associated with pre-frontal cortex (PFC) activation (D'Esposito & Postle, 2015). However, it remains possible that these tasks recruit neural regions that were affected by the current flow depicted in Figure 1B. Previous research has found that tasks performed during stimulation may affect performance on a task immediately after timing (see Horvath, Carter, & Forte, 2014 for discussion). For example, Nozari, Woodard, and Thompson-Schill (2014) reported differential effects on post-stimulation Flanker task performance when participants performed another task with high PFC involvement compared to a task with low PFC involvement during cathodal tDCS placed over the PFC. In particular, comparing cathodal tDCS to sham, participants in the cathodal condition were faster and more accurate at the Flanker after stimulation with a low PFC involvement task during stimulation, but slower when the task had high PFC involvement. These findings suggest that the task performed during stimulation, and the neural regions it recruits, can affect performance on later tasks. Although we attempted to select tasks that were expected not to engage regions most affected by the current, we cannot rule out that this played a role on the basis of the current design. Future work would need to systematically vary the tasks performed prior to the speech motor learning task to verify that the effects are not dependent on performing working memory tasks during the stimulation.

The group differences based on the timing of stimulation indicate the importance of testing a variety of parameter settings when considering whether tDCS can affect performance in a particular domain, as has been noted by others (de Aguiar, Paolazzi, & Miceli, 2015). Given that review papers (e.g., Horvath et al., 2015a, 2015b; Medina & Cason, 2017) often indicate that there is limited data supporting the use of tDCS to promote increases in performance, it is critical to consider the possibility that certain parameter combinations are more effective than others, particularly when there is a goal of translating the research to rehabilitation. For example, it remains unknown whether the timing differences obtained here would be observed had we used different stimulation parameters. We note that in the speech and language rehabilitation literature, most studies begin stimulation and treatment at the same time with treatment continuing for up to an hour after stimulation has ended (Wortman-Jutt & Edwards, 2017). In addition, it may also be necessary to individualize the stimulation approach for individuals with acquired neurological impairment subsequent to stroke (Shah-Basak et al., 2015).

4.2 tDCS and speech motor learning

This study presents promising results for using tDCS to enhance speech motor learning, although it is worth noting that this study was based on a single session of tDCS and a single motor learning practice session. Thus, it is unknown whether the differences based on tDCS timing would be maintained in a longer-term study with multiple sessions as is more typical in motor learning studies and in the stroke neurorehabilitation literature. In addition, the differences in improvement for the different tDCS groups were limited to the same day retention test. One possibility suggested by this pattern is that stimulation prior to the task somehow facilitates or expedites memory consolidation that the other tDCS group did not exhibit until the second session two days later.

In the one study we are aware of that examined the brain regions associated with non-native consonant sequence learning, Segawa et al. (2015) trained participants on nonwords that contained non-native consonant clusters, and then tested them on trained and untrained clusters during fMRI scanning. They reported increased activation in primarily cortical regions including left premotor cortex, frontal operculum, and posterior superior temporal cortex for the production of trained non-native consonant sequences compared to untrained non-native sequences. Given the tDCS montage that we used (see Figure 1c), it is likely that the left motor cortex, premotor cortex and frontal operculum all received some depolarization from the current, which may be related to the increased performance. Based on the DIVA model (Guenther, Ghosh, & Tourville, 2006), Segawa et al. hypothesized that repeated production of the non-native clusters during training leads to the formation of a new motor program (stored in premotor cortex) for the cluster rather than the assembly of the cluster from two separate motor programs. Similarly, the motor cortex and pre-motor cortex have been jointly implicated in the mechanisms supporting non-speech motor learning as well (Yokoi, Arbuckle, & Diedrichsen, 2018).

In a recently published study, Lametti et al. (2018) examined the effect of tDCS on sensorimotor adaptation in a perturbed auditory feedback paradigm. Participants received stimulation to the motor cortex (the same montage as used here) or the cerebellum, with active stimulation and sham control, and heard their F1 shifted upward (toward a different vowel/word) during a 12.5 minute learning phase. In aggregate, the participants receiving stimulation to the cerebellum responded by countering the shift (lowering their F1 with little change in the unaltered F2 dimension) while the participants receiving motor cortex stimulation responded by shifting both F1 and F2 away from the perceived vowel error. These results demonstrated a complex cortico-cerebellar network involved in speech production and sensorimotor adaptation. Taken together with the present findings, these studies demonstrate that tDCS can have measurable impact on learning and adaptation in speech production in even a single session, although more research is clearly required to demonstrate the consistency of these findings.

Two additional patterns in the data may help us understand aspects of speech motor learning. First, in addition to changes in accuracy, there were changes in the error productions of the participants over the course of the practice. In particular, when participants in the Before group made schwa insertion errors in producing /gd/ clusters, their errors became closer to the target, which may reflect that learning to coordinate these non-native speech sequences is a gradual (rather than a discrete) process with learners continuing to improve at approximating the target. In this case, only the group receiving tDCS before the task exhibited significant decreases in schwa duration in these error tokens (i.e., the errors became closer to the target) during practice, and the differences between practice and longer-term retention may have reflected a performance decrement on this measure for the During group, a finding for which we have no clear explanation.

Second, there was no accuracy difference obtained between trained and untrained items. In other words, learning to produce a non-native cluster in some nonwords generalized to the production of that cluster in other, untrained nonwords. This suggests that the learning is occurring at the level of the novel speech sequence (the cluster) rather than at the level of the

token itself. This finding is consistent with other investigations of speech motor learning in impaired speakers with apraxia of speech where it has been reported that improvement in the production of a speech sound or a sequence of sounds trained in one set of stimuli generalizes to the production of that sound or sequence in untrained stimuli (Aichert & Ziegler, 2008; Buchwald, Gagnon, & Miozzo, 2017).

4.3 Summary

A single-session of tDCS as an adjunct to a speech motor learning task indicated that unimpaired individuals receiving active stimulation prior to the task exhibited increased learning relative to individuals receiving sham stimulation or those receiving active stimulation during a task. This research provides preliminary evidence that tDCS may enhance learning performance in the speech domain under specific stimulation conditions. This finding warrants further research to replicate the finding that is aimed at understanding the neurocognitive mechanism underlying this effect and the difference between the conditions.

Acknowledgments

The authors thank Emily Atkinson, Krystal Canellas, Erin Clancy, Megan Cummings, Christina Grisanzio, Kasey Fahy, Kelly Karpus, Christine Laganella, Joely Mass, Melyssa Weller, Melissa Wells, and Joshua Wittstein for assistance with data collection and coding.

Funding

This work was supported by an award from the National Institute on Deafness and other Communication Disorders (NIDCD) to AB (K01DC014298).

Appendix A:: International phonetic alphabet (IPA) transcription and orthography for experimental stimuli

Cluster	IPA	Orthography	IPA	Orthography
/gd/	gdivu	GDEEVOO	gdæpi	GDAPEE
	gdanæd	GDAHNAD	gduwam	GDOOWOM
	gdulsam	GDOOLSAHM	gdabput	GDOBPOOT
	gdinsten	GDEENSTAN	gdækbrit	GDAKBREET
/zg/	zginu	ZGEENOO	zgæni	ZGANEE
	zgadæn	ZGAHDAN	zgupæb	ZGOOPAB
	zgudmak	ZGOODMOCK	zgabdut	ZGOBDOOT
	zgipbram	ZGEEPBRAM	zgæmklid	ZGAMKLEED
/zb/	zbædi	ZBADEE	zbæda	ZBADAH
	zbudab	ZBOODOB	zbukib	ZBOOKKEEB
	zbalruk	ZBAHLMOOK	zbalgæn	ZBAHLGAN
	zbagskip	ZBOGSHKEEP	zbædskum	ZBADSKOOM
/pt/	ptægi	PTAGEE	ptiku	PTEEKOO
	ptuvæl	PTOOVAL	ptawæn	PTAHWAN
	ptamsuk	PTOMSOOK	ptulnap	PTOOLNOP

	ptælswin	PTALSWEEN	ptintwæg	PTEENTWAG
	fniku	FNEEKOO	fnavi	FNAVEE
/fn/	fnadæp	FNAHDAP	fnubæz	FNOOBAZ
	fnuktas	FNOOKTASS	fnabzud	FNOBZOOD
	fnigdwap	FNEEGDWOP	fnætʃrig	FNATSHREEG
	fmæbi	FMABEE	fminu	FMEENOO
/fm/	fmudæp	FMOODAP	fmapæg	FMAHPAG
	fmadzud	FMODZOOD	fmukbam	FMOOKBOM
	fmæbglip	FMABGLEEP	fmibklæt	FMEEBKLAT
	vmidu	VMEEDOO	vmæki	VMAKEE
/vm/	vmasæb	VMAHSAB	vmubæp	VMOOBAP
	vmugpab	VMOOGPOB	vmaldub	VMAHLDOOB
	vmipstem	VMEEPSTAM	vmækswig	MAKSWEEG
	vnibu	VNEEBOO	vnæzi	VNAZEE
/vn/	vnapæm	VNAHPAM	vnugæn	VNOOGAN
	vnuʃpak	VNOOSHPOK	vnafmuk	VNAHFMOOK
	vnikspæd	VNEEKSPAD	vnægdwim	VNAGDWEEM

REFERENCES

- Aichert I, & Ziegler W (2008). Learning a syllable from its parts: Cross-syllabic generalisation effects in patients with apraxia of speech. *Aphasiology*, 22(11), 1216–1229.
- Allman C, Amadi U, Winkler AM, Wilkins L, Filippini N, Kischka U, et al. (2016). Ipsilesional anodal tDCS enhances the functional benefits of rehabilitation in patients after stroke. *Science Translational Medicine*, 8(330), 330re331–330re331.
- Ambrus GG, Al-Moyed H, Chaieb L, Sarp L, Antal A, & Paulus W (2012). The fade-in – Short stimulation – Fade out approach to sham tDCS – Reliable at 1 mA for naïve and experienced subjects, but not investigators. *Brain Stimulation*, 5(4), 499–504. [PubMed: 22405745]
- Baayen RH, Davidson DJ, & Bates DM (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(4), 390–412.
- Baker JM, Rorden C, & Fridriksson J (2010). Using transcranial direct-current stimulation to treat stroke patients with aphasia. *Stroke*, 41(6), 1229–1236. [PubMed: 20395612]
- Berent I, Steriade D, Lennertz T, & Vaknin V (2007). What we know about what we have never heard: Evidence from perceptual illusions. *Cognition*, 104, 638–643.
- Bestmann S, & Walsh V (2017). Transcranial electrical stimulation. *Current Biology*, 27(23), R1258–R1262. [PubMed: 29207262]
- Bikson M, Grossman P, Thomas C, Zannou AL, Jiang J, Adnan T, et al. (2016). Safety of transcranial direct current stimulation: evidence based update 2016. *Brain stimulation*, 9(5), 641–661. [PubMed: 27372845]
- Boersma P, & Weenink D (2017). Praat: doing phonetics by computer (Version 6.0.36) [Computer program]. Retrieved from <http://www.praat.org/>.
- Buchwald A, Gagnon B, & Miozzo M (2017). Identification and Remediation of Phonological and Motor Errors in Acquired Sound Production Impairment. *Journal of Speech, Language, and Hearing Research*, 60(6S), 1726–1738.
- Chesters J, Möttönen R, & Watkins KE (2018). Transcranial direct current stimulation over left inferior frontal cortex improves speech fluency in adults who stutter. *Brain*, 1161–1171. [PubMed: 29394325]
- D’Esposito M, & Postle BR (2015). The Cognitive Neuroscience of Working Memory. *Annual Review of Psychology*, 66(1), 115–142.

- Datta A, Baker JM, Bikson M, & Fridriksson J (2011). Individualized model predicts brain current flow during transcranial direct-current stimulation treatment in responsive stroke patient. *Brain Stimulation*, 4(3), 169–174. [PubMed: 21777878]
- Datta A, Truong D, Minhas P, Parra LC, & Bikson M (2012). Inter-Individual Variation during Transcranial Direct Current Stimulation and Normalization of Dose Using MRI-Derived Computational Models. *Front Psychiatry*, 3, 91. [PubMed: 23097644]
- Davidson L (2006a). Phonology, phonetics or frequency: Influences on the production of non-native sequences. *Journal of Phonetics*, 34(1), 104–137.
- Davidson L (2006b). Phonotactics and articulatory coordination interact in phonology: evidence from non-native production. *Cognitive Science*, 30(5), 837–862. [PubMed: 21702839]
- Davidson L (2007). The relationship between the perception of non-native phonotactics and loanword adaptation. *Phonology*, 24(2), 261–286.
- Davidson L (2010). Phonetic bases of similarities in cross-language production: Evidence from English and Catalan. *Journal of Phonetics*, 38(2), 272–288.
- de Aguiar V, Paolazzi CL, & Miceli G (2015). tDCS in post-stroke aphasia: The role of stimulation parameters, behavioral treatment and patient characteristics. *Cortex*, 63, 296–316. [PubMed: 25460496]
- Edwards D, Krebs H, Rykman A, Zipse J, Thickbroom G, Mastaglia F, et al. (2009). Raised corticomotor excitability of M1 forearm area following anodal tDCS is sustained during robotic wrist therapy in chronic stroke. *Restorative neurology and neuroscience*, 27(3), 199–207. [PubMed: 19531875]
- Flöel A (2014). tDCS-enhanced motor and cognitive function in neurological diseases. *NeuroImage*, 85, 934–947. [PubMed: 23727025]
- Giacobbe V, Krebs H, Volpe B, Pascual-Leone A, Rykman A, Zeierati G, et al. (2013). Transcranial direct current stimulation (tDCS) and robotic practice in chronic stroke: The dimension of timing. *NeuroRehabilitation*, 33(1), 49–56. [PubMed: 23949028]
- Guenther FH, Ghosh SS, & Tourville JA (2006). Neural Modeling and Imaging of the Cortical Interactions Underlying Syllable Production. *Brain and language*, 96(3), 280–301. [PubMed: 16040108]
- Holland R, & Crinion J (2011). Can tDCS enhance treatment of aphasia after stroke? *Aphasiology*, 26(9), 1169–1191. [PubMed: 23060684]
- Horvath JC, Carter O, & Forte JD (2014). Transcranial direct current stimulation: five important issues we aren't discussing (but probably should be). *Frontiers in systems neuroscience*, 8, 2–2. [PubMed: 24478640]
- Horvath JC, Forte JD, & Carter O (2015a). Evidence that transcranial direct current stimulation (tDCS) generates little-to-no reliable neurophysiologic effect beyond MEP amplitude modulation in healthy human subjects: A systematic review. *Neuropsychologia*, 66, 213–236. [PubMed: 25448853]
- Horvath JC, Forte JD, & Carter O (2015b). Quantitative Review Finds No Evidence of Cognitive Effects in Healthy Populations From Single-session Transcranial Direct Current Stimulation (tDCS). *Brain Stimulation*, 8(3), 535–550. [PubMed: 25701175]
- Kang N, Summers JJ, & Cauraugh JH (2015). Transcranial direct current stimulation facilitates motor learning post-stroke: a systematic review and meta-analysis. *Journal of Neurology, Neurosurgery & Psychiatry*.
- Katz B, Au J, Buschkuehl M, Abagis T, Zabel C, Jaeggi SM, et al. (2017). Individual Differences and Long-term Consequences of tDCS-augmented Cognitive Training. *Journal of Cognitive Neuroscience*, 29(9), 1498–1508. [PubMed: 28253083]
- Lametti DR, Smith HJ, Freidin PF, & Watkins KE (2018). Cortico-cerebellar Networks Drive Sensorimotor Learning in Speech. *Journal of Cognitive Neuroscience*, 30(4), 540–551. [PubMed: 29211651]
- Maas E, Robin DA, Hula SNA, Freedman SE, Wulf G, Ballard KJ, et al. (2008). Principles of motor learning in treatment of motor speech disorders. *American Journal of Speech-Language Pathology*, 17(3), 277–298. [PubMed: 18663111]

- Medina J, & Cason S (2017). No evidential value in samples of transcranial direct current stimulation (tDCS) studies of cognition and working memory in healthy populations. *Cortex*, 94, 131–141. [PubMed: 28759803]
- Meinzer M, Jähnigen S, Copland DA, Darkow R, Grittner U, Avirame K, et al. (2014). Transcranial direct current stimulation over multiple days improves learning and maintenance of a novel vocabulary. *Cortex*, 50, 137–147. [PubMed: 23988131]
- Meinzer M, Lindenbergh R, Sieg M, Nachtigall L, Ulm L, & Flöel A (2014). Transcranial direct current stimulation of the primary motor cortex improves word-retrieval in older adults. *Frontiers in Aging Neuroscience*, 6.
- Nilsson J, Lebedev AV, Rydström A, & Lövdén M (2017). Direct-Current Stimulation Does Little to Improve the Outcome of Working Memory Training in Older Adults. *Psychological Science*, 28(7), 907–920. [PubMed: 28509625]
- Nitsche MA, & Paulus W (2001). Sustained excitability elevations induced by transcranial DC motor cortex stimulation in humans. *Neurology*, 57(10), 1899–1901. [PubMed: 11723286]
- Nitsche MA, Schauenburg A, Lang N, Liebetanz D, Exner C, Paulus W, et al. (2003). Facilitation of Implicit Motor Learning by Weak Transcranial Direct Current Stimulation of the Primary Motor Cortex in the Human. *Journal of Cognitive Neuroscience*, 15(4), 619–626. [PubMed: 12803972]
- Nozari N, Woodard K, & Thompson-Schill SL (2014). Consequences of Cathodal Stimulation for Behavior: When Does It Help and When Does It Hurt Performance? *PLoS ONE*, 9(1), e84338. [PubMed: 24409291]
- Orban de Xivry J-J, & Shadmehr R (2014). Electrifying the motor engram: effects of tDCS on motor learning and control. *Experimental Brain Research*, 232(11), 3379–3395. [PubMed: 25200178]
- Psychology Software Tools Inc. (2016). Eprime 3.0. Pittsburgh, PA.
- R Core Team. (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Sadagopan N, & Smith A (2013). Age Differences in Speech Motor Performance on a Novel Speech Task. *Journal of Speech, Language, and Hearing Research*, 56(5), 1552–1566.
- Sasisekaran J, Smith A, Sadagopan N, & Weber-Fox C (2010). Nonword repetition in children and adults: effects on movement coordination. *Developmental Science*, 13(3), 521–532. [PubMed: 20443972]
- Schmidt RA, & Lee T (2005). *Motor Control and Learning: A behavioral emphasis*. Champaign, IL: Human kinetics.
- Segawa JA, Tourville JA, Beal DS, & Guenther FH (2015). The neural correlates of speech motor sequence learning. *Journal of cognitive neuroscience*, 27(4), 819–831. [PubMed: 25313656]
- Shah-Basak PP, Norise C, Garcia G, Torres J, Faseyitan O, & Hamilton RH (2015). Individualized treatment with transcranial direct current stimulation in patients with chronic non-fluent aphasia due to stroke. *Frontiers in Human Neuroscience*, 9, 201. [PubMed: 25954178]
- Sriraman A, Oishi T, & Madhavan S (2014). Timing-dependent priming effects of tDCS on ankle motor skill learning. *Brain research*, 1581, 23–29. [PubMed: 25063361]
- Stagg CJ, Jayaram G, Pastor D, Kincses ZT, Matthews PM, & Johansen-Berg H (2011). Polarity and timing-dependent effects of transcranial direct current stimulation in explicit motor learning. *Neuropsychologia*, 49(5), 800–804. [PubMed: 21335013]
- Stagg CJ, & Nitsche MA (2011). Physiological basis of transcranial direct current stimulation. *Neuroscientist*, 17(1), 37–53. [PubMed: 21343407]
- Steinberg Lowe M, & Buchwald A (2017). The Impact of Feedback Frequency on Performance in a Novel Speech Motor Learning Task. *Journal of Speech, Language, and Hearing Research*, 60(6S), 1712–1725.
- Vöröslakos M, Takeuchi Y, Brinyiczki K, Zombori T, Oliva A, Fernández-Ruiz A, et al. (2018). Direct effects of transcranial electric stimulation on brain circuits in rats and humans. *Nature Communications*, 9(1), 483.
- Westwood SJ, Olson A, Miall RC, Nappo R, & Romani C (2017). Limits to tDCS effects in language: Failures to modulate word production in healthy participants with frontal or temporal tDCS. *Cortex*, 86, 64–82. [PubMed: 27912107]

- Wiethoff S, Hamada M, & Rothwell JC (2014). Variability in Response to Transcranial Direct Current Stimulation of the Motor Cortex. *Brain Stimulation*, 7(3), 468–475. [PubMed: 24630848]
- Wilson C, Davidson L, & Martin S (2014). Effects of acoustic–phonetic detail on cross-language speech production. *Journal of Memory and Language*, 77, 1–24.
- Wortman-Jutt S, & Edwards DJ (2017). Transcranial Direct Current Stimulation in Poststroke Aphasia Recovery. *Stroke*, 48(3), 820–826. [PubMed: 28174328]
- Yokoi A, Arbuckle SA, & Diedrichsen J (2018). The role of human primary motor cortex in the production of skilled finger sequences. *J Neurosci*.

Figure 1:

Depiction of the procedure for the experiment. All participants wore the tDCS set-up for 40 minutes (during the non-language tasks and the Speech learning task). For the two Before groups, tDCS was turned on with active or sham stimulation and ran for 20 minutes prior to any sessions of the speech motor learning task. For the two During groups, the tDCS was turned on with active or sham stimulation as the pre-practice session began.

Figure 2:

Electrode montage and current flow estimation for the Experiment. (A) depicts the electrode montage on a model head. (B) provides coronal and sagittal images of the output of the HD-Explore software for that montage on a standardized brain (see Datta, Baker, Bikson, & Fridriksson, 2011; and Datta, Truong, Minhas, Parra, & Bikson, 2012 for details on the simulations).

Figure 3:

Waveform, spectrogram, and segmentation of two tokens of /fniku/. (A) The top token was produced accurately with no epenthetic vowel. (B) The bottom token was produced with an epenthetic vowel in the cluster, indicated by the vocoid in the waveform and formant structure in the spectrogram.

Figure 4. Change in performance on accuracy measures across acquisition and retention. (A) depicts overall whole nonword accuracy for each quintile during practice (acquisition), as well as at shorter-term retention (R1) and longer-term retention (R2). (B) depicts overall cluster accuracy for each quintile during practice (acquisition), as well as R1 and R2. For all data points, error bars represent standard error.

Figure 5. Change in performance on duration measures across acquisition and retention. Figure depicts overall epenthesis error duration for each quintile during practice (acquisition), as well as R1 and R2. Error bars represent standard error.

Table 1:

Summary of logistic mixed-effects model output for within-group accuracy comparisons. Statistically significant differences are shown in bold.

<i>Within-group whole nonword accuracy changes</i>										
		<i>Before</i>			<i>During</i>			<i>Sham</i>		
		$\hat{\beta}$	z	p	$\hat{\beta}$	z	p	$\hat{\beta}$	z	p
	Quintile	0.14	4.43	<0.0001	0.16	4.93	<0.0001	0.07	2.98	0.0029
Q1 vs.	R1	0.59	5.56	<0.0001	0.27	2.53	0.0114	0.25	3.30	0.0009
	R2	0.53	4.98	<0.0001	0.53	4.87	<0.0001	0.39	5.08	<0.0001
Q5 vs.	R1	0.07	0.64	0.5207	-0.43	-4.00	<0.0001	-0.03	-0.39	0.6962
	R2	0.01	0.06	0.9550	-0.18	-1.64	0.1013	0.10	1.41	0.1586

<i>Within-group cluster accuracy changes</i>										
		<i>Before</i>			<i>During</i>			<i>Sham</i>		
		$\hat{\beta}$	z	p	$\hat{\beta}$	z	p	$\hat{\beta}$	z	p
	Quintile	0.10	3.09	0.0020	0.09	2.76	0.0058	0.04	1.81	0.0703
Q1 vs.	R1	0.36	3.37	0.0007	0.01	0.11	0.9116	0.16	2.06	0.0396
	R2	0.29	2.72	0.0066	0.22	1.99	0.0457	0.28	3.58	0.0003
Q5 vs.	R1	0.04	0.37	0.7153	-0.35	-3.20	0.0014	0.01	0.09	0.9283
	R2	0.03	0.29	0.7740	0.14	-1.29	0.1981	0.13	1.63	0.1648