Comparing the effects of frontal and temporal neurostimulation on second language learning

Kinsey Bice & Chantel S. Prat
{klbice, csprat}@uw.edu
Department of Psychology and
Institute for Learning & Brain Sciences,
University of Washington

Abstract

Successful language learning requires a dynamic balance between declarative and procedural mechanisms, yet individuals may engage them in less than optimal ways. The goal of the current experiment was to determine whether transcranial direct current stimulation (tDCS) can tip the balance, specifically facilitating declarative or procedural learning. Seventy-nine subjects (31 no stimulation, 16 sham stimulation, 16 temporal, 16 frontal) completed an artificial grammar learning task followed by a two-alternative forced-choice test measuring sensitivity to the underlying grammar (procedural) versus the surface form (declarative). The pattern of results is consistent with separate engagement of declarative and procedural systems. Left temporal stimulation resulted in higher selection of strings with familiar surface features. In contrast, frontal stimulation resulted in a slower learning trajectory and more frequent selection of grammatical letter strings. We conclude that tDCS may be used to facilitate engagement of different learning systems required for language learning.

Keywords: language learning; tDCS; declarative learning; procedural learning

Introduction

Learning a new language as an adult is a difficult, yet often desirable process, which leads to highly variable outcomes between individuals. Successful language learning depends on the coordinated deployment of declarative and procedural systems at different time points, and for different types of knowledge acquisition (e.g., vocabulary vs. grammar learning). It has been proposed that individual differences in learning success may relate to differences in the degree to which adults employ declarative versus procedural learning mechanisms (e.g., Morgan-Short, Faretta-Stutenberg, Brill-Schuetz, Carpenter, & Wong, 2014). Because these systems rely on distinct brain circuits, the current study investigated the possibility of pushing learners to use one system over the other by employing neurostimulation techniques.

Declarative and procedural systems have distinguishing features and play independent roles in language learning. Explicit, declarative learning tends to be quite fast, even one-shot learning, and depends on capacity-limited resources like working memory. The knowledge produced by declarative systems is verbalizable, such that declarative memory is critical for successful vocabulary acquisition. One of the primary regions of the brain associated with declarative learning is the medial temporal lobe (Ullman, 2001), and one of the primary regions associated with vocabulary knowledge (word forms, form-meaning mappings) is the (left) temporal lobe (Price, 2010). In contrast, procedural systems are primarily responsible for implicit knowledge that cannot be easily verbalized, such that they are critical for achieving grammatical fluency. Procedural learning is considerably slower and requires many repeated iterations from which one can develop expectations and learn from errors or feedback. Yet retrieval from declarative memory tends to be relatively slower than procedural memory, which can feel almost reflexive. The dorso-lateral prefrontal cortex is one of the primary regions involved in grammatical fluency (Price, 2010), and other types of procedural learning, such as reinforcement learning, have been found to depend on other frontal and subcortical regions (Ullman, 2001).

It is important to note that during complex skill acquisition, declarative and procedural systems can compete with each other; for instance, engaging declarative learning processes can inhibit access to procedural learning (Ashby & Crossley, 2010; Collins, 2018; Foerde, Knowlton, & Poldrack, 2006). The declarative-procedural model of Ullman (2001) proposes that adults may be biased to rely on declarative learning due to their highly developed cognitive resources, which may inhibit the grammar from gaining access to the more optimal procedural learning system.

The goal of the current study was to determine whether we could drive reliance on procedural or declarative learning mechanisms using neurostimulation parameters (transcranial direct current stimulation, or tDCS) designed to excite one pathway over the other. We examined whether anodal tDCS across the left temporal lobe specifically impacted measures of declarative learning and whether anodal tDCS traveling through medial frontal and subcortical regions affected measures of procedural learning. The ultimate goal of this research is to develop methods for using neuroengineering to enhance language learning in adults by influencing their use of the optimal learning pathways at particular points in time.

Transcranial Direct Current Stimulation

tDCS is a neurostimulation method that sends a constant flow of low-voltage electricity from the anode to the cathode. Depending on where the anode and cathode are placed on the head, the current follows a path through the underlying brain regions, and has been shown to affect cortical excitability depending on the polarity of the current (anodal vs. cathodal; Lang, Nitsche, Paulus, Rothwell, & Lemon, 2004). Past studies have demonstrated that tDCS can influence declarative and procedural processes in healthy adults. Anodal tDCS (atDCS) over the left temporal lobe

(Wernicke's area) has been shown to facilitate associative verbal learning and lexical access (Monti et al., 2013), both of which are declarative processes. Neuroimaging research has also highlighted the importance of the left temporal lobe for word form representation and for the mapping between word forms and meanings (Price, 2010) and the medial temporal lobe for declarative learning (Ullman, 2001). Therefore, we hypothesized that anodal stimulation of the left temporal lobe would specifically enhance declarative learning related to word forms or vocabulary.

Other studies have shown that atDCS over frontal regions affects procedural processes. Stimulation of the left frontal lobe facilitates artificial grammar learning and grammatical decisions (Monti et al., 2013). Similarly, atDCS over the frontal midline region enhanced indices of brain activity that are related to reinforcement learning (Reinhart & Woodman, 2015), including error-related and feedback-related negativity, which may help with the incremental process of acquiring and generalizing grammar. Therefore, we predicted that atDCS of the left medial frontal lobe would enhance procedural learning related to pattern extraction or grammar.

Artificial Grammar Learning

Artifical grammar learning (AGL) tasks, first reported by Reber (1967), are widely used to study procedural learning. They are an incidental learning task in which participants reproduce (copy and/or type) letter strings that were formed by following legal pathways through a finite-state grammar under the guise of some other task (e.g., memory). At test, they are presented with a two alternative-forced-choice (2AFC) task in which they must select which of two novel letter strings fit the pattern. Participants reliably perform above chance at selecting grammatical strings, and studies have demonstrated that a successfully learned artificial grammar engages the same regions as the grammar of a natural language (Petersson, Folia, & Hagoort, 2010).

Since Reber's initial study, a field of research has developed to investigate various aspects of AGL tasks. While the task was developed to specifically measure procedural learning, some have pointed out that one can accomplish the task using more explicit, declarative strategies. One declarative strategy for an AGL task is to attend to the chunk strength of the items during training and use it as a cue to distinguish items at test. Chunk strength refers to the surface similarity; by repeatedly following pathways through the artificial grammar, common "chunks" of letters appear more often than others, providing a clue that those chunks are grammatical. Essentially, it refers to the strategy of trying to memorize the training items, which contain these frequent chunks, and then comparing the novel items at test to those training items in memory based on similarity. It is considered a declarative strategy because it depends on episodic (explicit) memories of repeated exemplars. More evidence that chunk strength depends on declarative systems comes from AGL studies using amnesiac patients, who can reliably make grammaticality judgments but do not distinguish high and low chunk strength items (Knowlton, Ramus, & Squire, 1992). Interestingly, some research has found individual differences in cue bias (chunk strength vs. grammar), showing that some individuals attend more to the chunk strength cues than the grammar of the letter strings in making their decisions at test (McAndrews & Moscovitch, 1985).

Of relevance for the current study, the test items can be orthogonally manipulated on the dimensions of grammaticality and chunk strength. By presenting a 2AFC test between items that are similar in chunk strength and only differ with respect to grammaticality, an individual's reliance on the procedural learning system can be isolated; likewise, presenting two items that are both (non)grammatical and only differ in chunk strength can isolate an individual's reliance on more declarative learning systems.

In the current study, we hypothesized that neurostimulation would influence individuals' biases to attend to chunk strength or grammar in the AGL task. Specifically, the prediction was that participants who received left temporal atDCS should demonstrate patterns of learning and test consistent with declarative learning; they should learn rapidly and exhibit greater sensitivity to chunk strength at test, similar to the word form knowledge of vocabulary. The prediction for participants who received frontal atDCS was that they should demonstrate patterns of learning and test consistent with procedural learning; learning may be initially slow, and they should exhibit greater sensitivity to the grammaticality of an item at test.

Methods

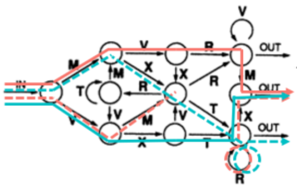
Participants

A total of 96 people participated in the study, of which 17 had to be excluded due to software difficulties (n = 10) and poor atDCS contact quality (n = 6). One subject in the sham stimulation condition was excluded for noticing that the stimulation was off during most of the task, due to the fact that the stimulation induced phosphenes.

Of the remaining 79 participants, 31 received no stimulation (19 female), 16 received sham stimulation (10 female), 16 received anodal left temporal lobe stimulation (8 female), and 16 received anodal frontal lobe stimulation (6 female). The participants with no stimulation were recruited separately (through the university subject pool) from those who received sham, temporal, or frontal stimulation, due to the different eligibility requirements related to the use of neurostimulation. All participants spoke English as their native language, and all participants who received sham, temporal, or frontal stimulation were right-handed in addition to meeting the safety eligibility requirements associated with the use of the tDCS (Bikson et al., 2016).

Artificial Grammar Learning

The grammar used for the current study was reported in Vokey and Brooks (1992; Figure 1). A total of 24 grammatical letter strings were created and used for training by following legal pathways through the grammar. Letter strings ranged from 3 letters to 8 letters long.



2AFC Test Conditions

Measure	Example
Chunk strength isolated (grammaticality constant)	VXTX vs. MXTR
Grammaticality isolated (chunk strength constant)	MVRM vs. VMRX

Figure 1: Artificial grammar used (Vokey & Brooks, 1992) with examples of a "correct" pathway (solid lines) and an "incorrect" pathway of the 2AFC test (dashed lines) for each condition of interest. While the incorrect grammatical pathway can be visualized easily, the incorrect (low chunk strength) pathway depends upon items presented at training.

During training, a single letter string appeared in the center of the screen for 5 s, during which participants were told to memorize it. A brief (500 ms) blank screen intervened, after which participants were prompted to re-type the letter string and press enter to submit their response. They were provided feedback (correct/incorrect). Participants who incorrectly typed a letter string had to re-do the trial until they were correct. Each letter string was presented (until correctly re-typed) twice, once during each of the two blocks of training.

At test, four types of items were created for presentation that varied orthogonally on grammaticality (grammatical, nongrammatical) and chunk strength (low, high). Figure 1 illustrates how these test items were created from the grammar. Grammatical, high chunk strength test items followed legal but novel pathways through the grammar (e.g., MVRM). Grammatical, low chunk strength test items followed legal pathways through the grammar that had not been frequently presented in the process of creating training items (e.g., MXTR). Nongrammatical test strings were created by following pathways through the grammar and then transposing letters that rendered the string illegal or by shuffling grammatical letter strings (e.g., VMRX, MRXV), and were sorted into high and low chunk strength by their similarity to training items.

The training items were used as the database for calculating chunk strength. Each training item was split into every possible smaller chunk (e.g., MVXT was split into MVX, VXT, MV, VX, and XT). The frequency of each chunk was calculated (e.g., MVXT appeared 4 times across all training items, MVX appeared 8 times, etc.). A similar operation was conducted on the test items to determine how frequent each of the chunks had appeared during training. Chunk strength was calculated by summing the frequency of each chunk of the test item as it appeared in the training items, divided by the number of chunks in the test item. For example, the test item MVXR had a chunk strength of 9 (MVXR = 4, MVX = 8, VXR = 4, MV = 11, VX = 18, XR = 9, divided by a total of 6 chunks). An ANOVA on the test items' chunk strength values with grammaticality (grammatical, nongrammatical) and chunk strength (high, low) was conducted to verify that only the chunk strength manipulation differed. The results showed a main effect of chunk strength ($F(1, 92) = 12.05$, $p < .01$) in the expected direction, and no main effect of grammaticality ($F(1, 92) = 0.71$, $p = .4$) nor an interaction ($F(1, 92) = 0.39$, $p = .53$).

The test items were presented in pairs during the 2AFC test. Participants read each letter string and used the left/right arrow keys to indicate which of the two letter strings they believed was a "better fit" for the pattern of items during training. The 2AFC test had two conditions of interest. One condition was designed to measure grammatical sensitivity, by presenting two items that differed in grammaticality but had similar chunk strength (both items had high chunk strength or low chunk strength). The other condition measured chunk strength sensitivity, by presenting two items that differed in chunk strength but were both either grammatical or nongrammatical.

atDCS Stimulation Procedures and Parameters

Two 3 x 5 cm sponges were soaked with saline solution, and electrodes were inserted into the sponges. Head measurements were taken to record the circumference and nasion-inion distance for subsequent electrode placement. For frontal stimulation, the anode was placed by following 30% of the nasion-inion distance from the nasion on the midline, approximately the location of electrode Fz in the 10-20 international electrode placement system. The sponge was placed horizontally (i.e., the long side of the sponge was perpendicular to the midline) and secured with an elastic

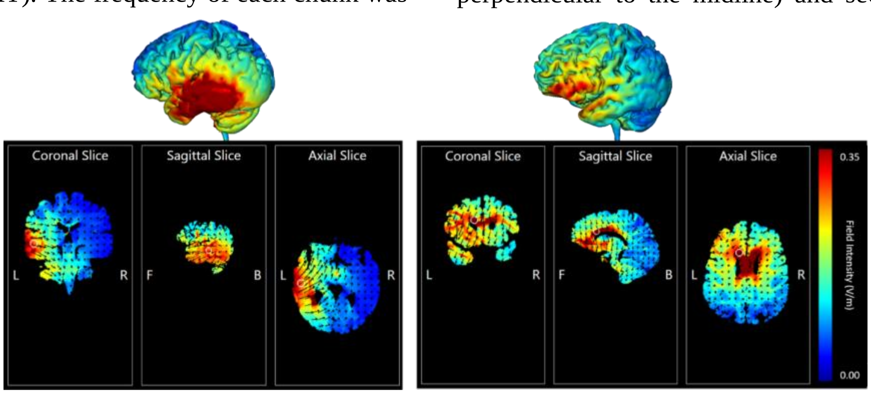


Figure 2: Current flow model generated by Soterix HD Explore software for 1.5 mA of current. Left: temporal condition with anode electrode placement over P7 and TP7 and cathode electrode placement over EX11 . Right: frontal condition with anode electrode placement over Fz, F1, and F2, and cathode over EX11.

headband. The cathode was placed on the left cheek between the jawline and cheekbone, vertically, and was also secured with an elastic band. For temporal stimulation, the anode placement was determined by following 10% of the nasion-inion distance from the inion up the midline (approximately electrode Oz), and then 10% of the head circumference from the previous Oz location to the left along the circumference, approximately the location between electrodes P7 and TP7 of the international 10-20 system. The cathode was once again placed on the left cheek, vertically between the jawline and cheekbone. Both sponges were secured with an elastic band. Figure 2 illustrates the model of the current flow given the current intensity and electrode placement for the two stimulation conditions. For sham stimulation, the 'anode' was placed over the right hemisphere, approximately the location between electrodes P8 and TP8. The 'cathode' was placed on the right supraorbital area.

Once the electrode sponges were placed and secured, the impedance level of the electrodes was reduced to an acceptable level by parting the hair directly beneath the sponge and by applying additional saline solution as needed. Frontal and temporal stimulation reached 1.5 mA and lasted for 20 minutes. For sham stimulation, the direct current ramped up to 1.5 mV, taking approximately 30 s, and then immediately ramped back down, at the very beginning and very end of the 20 minutes.

Procedure

Upon arrival, participants first provided informed consent, approved by the Institutional Review Board at the University of Washington. Participants completed three computerized questionnaires: the tDCS safety form (sham/stim groups only), a demographics form, and a shortened version of the Edinburgh handedness questionnaire. Participants also completed an additional task to measure phonological working memory (LLAMA-D; Meara, 2005).

Participants were told that their goal was to accurately reproduce the letter strings on the screen, under the guise of a (working) memory task. The stimulation commenced simultaneously with the task and continued throughout the duration of AGL training and test. Upon completing the AGL training, participants were informed that the letter strings that they were previously re-typing followed a pattern. They were instructed that they would view two novel letter strings on the screen and should indicate with the arrow keys (left/right) which of the two letter strings they believe were a better fit for the pattern in the letter strings presented at training. They were asked to make their judgement quickly, to not overthink the decision too closely but instead to use their intuition.

After the AGL task was completed, the electrode sponges were removed and participants completed a final set of questionnaires: a modified version of the LEAP-Q to report any previous language-learning experience (Marian, Blumenfeld, & Kaushanskaya, 2007) and a debriefing form to report their experience with the stimulation. The entire session lasted one hour on average. Participants who received sham, temporal, or frontal stimulation were paid for their time, and participants in the no stimulation condition were awarded course credits.

Results

Predictions

The prediction was that left temporal stimulation should facilitate the engagement of declarative learning pathways. Therefore, we expected to see fast learning and evidence that surface features of the AGL task were preferentially learned (i.e., chunk strength).
In contrast, frontal stimulation was predicted to facilitate procedural learning. Therefore, we expected the patterns to reflect a slower learning trajectory paired with evidence that the underlying grammatical patterns were preferentially learned.

Dependent Variables

For training, the primary dependent variable was training accuracy. Accuracy was calculated by determining the number of typing attempts during training that were accurate. Often, a participant had to re-type the same letter string up to four or five times, and each counted against their accuracy.

For the 2AFC test, the primary dependent variable was also accuracy. Accuracy rates were calculated as the proportion of trials in which the subjects selected the "correct" option. For the condition in which grammaticality was isolated (chunk strength held constant), the correct option corresponded to the grammatical option; for the condition in which chunk strength was isolated (grammaticality held constant), the high chunk strength item was the correct option.

Group Comparisons

The purpose of collecting data from the no stimulation group was to verify that sham stimulation did not influence performance, but also that the participants receiving sham stimulation still believed they had received real stimulation. The results of the debriefing form indicated that the participants in the sham condition believed they experienced true stimulation; a similar proportion of participants in the sham condition reported "discomfort or pain" related to the tDCS (11/16 subjects, compared to 8/16 in frontal and 10/16 in temporal), and their rating of the discomfort or pain was at similar levels as the real stimulation groups (scale: 1 (mild discomfort) to 5 (significant pain); $M_{\text{sham}} = 1.17$, $M_{\text{frontal}} = 1.20$, $M_{\text{temporal}} = 1.36$; all ps > .4).

To verify that the sham stimulation did not significantly influence performance, we first conducted a set of analyses to compare the performance between the no stimulation and sham stimulation groups on the measures of interest. No differences were found in AGL performance between the no stimulation and sham stimulation groups, whereas there were a number of demographic differences between the no stimulation group and the three other groups due to recruitment and sampling differences (results reported below). Therefore, in order to maintain balanced groups, and to ensure comparisons between maximally similar experiences, only the sham group was included as the baseline for the comparisons with the real stimulation groups.

Baseline and Demographics

The AGL task does not include any baseline learning block to consider between-group pre-existing differences in learning. However, past work has shown that a person's ability to rehearse the letter strings affects learnability of the AG (Andrade & Baddeley, 2011), and that previous experience with learning a second language leads to faster learning of a new language (Kaushanskaya & Marian, 2009). Therefore, to account for any pre-existing differences between groups, we investigated performance on the LLAMA task, measuring phonological language learning aptitude, and a number of background variables (age, previous second language learning experience).

A comparison of the no stimulation and sham stimulation groups revealed that that no stimulation group was significantly younger than the sham stimulation group ($t(45) = 3.81$, $p < .01$). The no stimulation group also had marginally more and earlier experience with learning a second language (L2) [L2 speaking: $t(30) = 1.87$, $p = .07$; L2 understanding: $t(30) = 1.78$, $p = .08$; L2 age of acquisition: $t(30) = 2.03$, $p = .05$]. The two groups did not significantly differ in their performance on the LLAMA task ($t(45) = 0.96$, $p > .1$).

A comparison of the three stimulation groups using pairwise t-tests revealed no differences in age (all *ps* > .1) or in previous L2 experience [L2 speaking: all *ps* > .5; L2 understanding: all *ps* > .5; L2 age of acquisition: all *ps* > .3]. The three groups also had comparable performance on the LLAMA task (all *ps* > .5). The stimulation groups, recruited and tested under similar experimental conditions, were comparable populations without any known pre-existing differences that have been found to affect AGL performance.

Training

An ANOVA was conducted on training accuracy for each block, with group ($ANOVA_1$: no stimulation, sham; $ANOVA_2$: sham, temporal, frontal) as a between-subjects variable, and block (1, 2) as a within-subjects variable. The comparison between the no stimulation and sham stimulation group revealed a main effect of block ($F(1, 42) = 41.88.67$, $p < .01$) but there was no effect of group and no interaction (*ps* > .3). The ANOVA with the three stimulation groups revealed a main effect of block ($F(1, 40) = 48.86$, $p < .01$) that captured the fact that participants became more accurate in the second block than in the first block. There was also a significant interaction between group and block ($F(2, 41) = 4.50$, $p = .02$). Follow-up Holm-corrected pairwise t-tests uncovered that the interaction reflected between-group differences in accuracy in the first block but similar accuracy rates by the second block (see Figure 3). Specifically, the pairwise comparisons of the accuracy in the first block between stimulation conditions revealed significantly higher accuracy in the temporal than frontal condition ($p = .03$) and marginally higher accuracy in the sham condition than the frontal condition ($p = .06$), but no differences were found between stimulation conditions for the accuracy rates in the second block (all *ps* > .9).

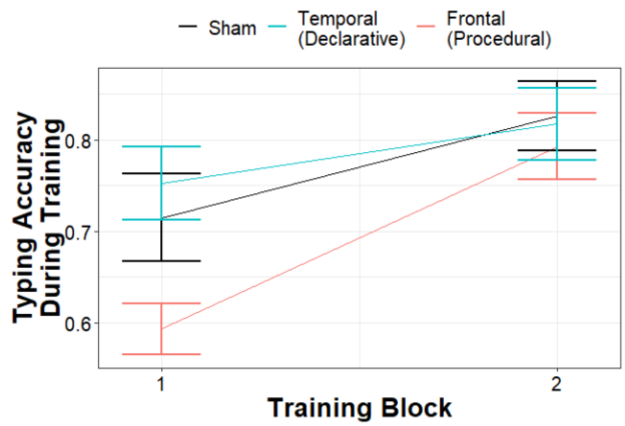


Figure 3: Typing accuracy rates during blocks 1 and 2 of training as a function of stimulation condition. Error bars represent one standard error.

Test

To analyze performance at test, an ANOVA was conducted with stimulation group as a between-subjects variable ($ANOVA_1$: no stimulation, sham; $ANOVA_2$: sham, frontal, temporal) and the 2AFC comparison as a within-subjects variable (grammar isolated, chunk strength isolated). The comparison between no stimulation and sham stimulation revealed no main effects and no interaction (all *ps* > .2). The comparison between the three stimulation groups revealed a marginally significant group x 2AFC comparison interaction ($F(2, 46) = 2.97$, $p = .06$). The results of the accuracy rates for the 2AFC test are depicted in Figure 4.

A priori one-tailed t-tests were conducted to evaluate the hypotheses that the frontal stimulation group would exhibit higher accuracy in the grammar isolated 2AFC condition and that the temporal stimulation group would exhibit higher accuracy in the chunk strength isolated 2AFC condition. The results of the t-test on the accuracy in the grammar isolated condition revealed that the frontal stimulation group had significantly higher accuracy than the temporal stimulation group ($t(31) = 1.78$, $p = .04$) and marginally higher accuracy than the sham stimulation group ($t(30) = 1.42$, $p = .08$). The results of the t-tests on the accuracy in the chunk strength isolated condition revealed that the temporal stimulation group had significantly higher accuracy than the frontal stimulation group ($t(31) = 1.83$, $p = .04$) and marginally higher accuracy than the sham stimulation group ($t(31) = 1.44$, $p = .08$).

Discussion

The current study used atDCS to shape how people learned new information. atDCS was applied to two different brain regions associated with declarative and procedural learning while participants performed an artificial grammar learning task. The results were as predicted in showing patterns of

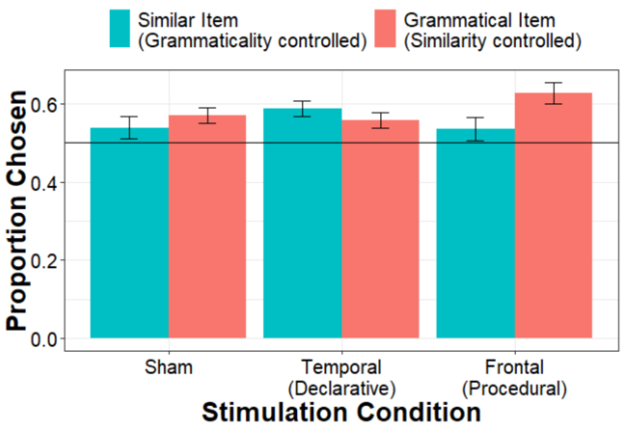


Figure 4: Accuracy rates for performance at test. Error bars represent one standard error. Horizontal black line represents chance (50%) performance. Red bars are the accuracy rates for the 2AFC condition in which grammaticality was isolated; blue bars are the accuracy rates for the 2AFC condition in which chunk strength was isolated.

learning and memory consistent with declarative systems for the group who received temporal stimulation, and with procedural systems for the group who received frontal stimulation. The temporal stimulation group was specifically more accurate at selecting high chunk strength items at test when grammaticality was controlled for, suggesting that their learned representations included more explicit information about frequent letter chunks. In contrast, the frontal stimulation group had significantly slower initial learning, but caught up to the performance of other groups after a number of repeated iterations. This pattern is consistent with the trajectory of procedural learning, which depends heavily on repeated experience with errors and feedback. Likewise, the frontal stimulation group demonstrated higher accuracy specifically for selecting grammatical items at test when the chunk strength was controlled for. Taken together, the pattern of results observed in the temporal and frontal stimulation groups was consistent with what would be expected based on increased declarative and procedural learning systems, respectively.

It is important to note that differences in training accuracy emerged early, during the first block of training. The fact that the temporal and sham stimulation groups performed similarly, while the frontal stimulation group had significantly lower accuracy in the first block only, suggests a specific effect of frontal stimulation on early learning. Such a pattern could be interpreted in light of studies that have found competition between declarative and procedural learning systems (e.g., Ashby & Crossley, 2010; Foerde, Knowlton, & Poldrack, 2006). By shifting the learning pathways through procedural routes, declarative resources may be relatively inhibited, resulting in worse initial accuracy. Because procedural learning proceeds via error and feedback, it is slower and relies on repetition. However, once the underlying patterns are configured, procedural learning can be quite robust, which is consistent with the finding that the frontal stimulation group did not perform significantly differently than the temporal or sham groups during the second training block.

As a between-subjects design, we cannot rule out the possibility that the early differences observed during training were due to pre-existing differences between the individuals in each group rather than the stimulation. The nature of the AGL task prohibited a within-subjects design due to the surprise 2AFC test. However, the three stimulation groups were well-matched on other measured background (e.g., demographic, second-language experience) and cognitive (e.g., phonological working memory) measures. The idea that the early differences in training were due to stimulation is also supported by studies using event-related designs that report tDCS effects on the scale of (milli)seconds (Furubayashi et al., 2008; Javadi, Cheng, & Walsh, 2012). Given that the difference between groups was eliminated by the second block of training, rather than persisting throughout the task, the pattern of results is more consistent with the idea that the declarative and procedural systems compete. Future research should consider applying stimulation prior to training and test, and/or to extend learning across multiple sessions by using a more complex grammar. A longer training period would also enable a closer look at the trajectory of learning, especially given that a protracted learning trajectory is a core assumption of procedural learning models.

The reliability of tDCS has come under question in recent years, with doubt surrounding the efficacy and specificity of the effects (Horvath, Carter, & Forte, 2014). Most tDCS studies use repeated measures designs in which participants return for multiple sessions and receive counterbalanced administration of anodal, cathodal, and sometimes sham stimulation. Research designs such as those are important for establishing how polarity affects the levels of excitability of the targeted regions and whether a targeted region causally contributes to a cognitive process. The design of the current study would not allow a repeated-measures design because the surprise component of the AGL task prohibits multiple administrations (i.e., participants would know the second time they performed the task to attend to the underlying patterns). However, the question in the current study was not regarding polarity or the relevance of a given region; instead, the results demonstrated that tDCS can incur relatively specific benefits by showing the clear dissociation between procedural and declarative learning. These results are an important first step toward building tools to facilitate learning, and especially language learning, for adults.

It will be important for future research to demonstrate the generalizability of these findings for a natural language. The complexity of natural language learning is immense. It may not be that procedural learning is overall better, or even always better for grammar; instead, future research should focus on the interplay between declarative and procedural systems over time and how they may interact with individual differences prior to learning.

In summary, the present study demonstrated that tDCS can shape learning to engage pathways that are important for vocabulary and grammar processing. Anodal stimulation of the left temporal lobe enhanced learners' attention and memory for explicit chunks of letters in an artificial grammar learning task, akin to word forms. Anodal stimulation of the medial-left frontal lobe facilitated learners' ability to correctly identify grammatical letter strings. These findings are an important first step toward the development of tools to facilitate language learning for adults.

Acknowledgments

This work was supported by the Washington Research Foundation Innovation Postdoctoral Fellowship in Neuroengineering to KB and CSP and by an award from the Office of Naval Research (ONRBAA13-003) to CSP.

References

Andrade, J., & Baddeley, A. (2011). The contribution of phonological short-term memory to artificial grammar learning. *The Quarterly Journal of Experimental Psychology, 64*(5), 960-974.

Ashby, F. G., & Crossley, M. J. (2010). Interactions between declarative and procedural-learning categorization systems. *Neurobiology of Learning and Memory, 94*(1), 1-12.

Bikson, M., Grossman, P., Thomas, C., Zannou, A. L., Jiang, J., Adnan, T., ... & Brunoni, A. R. (2016). Safety of transcranial direct current stimulation: evidence based update 2016. *Brain Stimulation, 9*(5), 641-661.

Collins, A. G. (2018). The tortoise and the hare: Interactions between reinforcement learning and working memory. *Journal of Cognitive Neuroscience, 30*(10), 1422-1432.

Foerde, K., Knowlton, B. J., & Poldrack, R. A. (2006). Modulation of competing memory systems by distraction. *Proceedings of the National Academy of Sciences, 103*(31), 11778-11783.

Furubayashi, T., Terao, Y., Arai, N., Okabe, S., Mochizuki, H., Hanajima, R., ... & Ugawa, Y. (2008). Short and long duration transcranial direct current stimulation (tDCS) over the human hand motor area. *Experimental Brain Research, 185*(2), 279-286.

Horvath, J. C., Carter, O., & Forte, J. D. (2014). Transcranial direct current stimulation: five important issues we aren't discussing (but probably should be). *Frontiers in Systems Neuroscience, 8*, 2.

Javadi, A. H., Cheng, P., & Walsh, V. (2012). Short duration transcranial direct current stimulation (tDCS) modulates verbal memory. *Brain Stimulation, 5*(4), 468-474.

Kaushanskaya, M., & Marian, V. (2009). The bilingual advantage in novel word learning. *Psychonomic Bulletin & Review, 16*(4), 705-710.

Knowlton, B. J., Ramus, S. J., & Squire, L. R. (1992). Intact artificial grammar learning in amnesia: Dissociation of classification learning and explicit memory for specific instances. *Psychological Science, 3*(3), 172-179.

Lang, N., Nitsche, M. A., Paulus, W., Rothwell, J. C., & Lemon, R. N. (2004). Effects of transcranial direct current stimulation over the human motor cortex on corticospinal and transcallosal excitability. *Experimental Brain Research, 156*(4), 439-443.

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research, 50*(4), 940-967.

McAndrews, M. P., & Moscovitch, M. (1985). Rule-based and exemplar-based classification in artificial grammar learning. *Memory & Cognition, 13*(5), 469-475.

Meara, P. M. (2005). *Llama Language Aptitude Tests*. Swansea: Lognostics 2005.

Monti, A., Ferrucci, R., Fumagalli, M., Mameli, F., Cogiamanian, F., Ardolino, G., & Priori, A. (2013). Transcranial direct current stimulation (tDCS) and language. *Journal of Neurology, Neurosurgery, & Psychiatry, 84*(8), 832-842.

Morgan-Short, K., Faretta-Stutenberg, M., Brill-Schuetz, K. A., Carpenter, H., & Wong, P. C. (2014). Declarative and procedural memory as individual differences in second language acquisition. *Bilingualism: Language and Cognition, 17*(1), 56-72.

Petersson, K. M., Folia, V., & Hagoort, P. (2012). What artificial grammar learning reveals about the neurobiology of syntax. *Brain and Language, 120*(2), 83-95.

Price, C. J. (2010). The anatomy of language: a review of 100 fMRI studies published in 2009. *Annals of the New York Academy of Sciences, 1191*(1), 62-88.

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior, 6*(6), 855-863.

Reinhart, R. M., & Woodman, G. F. (2015). The surprising temporal specificity of direct-current stimulation. *Trends in Neurosciences, 38*(8), 459-461.

Ullman, M. T. (2001). The neural basis of lexicon and grammar in first and second language: The declarative/procedural model. *Bilingualism: Language and Cognition, 4*(2), 105-122.

Vokey, J. R., & Brooks, L. R. (1992). Salience of item knowledge in learning artificial grammars. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*(2), 328.