JSLHR

Research Note

A Neural Marker of Speech Intention: Evidence From Contingent Negative Variation

Juhi Kidwai,[a] Jonathan S. Brumberg,[b] and Brianna M. Marsh[c]

Purpose: This study investigated whether changes in brain activity preceding spoken words can be used as a neural marker of speech intention. Specifically, changes in the contingent negative variation (CNV) were examined prior to speech production in three different study designs to determine a method that maximizes signal detection in a speaking task.
Method: Electroencephalography data were collected in three different protocols to elicit the CNV in a spoken word task that varied the timing and type of linguistic information. The first protocol provided participants with the word to be spoken before the instruction of whether or not to speak, the second provided both the word and the instruction to speak, and the third provided the instruction to speak before the word. Participants (N = 18) were split into three groups (one for each protocol) and were instructed to either speak (Go) or refrain from speaking (NoGo) each word according to task instructions. The CNV was measured by analyzing the difference in slope between Go and NoGo trials.
Results: Statistically significant effects of hemispheric laterality on the CNV slope confirm the third protocol where the participants know they will speak in advance of the word, as the paradigm that reliably elicits a CNV response related to speech intention.
Conclusions: The maximal CNV response when the instruction is known before the word indicates the neural processing measured in this protocol may reflect a generalized speech intention process in which the speech-language systems become prepared to speak and then execute production once the word information is provided. Further analysis of the optimal protocol identified in this study requires additional experimental investigation to confirm its role in eliciting an objective marker of speech intention.
Supplemental Material: https://doi.org/10.23641/asha.14111468

Speech production is a complex process driven by a speaker's communicative intent to translate their ideas and thoughts to a listener and respond to others (Bara, 2010; Grice, 1975; Sperber & Wilson, 1995). The process of speech production consists of many overlapping components and has been studied from a range of theoretical contexts that tend to focus on specific parts of the whole process (Hickok, 2014). For instance, prior models of speech production have focused on linguistic aspects (Dell et al., 1997; Indefrey & Levelt, 2004; Levelt et al., 1999), while others focused on sensorimotor control (Golfinopoulos et al., 2010; Guenther et al., 2006). An integrated model of speech production often interfaces between the linguistic and sensorimotor control components (Civier et al., 2013; Hickok, 2012). The goal of this pilot study is to identify an electrophysiological neural marker that represents a connection between linguistic and motor processes during speech production. Many integrated models of verbal communication use an intention signal to begin linguistic processing that converts communicative thoughts into linguistic units (Foygel & Dell, 2000; Hickok, 2012; Walker & Hickok, 2016). As the word is selected during linguistic processing, lexical–auditory and lexical–motor targets are generated that work through feedforward and feedbackward loops in the dynamic articulatory motor system for word production (Walker & Hickok, 2016).

In a general model illustrating the components of speech production (see Figure 1), speech intention forms a vital link between overlapping linguistic (Levelt et al., 1999) and motor

[a]Department of Speech-Language-Hearing: Sciences & Disorders, The University of Kansas, Lawrence
[b]Department of Speech-Language-Hearing: Sciences & Disorders, Neuroscience Graduate Program, The University of Kansas, Lawrence
[c]Department of Neurosciences, University of California San Diego

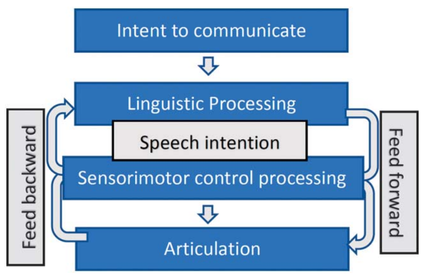
Correspondence to Jonathan S. Brumberg: brumberg@ku.edu

Editor-in-Chief: Cara E. Stepp
Editor: Ayoub Daliri

Received May 21, 2020
Revision received August 6, 2020
Accepted October 19, 2020
https://doi.org/10.1044/2020_JSLHR-20-00277
Publisher Note: This article is part of the Special Issue: Selected Papers From the 2020 Conference on Motor Speech—Basic Science and Clinical Innovation.

Disclosure: *The authors have declared that no competing interests existed at the time of publication.*

Figure 1. A general model of speech production highlighting speech intention.



speech processes (Guenther, 2016; van Der Merwe, 1997) and, if severed, can result in disrupted speech production. In this type of model, message compilation in linguistic planning involves morphosyntactic and phonological planning that is transmitted to the motor speech system for development of the motor program. This feedforward system is accompanied by a feedbackward system that integrates auditory and somatosensory information of the executed motor commands for comparison with intended speech and linguistic targets (e.g., Guenther, 2016). Key to this functionality is a simultaneous process of speech intention that facilitates feedforward communication between the linguistic and speech motor subsystems.

An objective measure of speech intention may help improve our understanding of the relationships between neural processes involved in language and speech production and narrow the effects of neural dysfunction to linguistic only (prior to speech intention), speech only (following speech intention), or the transition between the two processes (the intention process itself). Electroencephalography (EEG) is one objective measure of brain function that can be used to examine this transition stage of speech intention in the midst of the rapidly and simultaneously occurring linguistic and speech motor processes (Beres, 2017). Specifically, the contingent negative variation (CNV), an event-related potential that reflects anticipation of a motor response, is particularly well suited with prior known effects due to speech and language production (Ning et al., 2017; Vanhoutte et al., 2016, 2015; Wu & Thierry, 2017).

The CNV is an EEG slow cortical potential observed during a two-cue motor paradigm where the second (S2), imperative stimulus is contingent on a first (S1), warning stimulus (Walter et al., 1964). The S1 (warning) elicits an early orienting response with a late expectancy wave generated prior to the S2 (imperative; Loveless & Sanford, 1974; Rohrbaugh et al., 1976). The slope or mean amplitude of the late CNV portion occurring just before S2 represents preparation for a motor act like moving your arm (Bareš et al., 2007; Birbaumer et al., 1990; Fan et al., 2007; McCallum, 1988). While the CNV was originally reported to indicate anticipation and preparation of a motor response (Walter et al., 1964), it has been recently observed during expectation and anticipation of a linguistic stimulus (Mnatsakanian & Tarkka, 2002; Tarkka & Basile, 1998) and behavioral performance and perceptual timing (e.g., decision making between auditory signals as "short" or "long"; He & Zempel, 2013). The CNV is also sensitive to anticipation of complex speech movements compared to simple lip stretching and rounding, which further supports its use in measuring speech intention (Wohlert, 1993). The late CNV component was chosen in this pilot study to investigate speech intention as the transitional link between linguistic and sensorimotor control, as suggested in Figure 1, since speech intention may be represented as an anticipation or expectancy of speech motor control (e.g., imperative stimulus) contingent on some speech production task (e.g., warning stimulus). In the current study, the slope of the late CNV component is our primary dependent measure of speech intention in order to capture overall trends in greater negativity associated with prominent late CNV components, rather than average amplitude that may be affected by earlier components of the CNV.

For the current investigation, three different visual presentation protocols were used to elicit the CNV in response to speech intention that varied the amount and type of information available to participants at the warning and imperative stimuli of the classical CNV paradigm. By varying the amount, type, and timing of information in our three different stimulus presentations, the goal was to determine the combination that invoked the greatest CNV negativity (i.e., steepest slope) for a spoken word production task while faithfully reflecting speech intention processing. In each of the protocols, the warning stimulus (S1) was used to indicate the decision to speak, the word to speak, or both, once the imperative stimulus (S2) appeared. As the CNV is maximally negative prior to the imperative stimulus (S2), it is hypothesized that the presence of a CNV prior to speaking (S2) and differences between the Go/NoGo trials are an indication of speech intention and transmission of linguistic commands to the motor speech system. In particular, greater differences in CNV slope between speaking trials versus not speaking was predicted for Protocols 2 and 3, in which the decision to speak is made during the warning stimulus (S1) representing speech and language planning and anticipation of word production, respectively, compared to Protocol 1 where participants did not know if they would have to speak the word or withhold response until the imperative stimulus (S2) appeared. Furthermore, Protocol 2 provided both the instruction to speak and the word to speak at the warning stimulus (S1), which was designed to represent rehearsal and speech intention processing (Ludyga et al., 2018; Vanhoutte et al., 2014), while Protocol 3 presented the instruction to speak (S1) before the word presentation (S2) and was designed to prime the speech intention system only. Protocol 1 was designed similar to previous CNV experiments in which the word was provided before the instruction to speak; this protocol reflects a general motor gating (Wohlert, 1993). A summary of the paradigm differences is shown in Table 1.

Materials and Method

Participants

Eighteen healthy young adults in the age range of 18–36 years (M = 24.6, SD = 4.27) were recruited into

Table 1. Hypothesized linguistic and speech motor processing at the warning (S1) and imperative (S2) stimuli as well as the S1–S2 interval for each protocol in this study.

Protocol	S1 processing	S2 processing	S1–S2 interval processing
Protocol 1	Linguistic	Speech motor + gating	General motor preparation
Protocol 2	Linguistic and speech motor	Gating only	Rehearsal, speech preparation
Protocol 3	Speech motor	Linguistic + gating	Speech preparation/initiation

Note. S1 = Stimulus 1; S2 = Stimulus 2.

three groups of six participants each for the three presentation protocols (nine women, all right-handed). All participants were native speakers of American English, with self-reported normal or corrected vision, speech, language, and hearing and no reported neurological or neuromotor complaints. All participants provided their written informed consent to participate in our study that was approved by the institutional review board of the University of Kansas. Data from one participant were not included in this analysis due to technical recording errors during data collection, leaving data from 17 participants for further analysis.

Presentation Protocols

The three presentation protocols investigated in this study (shown in Figure 2) varied the information presented to participants at each of the two stimulus cues—warning and imperative. In the first presentation protocol, the warning stimulus (S1), included only word information (colored white) for participants who then received an imperative stimulus (S2) to either speak (change to green) or withhold response (change to red) and should elicit responses for motor preparation and gating. The CNV is most commonly elicited in paradigms, such as this, where the participant receives an S1 to orient to the task and then performs the task at S2 (Kowalski et al., 2018; Lasaponara et al., 2019; Neuhaus, 2019; Ning et al., 2017). In the second presentation protocol, the warning stimulus (S1) provided both the word and the task instruction to speak (green-colored) or to withhold response (red-colored) at the imperative stimulus (S2) at which time the word color changed to white. We hypothesize participants practice the target word through mental rehearsal processes as well as engaging speech intention. Protocol 2 is very similar to other CNV paradigms focused on speech and language production where the participant is aware of the exact word (S1) to be produced at a later time (S2; Ludyga et al., 2018; Vanhoutte et al., 2016, 2015, 2014). In the third presentation protocol, we combined the motor initiation and speech-focused qualities of the first two protocols to focus in on speech-specific initiation processing without confounding mental rehearsal. For the third protocol, participants viewed a green- or red-colored circle warning stimulus (S1) that instructed participants to either speak or withhold response for a target word provided at the imperative stimulus (S2) through a change in color to white. By providing only the instruction to speak without content, participants can only prime the speech production and initiation systems without rehearsal. The third presentation protocol follows a novel paradigm to elicit CNV associated with speech motor preparation by defining the task at S1 and then provide the word at S2 (cf. Maxfield et al., 2015; Wu & Thierry, 2017), eliminating linguistic rehearsal focusing on speech motor anticipation and intention. All three protocols recorded the CNV before speech production to eliminate contamination of the recorded EEG due to electromyographical artifacts associated with orofacial muscle contractions during speech production.

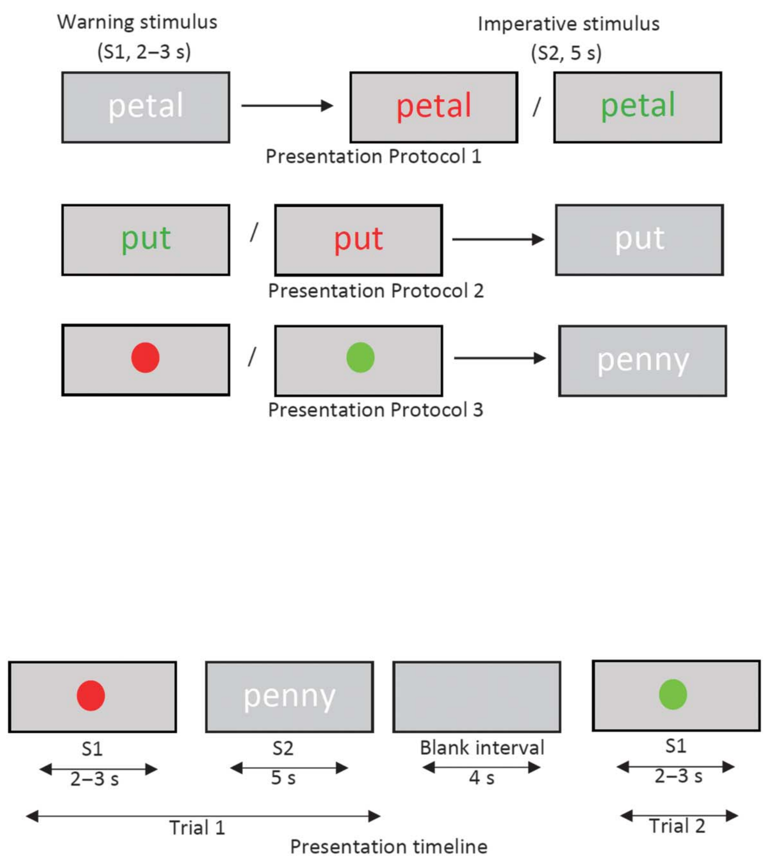
Stimuli

A total of 90 different words were included in this study, each beginning with an initial /p/ to help minimize orofacial artifacts. The words were grouped according to three syllable structures (CVC, CVCV, CVCVC) and three levels (low, medium, high) of word frequency of occurrence in English language as determined by $\mathrm{SUBTLEX}_{\mathrm{US}}$ (Brysbert & New, 2009) and randomly presented through PsychoPy (Peirce et al., 2019). All words were repeated twice, once each for Go and NoGo conditions for a total of 180 trials.

Procedure

EEG recordings took place in a sound-treated booth using 62 active electrodes (g.HIamp, Guger Technologies) arranged uniformly according to the 10–10 standard (Oostenveld & Praamstra, 2001). In the first presentation protocol, the participants were presented with a white-colored word warning stimulus (S1) on the screen, then instructed to speak immediately at the imperative stimulus (S2) if the color of the word changed to green, or withhold their response if the color of the word changed to red. For the second presentation protocol, participants were presented with a green- or red-colored word on the screen at the warning stimulus (S1). If the word color was green, participants were instructed to immediately speak the word aloud when it turned white at the imperative stimulus (S2) and to withhold their response at the imperative stimulus if the initial word color was red. For the third presentation protocol, participants were presented with a green- or red-colored circle in the center of the screen at the warning stimulus (S1). If the circle was green, then participants were to speak the upcoming white-colored word presented at the imperative stimulus (S2) displayed on the screen, and if the

Figure 2. Presentation protocols to elicit the contingent negative variation and a flow of stimulus presentation as seen by the participants. S1 = Stimulus 1; S2 = Stimulus 2.



circle was red, they were to remain silent. In all protocols, the warning stimulus was presented for a random duration between 2 and 3 s, followed by the imperative stimulus (S2), which stayed on the screen for 5 s. Each trial was separated by a 4-s blank screen interval (see Figure 2).

EEG Analysis

EEG analysis was performed in MATLAB, and statistical analyses were performend in R. EEG data were recorded at a sampling rate of 512 Hz referenced to the left earlobe with a ground electrode placed on the forehead just below location AFz. A visual synchronization marker was additionally presented in PsychoPy and recorded simultaneously with the EEG signals by a photodiode to ensure precise alignment of each trial and stimulus (imperative and warning). The resultant signals were bandpass filtered from 0.1 to 30 Hz and down-sampled to 128 Hz, after which ocular artifacts were removed using independent component analysis (Bell & Sejnowski, 1995), where independent components were removed by visual inspection for wave morphology, power spectral density, and spatial concentration around electrodes over the eyes and forehead. The continuous signal was separated into epoch windows of 1.5 s that included the time range from 1.5 s before the imperative stimulus, with baseline correction using the mean amplitude in the 0.5 s prior to the warning stimulus of each trial. Since the trials were separated by a 4-s blank interval, we were able to use a relatively long baseline period (0.5 s) to improve stability of the baseline while avoiding possible influences from the previous trial.

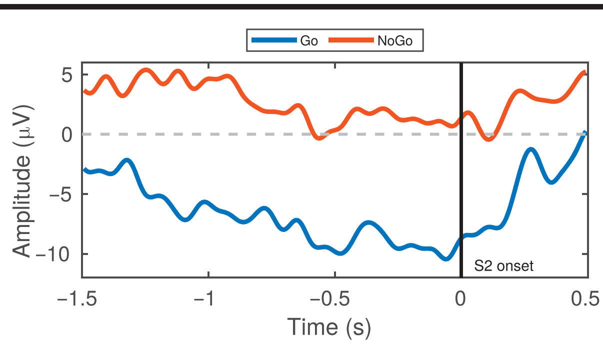
Statistical Analysis

Preprocessed EEG data from 17 participants (n = 6, Protocol 1; n = 5, Protocol 2; n = 5, Protocol 3) were analyzed to first derive a subject average CNV event-related potential separately for Go/NoGo trials for 15 EEG channels with hemisphere locations left (electrode ID 3), right (electrode ID 4), and midline (electrode ID z) and scalp locations FC, F, C, CP, and P according to the 10–10 standard. We then computed slopes from the linear regression of participant average CNV responses in Go and NoGo trials separately in the interval −1.5 to 0 s relative to S2, and their difference was computed as the dependent measure for statistical analysis. The slope was taken as a measure of increasing CNV negativity with differences in slope, indicating differences in processing between Go and NoGo trials. Statistical analysis was performed for the 15 EEG channels per participant grouped according to laterality, left, center (midline), and right. A mixed effects model was used to evaluate the main effects of protocol type and hemispheric laterality and their interaction, with participant ID as a random factor. Statistically significant results at $p < .05$ were separately examined in post hoc comparisons using Tukey's adjustment.

Results

An example of a CNV response for Go and NoGo conditions is shown in Figure 3, where the average of all Go trials is shown in blue and that of NoGo trials is shown in red, which demonstrates amplitude decreasing toward the S2 alignment point (e.g., negative slope). A summary of average CNV slope differences is represented through box plots with mean values (black dots) for each electrode for each subject in Figure 4. Though our protocol included target words of a variety of frequencies and complexities, an initial analysis revealed no statistically significant effects of syllable structure and word frequency on the CNV and is not discussed in the remainder of these results. We found a main effect of hemispheric laterality, $F(2, 232) = 7.30$, $p = .03$, and an interaction effect of Protocol × Hemispheric Laterality, $F(4, 232) = 2.98$, $p = .01$, for CNV slope differences. Though a main effect of protocol was not significant, $F(2, 14) = 2.25$, $p = .38$, post hoc comparisons of slope differences revealed the statistically significant main effect of hemisphere, and the interaction effect was driven by statistically significant differences between right and left hemisphere electrodes (left: 0.139 μV/s, right: −0.632 μV/s), $t(232) = 2.59$, $p = .028$, and between right hemisphere and center electrodes (right: −0.632 μV/s, center: 0.175 μV/s), $t(232) = 2.71$, $p = .020$, in Protocol 3. These values suggest a main result that the CNV differences for Go and NoGo conditions occur most reliably in Protocol 3. Additionally, these results show that right hemisphere electrodes appear to be consistently more negative-going (due to negative slope) than midline and left electrodes, which is corroborated by visual inspection of grand-averaged CNV responses by protocol and laterality (Supplemental Material S1).

Figure 3. Example of the contingent negative variation for Go (blue) and NoGo (red) trials. S2 = Stimulus 2.



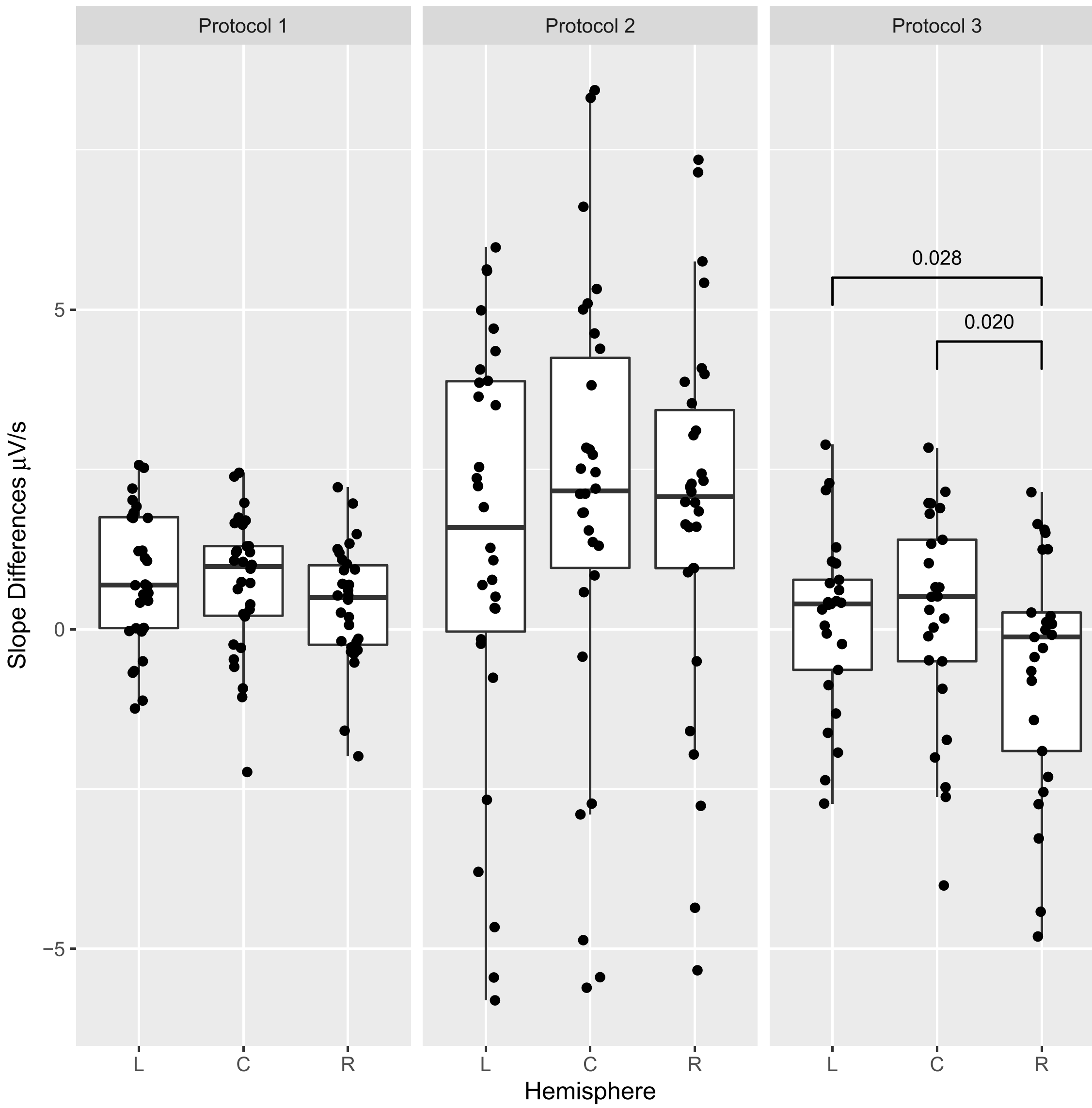
Discussion

This study evaluated speech motor preparatory activity as reflected by the CNV preceding spoken words. Specifically, speech intention was examined by the slope of the late CNV response elicited in a word production task. The analysis focused on determining the factors in each of three presentation protocols that differed in the amount and type of information provided to participants at the warning and imperative stimuli of a classical CNV paradigm for maximizing differences in the late CNV component between Go and NoGo trial conditions. Our main hypothesis was that Protocol 1 would reflect general motor preparation and gating, Protocol 2 would reflect mental rehearsal, and Protocol 3 would reflect priming for intention to speak in healthy individuals.

In this research note, we narrowly focus on investigating the CNV response as a possible objective measure for quantifying speech intention. Our larger goal is to validate this measure with future study of Protocol 3 for use in tracking speech intention through the complex and overlapping processes involved in fluent speech production and speech production in adverse conditions, such as mismatches between feedback and expectations and in populations with difficulty initiating and producing expressive speech and language. The CNV has been used previously to assess motor intention and speech preparation; therefore, it is ideally suited to quantitatively measure intentional processes in speech (Ning et al., 2017; Vanhoutte et al., 2016, 2015; Wu & Thierry, 2017). Thus, the goal of the current pilot study was to identify a CNV paradigm that maximized CNV differences between spoken utterances (Go) and silence (NoGo) and reflect speech intentional processes between the warning and imperative stimulus as a reflection of the link between linguistic and motor speech processes during speech production.

For this pilot study, only Protocol 3 resulted in statistically significant differences between Go and NoGo trials in terms of the average differences between Go/NoGo CNV waveform slopes for hemispheric laterality comparisons between left and right, and center and right electrodes. Protocol 2 provided participants with both the task goal and the target word content at S1, where participants likely used the warning–imperative stimulus interval to rehearse the target word and prime for production at the imperative stimulus possibly confounding observation of the intention signal alone. Since all three protocols involved a speaking task, we may assume that speech preparation is a default process during the S1–S2 interval, but Protocols 1 and 2 are burdened with additional linguistic processing of the word stimulus. The combination of default speech preparation and linguistic processing in Protocol 2 may be why we did not see reliable effects of Go/NoGo slope differences as were observed in Protocol 3, which were counter to our initial hypotheses.

Figure 4. Summary of average contingent negative variation slope differences for each protocol (1–3) and electrode location by hemisphere (L = left; R = right; C = center/midline), with *p* values provided for multiple comparisons testing. The data points represent the individual electrode channels used to define the hemispheric categories of left, right, and center for each subject.



On the other hand, Protocol 3 provided information about task goals (to speak or to remain silent) at the warning stimulus but withheld the content of the word to be spoken until the imperative stimulus. As a result, participants engaged in both linguistic and speech motor processing at S2 and were limited to priming their production systems to process the upcoming stimulus during the warning–imperative stimulus interval. The hemispheric laterality effects seen in Protocol 3 may be attributed to anticipation of the word on the screen or execution of the decision-making process or a general intent to communicate, which are all related to the speech intention process. Finally, Protocol 1 provided the target word without the task goal at the warning stimulus, so participants may have rehearsed the target word without priming/initiating the production system. Based on our criteria, the results of this pilot work suggest using Protocol 3 to isolate neural processes of speech intention for further investigation in healthy young participants and others, including those with difficulty initiating expressive productions as opposed to Protocol 2, which is unable to separate the overlapping functions of linguistic processing and speech intention. Finding an objective neural marker of speech intention will provide a way to quantify speech intention in current models of speech production and track intentional mechanisms through the complex overlapping processes involved in speech production.

Conclusions

The current study examines speech intention as a transition process between lexical processing and speech motor production with distinct neural components that can be quantified using an electrophysiological marker. The findings of our pilot study provide evidence for speech intentional processes through EEG evaluation of three speech production paradigms eliciting a CNV response. Using the CNV paradigm, we were able to manipulate the type and amount of information provided to participants as they prepared to speak and then executed speech motor production. The paradigm that maximized differences in CNV slopes between Go and NoGo trials targeted speech intention processes by providing participants only with information about the task goal and withheld the production stimulus until a later time. This paradigm configuration forced participants to limit their preparatory activities to priming and initiating speech production. These findings for Protocol 3 not only provide a way to quantify and track intentional mechanisms as the link between linguistic and speech processes using EEG in healthy individuals but may also provide an important measure for assessment in individuals with difficulty producing/initiating speech, such as those with nonfluent aphasia who know the word they wish to speak but are unable to produce it. The identification of speech intention in individuals who have had a stroke with resultant speech disruption may be used to investigate whether the disruption is primarily in linguistic processing, speech motor processes, or their link.

Author Contributions

Juhi Kidwai: Investigation (Lead), Writing – original draft (Lead), Writing – review & editing (Equal). **Jonathan S. Brumberg:** Formal analysis (Equal), Investigation (Equal), Software (Lead), Supervision (Lead), Writing – review & editing (Lead). **Brianna M. Marsh:** Investigation (Equal), Methodology (Equal).

Acknowledgments

This research was supported in part by the National Institute on Deafness and Other Communication Disorders (RO1DC016343; PI: Brumberg) and the Initiative for Maximizing Student Development (IMSD NIH R25GM62232). The authors are grateful to Rebekah Schneider for her programming assistance, Chavis Lickvar-Armstrong for reviewing a draft of this research note, and all the participants for their time.

References

Bara, B. G. (2010). *Cognitive pragmatics: The mental process of communication*. MIT Press. https://doi.org/10.7551/mitpress/9780262014113.001.0001

Bareš, M., Nestrašil, I., & Rektor, I. (2007). The effect of response type (motor output versus mental counting) on the intracerebral distribution of the slow cortical potentials in an externally cued (CNV) paradigm. *Brain Research Bulletin, 71*(4), 428–435. https://doi.org/10.1016/j.brainresbull.2006.10.012

Bell, A. J., & Sejnowski, T. J. (1995). An information–maximization approach to blind separation and blind deconvolution. *Neural Computation, 7*(6), 1129–1159. https://doi.org/10.1162/neco.1995.7.6.1129

Beres, A. M. (2017). Time is of the essence: A review of electroencephalography (EEG) and event-related brain potentials (ERPs) in language research. *Applied Psychophysiology and Biofeedback, 42*(4), 247–255. https://doi.org/10.1007/s10484-017-9371-3

Birbaumer, N., Elbert, T., Canavan, A. G., & Rockstroh, B. (1990). Slow potentials of the cerebral cortex and behavior. *Physiological Reviews, 70*(1), 1–41. https://doi.org/10.1152/physrev.1990.70.1.1

Brysbert, M., & New, B. (2009). Moving beyond Kucera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods, 41*(4), 977–990. https://doi.org/10.3758/BRM.41.4.977

Civier, O., Bullock, D., Max, L., & Guenther, F. H. (2013). Computational modeling of stuttering caused by impairments in a basal ganglia thalamo-cortical circuit involved in syllable selection and initiation. *Brain and Language, 126*(3), 263–278. https://doi.org/10.1016/j.bandl.2013.05.016

Dell, G. S., Burger, L. K., & Svec, W. R. (1997). Language production and serial order: A functional analysis and a model. *Psychological Review, 104*(1), 123–147. https://doi.org/10.1037/0033-295X.104.1.123

Fan, J., Kolster, R., Ghajar, J., Suh, M., Knight, R. T., Sarkar, R., & McCandliss, B. D. (2007). Response anticipation and response conflict: An event-related potential and functional magnetic resonance imaging study. *The Journal of Neuroscience, 27*(9), 2272–2282. https://doi.org/10.1523/JNEUROSCI.3470-06.2007

Foygel, D., & Dell, G. S. (2000). Models of impaired lexical access in speech production. *Journal of Memory and Language, 43*(2), 182–216. https://doi.org/10.1006/jmla.2000.2716

Golfinopoulos, E., Tourville, J. A., & Guenther, F. H. (2010). The integration of large-scale neural network modeling and functional brain imaging in speech motor control. *NeuroImage, 52*(3), 862–874. https://doi.org/10.1016/j.neuroimage.2009.10.023

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics: Speech acts* (Vol. 3, pp. 41–58). Academic Press.

Guenther, F. H. (2016). *Neural control of speech*. MIT Press. https://doi.org/10.7551/mitpress/10471.001.0001

Guenther, F. H., Ghosh, S. S., & Tourville, J. A. (2006). Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain and Language, 96*(3), 280–301. https://doi.org/10.1016/j.bandl.2005.06.001

He, B. J., & Zempel, J. M. (2013). Average is optimal: An inverted-U relationship between trial-to-trial brain activity and behavioral performance. *PLOS Computational Biology, 9*(11), e1003348. https://doi.org/10.1371/journal.pcbi.1003348

Hickok, G. (2012). Computational neuroanatomy of speech production. *Nature Reviews Neuroscience, 13*(2), 135–145. https://doi.org/10.1038/nrn3158

Hickok, G. (2014). The architecture of speech production and the role of the phoneme in speech processing. *Language, Cognition and Neuroscience, 29*(1), 2–20. https://doi.org/10.1080/01690965.2013.834370

Indefrey, P., & Levelt, W. J. M. (2004). The spatial and temporal signatures of word production components. *Cognition, 92*(1), 101–144. https://doi.org/10.1016/j.cognition.2002.06.001

Kowalski, J. T., Falkenhagen, F., Radtke, Y., Niederberger, U., Petermann, F., Darabaneanu, S., Steinmann, E., Gerber, W., & Siniatchkin, M. (2018). Contingent negative variation in acute trauma patients: A prospective exploratory study. *Biological*

Psychology, 138, 126–132. https://doi.org/10.1016/j.biopsycho.2018.09.001

Lasaponara, S., Glicksohn, J., Mauro, F., & Ben-Soussan, T. D. (2019). Chapter 4. Contingent negative variation and P3 modulations following mindful movement training. *Progress in Brain Research, 244,* 101–114. https://doi.org/10.1016/bs.pbr.2018.10.017

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*(1), 1–75. https://doi.org/10.1017/S0140525X99001776

Loveless, N. E., & Sanford, A. J. (1974). Effects of age on the contingent negative variation and preparatory set in a reaction-time task. *Journal of Gerontology, 29*(1), 52–63. https://doi.org/10.1093/geronj/29.1.52

Ludyga, S., Herrmann, C., Mücke, M., Andrä, C., Brand, S., Pühse, U., & Gerber, M. (2018). Contingent negative variation and working memory maintenance in adolescents with low and high motor competencies. *Neural Plasticity, 2018,* 9628787. https://doi.org/10.1155/2018/9628787

Maxfield, N., Morris, K., Frisch, S. A., Morphew, K., & Constantine, J. L. (2015). Real-time processing in picture naming in adults who stutter ERP evidence. *Clinical Neurophysiology, 126*(2), 284–296. https://doi.org/10.1016/j.clinph.2014.05.009

McCallum, W. (1988). Potentials related to expectancy, preparation, and motor activity. In T. Picton (Ed.), *Handbook of electroencephalography and clinical neurophysiology: Revised series* (Vol. 3, pp. 427–534). Elsevier.

Mnatsakanian, E., & Tarkka, I. (2002). Task-specific expectation is revealed in scalp-recorded slow potentials. *Brain Topography, 15,* 87–94. https://doi.org/10.1023/A:1021416406992

Neuhaus, C. (2019). "You should praise"–"You should kill": The contingent negative variation indicates moral goodness and badness. *Frontiers in Human Neuroscience, 13,* 432. https://doi.org/10.3389/fnhum.2019.00432

Ning, N., Peng, D., Liu, X., & Yang, S. (2017). Speech timing deficit of stuttering: Evidence from contingent negative variations. *PLOS ONE, 12*(1), e0168836. https://doi.org/10.1371/journal.pone.0168836

Oostenveld, R., & Praamstra, P. (2001). The five percent electrode system for high-resolution EEG and ERP measurements. *Clinical Neurophysiology, 112*(4), 713–719. https://doi.org/10.1016/S1388-2457(00)00527-7

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods, 51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Rohrbaugh, J. W., Syndulko, K., & Lindsley, D. B. (1976). Brain wave components of the contingent negative variation in humans. *Science, New Series, 191*(4231), 1055–1057. https://doi.org/10.1126/science.1251217

Sperber, D., & Wilson, D. (1995). *Relevance: Communication and cognition* (2nd ed.). Blackwell.

Tarkka, I. M., & Basile, L. F. H. (1998). Electric source localization adds evidence for task-specific CNVs. *Behavioural Neurology, 11,* 728634. https://doi.org/10.1155/1998/728634

van Der Merwe, A. (1997). A theoretical framework for the characterization of pathological speech sensorimotor control. In M. R. McNeil (Ed.), *Clinical management of sensorimotor speech disorders* (pp. 1–26). Thieme.

Vanhoutte, S., Cosyns, M., van Mierlo, P., Batens, K., Corthals, P., De Letter, M., Van Borsel, J., & Santens, P. (2016). When will a stuttering moment occur? The determining role of speech motor preparation. *Neuropsychologia, 86,* 93–102. https://doi.org/10.1016/j.neuropsychologia.2016.04.018

Vanhoutte, S., Santens, P., Cosyns, M., van Mierlo, P., Batens, K., Corthals, P., De Letter, M., & Van Borsel, J. (2015). Increased motor preparation activity during fluent single word production in DS: A correlate for stuttering frequency and severity. *Neuropsychologia, 75,* 1–10. https://doi.org/10.1016/j.neuropsychologia.2015.05.013

Vanhoutte, S., Van Borsel, J., Cosyns, M., Batens, K., van Mierlo, P., Hemelsoet, D., Van Roost, D., Corthals, P., De Letter, M., & Santens, P. (2014). CNV amplitude as a neural correlate for stuttering frequency: A case report of acquired stuttering. *Neuropsychologia, 64,* 349–359. https://doi.org/10.1016/j.neuropsychologia.2014.09.036

Walker, G. M., & Hickok, G. (2016). Bridging computational approaches to speech production: The semantic–lexical–auditory–motor model (SLAM). *Psychonomic Bulletin & Review, 23*(2), 339–352. https://doi.org/10.3758/s13423-015-0903-7

Walter, W. G., Cooper, R., Aldridge, V. J., McCallum, W. C., & Winter, A. L. (1964). Contingent negative variation: An electric sign of sensorimotor association and expectancy in the human brain. *Nature, 203,* 380–384. https://doi.org/10.1038/203380a0

Wohlert, A. B. (1993). Event-related brain potentials preceding speech and nonspeech oral movements of varying complexity. *Journal of Speech and Hearing Research, 36*(5), 897–905. https://doi.org/10.1044/jshr.3605.897

Wu, Y. J., & Thierry, G. (2017). Brain potentials predict language selection before speech onset in bilinguals. *Brain and Language, 171,* 23–30. https://doi.org/10.1016/j.bandl.2017.04.002