JNeurosci
THE JOURNAL OF NEUROSCIENCE

Research Articles: Behavioral/Cognitive

SUBTHALAMIC NUCLEUS NEURONS DIFFERENTIALLY ENCODE EARLY AND LATE ASPECTS OF SPEECH PRODUCTION

WJ Lipski[1], A Alhourani[1], T Pirnia[1], PW Jones[1], C Dastolfo-Hromack[1], LB Helou[2], DJ Crammond[1], S Shaiman[3], MW Dickey[3], LI Holt[4], RS Turner[2,6], JA Fiez[5,6] and RM Richardson[1,2,6]

[1]*Brain Modulation Lab, Department of Neurological Surgery, University of Pittsburgh School of Medicine*
[2]*Department of Neurobiology, University of Pittsburgh School of Medicine*
[3]*Department of Department of Communication Science and Disorders, University of Pittsburgh*
[4]*Department of Psychology, Carnegie Mellon University*
[5]*Department of Psychology, University of Pittsburgh*
[6]*University of Pittsburgh Brain Institute.*

DOI: 10.1523/JNEUROSCI.3480-17.2018

Received: 8 December 2017

Revised: 9 May 2018

Accepted: 15 May 2018

Published: 22 May 2018

Author contributions: W.J.L., L.B.H., J.A.F., and R.M.R. designed research; W.J.L., A.A., D.J.C., and R.M.R. performed research; W.J.L., A.A., T.P., P.J., C.D.-H., S.S., M.D., L.H., J.A.F., R.S.T., and R.M.R. analyzed data; W.J.L., C.D.-H., D.J.C., S.S., M.D., L.H., J.A.F., R.S.T., and R.M.R. wrote the paper; L.B.H. edited the paper.

Conflict of Interest: The authors declare no competing financial interests.

Funding was provided by NINDS U01NS098969 (PI: Richardson), the Hamot Health Foundation (PI: Richardson), and a University of Pittsburgh Brain Institute NeuroDiscovery Pilot Research Award (PI: Richardson). The authors are grateful for the expert intraoperative clinical assistance of Danielle Corson and Jim Sweat.

Correspondence should be addressed to To whom correspondence should be addressed: Mark Richardson, MD, PhD, richardsonrm@upmc.edu, Department of Neurological Surgery, University of Pittsburgh Medical Center, 200 Lothrop St., Suite B400, Pittsburgh, PA 15213

Cite as: J. Neurosci ; 10.1523/JNEUROSCI.3480-17.2018

SUBTHALAMIC NUCLEUS NEURONS DIFFERENTIALLY ENCODE EARLY AND LATE ASPECTS OF SPEECH PRODUCTION

Lipski WJ[1], Alhourani A[1], Pirnia T[1], Jones PW[1], Dastolfo-Hromack C[1], Helou LB[2], Crammond DJ[1], Shaiman S[3], Dickey MW[3], Holt LL[4], Turner RS[2,6], Fiez JA[5,6], Richardson RM[1,2,6]*

[1]Brain Modulation Lab, Department of Neurological Surgery, University of Pittsburgh School of Medicine; [2]Department of Neurobiology, University of Pittsburgh School of Medicine; [3]Department of Department of Communication Science and Disorders, University of Pittsburgh; [4]Department of Psychology, Carnegie Mellon University; [5]Department of Psychology, University of Pittsburgh; [6]University of Pittsburgh Brain Institute.

*To whom correspondence should be addressed:

Mark Richardson, MD, PhD
richardsonrm@upmc.edu
Department of Neurological Surgery
University of Pittsburgh Medical Center
200 Lothrop St., Suite B400
Pittsburgh, PA 15213

Acknowledgements

Funding was provided by NINDS U01NS098969 (PI: Richardson), the Hamot Health Foundation (PI: Richardson), and a University of Pittsburgh Brain Institute NeuroDiscovery Pilot Research Award (PI: Richardson). The authors are grateful for the expert intraoperative clinical assistance of Danielle Corson and Jim Sweat.

ABSTRACT

Basal ganglia-thalamocortical loops mediate all motor behavior, yet little detail is known about the role of basal ganglia nuclei in speech production. Using intracranial recording during deep brain stimulation surgery in humans with Parkinson's disease, we tested the hypothesis that the firing rate of subthalamic nucleus neurons is modulated in sync with motor execution aspects of speech. Nearly half of seventy-nine unit recordings exhibited firing rate modulation, during a syllable reading task across twelve subjects (male and female). Trial-to-trial timing of changes in subthalamic neuronal activity, relative to cue onset versus production onset, revealed that locking to cue presentation was associated more with units that decreased firing rate, while locking to speech onset was associated more with units that increased firing rate. These unique data indicate that subthalamic activity is dynamic during the production of speech, reflecting temporally-dependent inhibition and excitation of separate populations of subthalamic neurons.

SIGNIFICANCE STATEMENT

The basal ganglia are widely assumed to participate in speech production, yet no prior studies have reported detailed examination of speech-related activity in basal ganglia nuclei. Using microelectrode recordings from the subthalamic nucleus during a single syllable reading task, in awake humans undergoing deep brain stimulation implantation surgery, we show that the firing rate of subthalamic nucleus neurons is modulated in response to motor execution aspects of speech. These results are the first to establish a role for subthalamic nucleus neurons in encoding of aspects of speech production, and they lay the groundwork for launching a modern subfield to explore basal ganglia function in human speech.

INTRODUCTION

Producing speech is the most complex of human motor behaviors, requiring dynamic interaction between multiple brain regions. The segregated loop organization of basal ganglia-thalamocortical circuits suggests that the basal ganglia, including the subthalamic nucleus (STN), play a critical role in speech production. This concept is supported additionally by observations that impairments in speech production are common features of basal ganglia-associated degenerative disorders including Parkinson's disease, and that other disorders in speech production (e.g., stuttering) are associated with abnormalities in basal ganglia activity (Alm, 2004; Giraud et al., 2008; Toyomura et al., 2015). Additionally, an extensive body of work in song birds implicate bird-homologues of the basal ganglia (Doupe and Kuhl, 1999), including a homologue of the STN, in the learning and production of vocalizations (Jiao et al., 2000). Many prominent models of speech production nonetheless virtually ignore the basal ganglia (Hickok, 2012), as few studies have examined speech-related neural activity in these subcortical nuclei directly (Ziegler and Ackermann, 2017).

Electrophysiological recordings obtained during the implantation of leads for deep brain stimulation (DBS) represent the only clinically-indicated opportunity to measure neural activity directly from the basal ganglia in awake, behaving human subjects. Previous reports of STN unit activity, however, have been limited to only a single preliminary, qualitative analysis of speech production-related changes in STN firing rates (Watson and Montgomery, 2006). Thus, recording from STN neurons during speech production is a unique opportunity to test hypotheses about the role of this region in the control of complex motor function, where the basal ganglia have alternately been hypothesized to participate in action selection, movement gain and motor learning (Desmurget and Turner, 2010).

To begin to define the role of the STN in speech production more clearly, we established an intraoperative protocol for microelectrode recording during a task that required subjects to read aloud single syllables displayed on a computer screen. We then examined trial-to-trial timing of changes in STN unit activity relative to either the visual presentation of single syllables or to the onset of speech production. Such time-locking is considered as evidence for an underlying functional linkage between the behavioral event and the linked

neural discharge (Seal and Commenges, 1985; Anderson and Turner, 1991; DiCarlo and Maunsell, 2005). Our results suggest that aspects of speech production are encoded in the STN through the inhibition and excitation of functionally segregated neurons.

METHODS

Subjects. Subjects were 12 movement disorders patients (10 male) undergoing awake DBS surgery for Parkinson's disease. Unified Parkinson's disease rating scale (UPDRS) testing was administered by a neurologist within four months before DBS surgery. 10 of 12 subjects underwent bilateral DBS implantation (left lead inserted first), while two underwent unilateral implantation (one left). All subjects underwent overnight withdrawal from their dopaminergic medication prior to surgery. All participants provided written, informed consent in accordance with a protocol approved by the Institutional Review Board of the University of Pittsburgh (IRB Protocol # PRO13110420). In our practice, lead implantation is undertaken using a Leksell frame, with the patient in a semi-sitting position, and occurs first on the left side (for bilateral cases). In order to minimize strain on patients, these subjects were not offered research participation on the second (right brain) side. One subject underwent unilateral right-sided implantation.

Electrophysiological recordings. Unit recordings were carried out using the Neuro-Omega recording system and Parylene insulated, microphonics-free tungsten microelectrodes (Alpha Omega, Nazareth, Israel). Microelectrode impedances ranged from 200-600 kΩ. Targeting of the dorsolateral STN and microelectrode recording (MER) were performed using a standard combination of indirect (starting AC-PC coordinates of x = ±12, y = -3, z = -4) and direct (visualization of the STN in the z=-4 plane of a T2-weighted scan obtained on a 3-Tesla MRI scanner) targeting (Starr et al., 2002). For each subject, two to three simultaneous microelectrode recording passes were made, starting at 15 mm above the surgical target with manual advance of the microdrive in 0.1mm steps, using a center, and posterior and/or medial trajectories, with center-to-center spacing of 2mm in a standard cross-shaped Ben-Gun array. Microelectrode signals were band-pass filtered at 0.075 Hz to 10 kHz and digitized at 44 kHz (NeuroOmega, Alpha Omega, Nazareth, Israel).

Speech Task. The speech task was performed during pauses in the microelectrode recording portion of DBS lead implantation in which stable units were detected. Visual stimuli were created using Matlab software (MathWorks, Natick, MA) and Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997; Kleiner et al, 2007). Subjects were asked to read, in a normal manner, a consonant-vowel-consonant (CVC) syllable presented in white text on an otherwise dark computer screen. Each trial was initiated manually by the experimenter, beginning with presentation of a green fixation cross at the center of the screen (0-250 ms), followed by a variable-time delay (500-1000 ms) during which the screen remained dark. At the end of the delay, text denoting a unique CVC syllable appeared on the screen and remained visible until the subject completed their naming response. A white fixation cross was displayed on the screen during the inter-trial interval (ITI; Figure 1A). Subjects were instructed to respond as soon as the word cue appeared. The CVC stimuli were drawn from prior behavioral work (Moore 2012), and were matched along a number of dimensions, including phoneme recurrence, number of letters, phonological neighborhood density, orthographic neighborhood, and mean bigram frequency. Stimulus lists contained an equal portion of CVC words and non-words, and were composed of consonants drawn from a set of 7 early- or 7 late-developing consonant phonemes.

Audio recordings. Speech output was recorded using an omnidirectional microphone (8 subjects: Audio-Technica, Stow, OH; model ATR3350iS, frequency response 50-18,000 Hz; 4 subjects: Preosonus, Baton Rouge, LA, model PRM1 Precision Flat Frequency Mic, frequency response 20-20,000 Hz) oriented at an angle of approximately 45 degrees and a distance of approximately 8 cm to the subject's mouth. In the four cases where the Preosonus PRM1 microphone was used, a Zoom H6 digital recorder was used to digitize the audio at 96 kHz. In all cases, the audio signal was split out to a Grapevine Neural Interface Processor, where it was digitized at 30 kHz. The audio signal was synchronized with the neural recordings and with visual cue events using digital pulses delivered via a USB data acquisition unit (Measurement Computing, Norton, MA, model USB-1208FS).

Task performance. The audio signal was segmented into trials and responses were coded by a speech-language pathologist using a custom-designed graphical user interface implemented in MATLAB. The response epoch for each trial was defined to start at cue presentation and end at the start of the ITI. The audio signal within each response epoch was coded as follows: (1) production onset was identified, (2) production offset was identified, (3) the phonetic content was identified. Only trials that met the following criteria were included for further analyses: (1) the subject's entire response could clearly be identified within the response epoch; (2) the time from cue presentation to production onset (production latency) was less than the mean production latency (1.2 s) plus 3 standard deviations (0.93 s) for all subjects (threshold = 4.0 s); (3) the duration of the response was less than the mean production latency (0.60 s) plus 3 standard deviations (0.20 s) for all subjects (threshold = 1.19 s); (4) the subject's response was a CVC or CV syllable and was composed of phonemes within the target set or the included mismatch set. Of 2,200 total trials, 150 (6.8%) were rejected from further analysis on the basis of these response criteria. In 11 of the rejected trials, no response was recorded. In 608 trials (139 of which were rejected), the response did not match the target.

Spike sorting. Microelectrode recording data were imported into off-line spike-sorting software (Plexon, Dallas, TX). A 4-pole Butterworth high-pass filter with a cutoff frequency of 200 Hz was applied to the microelectrode recording signal and waveforms were detected by setting a negative threshold at an amplitude equal to approximately 3 times the standard deviation of the voltage signal; single- and multi-unit action potentials were then discriminated using principal components analysis. The results were graded according to the quality and stability of the spike sorting over the duration of the recording. An assignment of "A-sort" was given only to spike clusters that could be discriminated from background activity throughout the duration of a recording, and whose spikes were not strongly modulated by cardiac rate (see Figure 1C-F). A-sorts were further subdivided into single- and multi-unit subcategories. A cluster qualified as a single unit (SU) if: (1) the principal component cluster was clearly separated from other clusters associated with background activity and other units, (2) contained spike waveforms with a unimodal distribution in principal component space, and (3) displayed a refractory period of at least 3ms in its inter-spike interval distribution (Starr et al., 2003; Schrock et al., 2009). For some SU recordings, the location of the principal component cluster drifted gradually during the period of

the recording, likely due to a shift of the brain relative to the electrode. Other A-sorts were classified as multi-unit (MU) recordings because the principle components cluster appeared to include waveforms from multiple units, forming multimodal principal component distributions that could not be clearly separated on short time scales, or that failed to obey the 3 ms refractory period in their inter-spike interval distribution. An assignment of "B-sort" was given to spike clusters that violated the above criteria due to presence of a non-uniform or rapid (5 second time scale) shift of the waveform cluster in principal component space, or due to incomplete separation of the spike cluster from the cluster associated with background noise.

STN unit baseline activity. Baseline spike rates were estimated by averaging across trials the spike rates during the baseline epoch, defined as the 1 s portion of the ITI preceding cue presentation. Because the firing rate of MU recordings depends on the number of neurons contributing to the spike population and thus is difficult to interpret, we calculated baseline firing rates only for SU recordings.

STN unit activity during speech. We used two different estimates of unit activity to test for task-related changes in neuronal spike rate. We tested for task-related increases using a spike density function (SDF), which is a direct representation of a unit's mean instantaneous firing rate. We tested for task-related decreases using a function that reflects a unit's mean inter-spike interval (ISI), which scales with the reciprocal of instantaneous spike rate. This approach was chosen to avoid potential under-sensitivity for the detection of decreases in firing in SDFs due to floor effects (Alexander and Crutcher, 1990a). To construct an SDF function, spike time stamps were rounded to 1 ms. The resulting time series was then convolved with a Gaussian kernel (σ = 25 ms). The inter-spike interval (ISI) time series was computed from the 1 kHz binned time stamp time series by taking the value of the current ISI at each millisecond time point:

$$ISI(t) = ts_{i+1} - ts_i, \qquad \text{(Equation 1)}$$

for t between ts_i and ts_{i+1}, where ts is the set of consecutive time stamps for that spike population. Across-trial means of the SDF and ISI functions were constructed aligned on two epochs of interest: (1) from cue presentation to 0.5 s after the mean production onset for that session (aligned on cue presentation, termed the cue epoch), and (2) from the mean time of cue presentation to 0.5 s after production onset (aligned on production onset, termed the production epoch). A baseline period for each trial was defined as the 1 s portion of the ITI preceding cue presentation, and the trial-wise mean SDF and ISI functions during this epoch served

as baselines against which the test epochs were compared. Baseline firing rates for each unit were defined as the mean of discharge rate during the baseline period across trials.

A unit was considered to have significantly elevated firing during a given epoch if the mean spike density within that test epoch exceeded a threshold level for at least 100ms. The threshold was defined as the upper 5% of a normal distribution with a mean and σ of the baseline mean SDF, Bonferroni corrected for multiple comparisons (where the number of independent observations was considered to be the duration of the epoch of interest divided by the width of the Gaussian kernel, 50ms). Similarly, a unit was considered to have significantly reduced firing within a given epoch if the mean ISI time series exceeded a threshold ISI value for at least 100ms. The threshold ISI value was defined as the upper 5% tail of a normal distribution with a mean and σ of the baseline mean ISI time series, Bonferroni corrected for multiple comparisons (where the number of independent observations was the mean number of ISIs within the epoch of interest).

Speech onset- and cue- locking. For all units with significant changes in mean firing, we sought to determine whether the timing of these responses was more closely locked to the presentation of the cue or to the onset of the production, by examining the trial-to-trial relationship between RTs and neuronal response onsets. First, response onsets were estimated for individual trials. The trial-to-trial timing of an increase in firing was estimated by searching for bursts. The spiking pattern within each trial (after cue presentation) was examined to find a sequence of at least 3 spikes with the highest Poisson Surprise (PS) Burst index. For a given sequence of n spikes within time interval *T*, the PS Burst index was based on the probability of encountering n or more spikes within time interval *T*, given a Poisson-distributed spike train with a discharge rate *r*:

$$PS_{burst} = -\log_{10}\left(e^{-rT}\sum_{i=n}^{\infty}\frac{(rT)^i}{i!}\right). \qquad \text{(Equation 2)}$$

Similarly, the trial-to-trial timing of a decrease in firing was estimated by searching for pauses. The PS Pause index was based on the probability of encountering n or fewer spikes within time interval *T*, in a Poisson-distributed spike train:

$$PS_{pause} = -\log_{10}\left(e^{-rT}\sum_{i=0}^{n}\frac{(rT)^i}{i!}\right). \qquad \text{(Equation 3)}$$

For both increase and decrease indices, *r* was estimated separately for each trial as the discharge rate across the entire trial, and T was the duration of the trial. Only trials with burst or pause sequences whose PS indices exceeded those found in that trial's baseline epoch were considered for further analysis. For each trial, the onset time of the PS Burst (for units with significant excitatory responses) or PS Pause (for units with significant inhibitory responses) spike sequences was defined as the neuronal response increase or decrease onset, respectively.

Next, two intervals were correlated (Spearman rank correlation, MATLAB function corr) with the production latency across trials for each unit: 1. the interval between cue presentation and the neuronal response onset (neuronal response latency), and 2. the interval between the neuronal response onset and production onset (neuronal response to production interval). A unit's response was considered to be temporally-locked to: 1. cue onset, if a significant change in activity during the cue epoch was observed, and the corresponding neuronal response to production interval was correlated ($p<0.05$) with production latency (Figure 2A-B), or 2. the onset of speech, if significant change in activity in the production epoch was observed, and the corresponding neuronal response latency was correlated ($p<0.05$) with production latency (Figure 2C-D). If both correlations were significant, then the unit's response was considered to be both cue- and production -locked, i.e. its activity was temporally associated with both events.

Analysis of speech volume. Relative speech volume was computed based on the audio recording corresponding to the subject's response (speech) and the audio corresponding to the baseline epoch (baseline). The ratio of the speech to baseline root-mean-square (RMS) amplitudes was represented as a decibel statistic for each trial:

$$dB(SNR) = 20 \times \log_{10}\left(\frac{RMS_{speech}}{RMS_{baseline}}\right) \quad \text{(Equation 4)}$$

For all units with a significant speech-related modulation of firing, the relative speech volume was then correlated (Spearman rank correlation, MATLAB function corr) across trials with the mean firing rate during speech, i.e. between speech onset and speech offset for each trial. Because the timing of the firing rate modulation varied between units and between trials, an additional analysis was carried out to examine the

correlation between relative speech volume and the mean burst firing rate (for increase-type responses) or mean pause firing rate (for decrease-type responses; see Methods: Speech onset- and cue- locking). For each type of firing rate measure, the firing rate was z-scored against the baseline firing rate (within each trial) prior to computing the correlation.

Anatomical localization of recordings. Anatomical locations of microelectrode recordings were expressed in terms of the microelectrode recording-defined STN boundaries along each electrode trajectory. Thus, each microelectrode recording location was identified by its relative position within the Ben-Gun orientation (central, posterior or medial) and the percent depth through the STN within that trajectory (with 0% representing the ventral STN boundary and 100% representing the dorsal STN boundary). In addition, electrode localization was carried using the Lead-DBS toolbox (Horn and Kühn, 2015). Preoperative and postoperative magnetic resonance images were co-registered and normalized to Montreal Neurological Institute (MNI) space. MNI locations of DBS lead placements were determined from post-operative images, and intraoperative microelectrode locations were calculated based on their position relative to final lead placement. In order to test whether unit responses recorded within the STN were anatomically segregated according to their speech-related response types and locking types, linear discriminant analysis was used to classify units based on MNI coordinates (MATLAB function fitcdiscr). 10-fold cross validation was used to estimate classification accuracy.

Analysis of spike isolation and stability. In order to quantify the sort quality of STN units, two different measures were adapted from a method by Joshua and colleagues: signal-to-noise ratio (SNR) and isolation score (IS) (Joshua et al. 2007). Signal-to-noise was defined as

$$SNR = \frac{peak\text{-}to\text{-}peak}{Noise}\ , \qquad \text{(Equation 5)}$$

where $peak\text{-}to\text{-}peak$ indicates the signal amplitude, or difference between the minimum and maximum of the average spike waveform, and the $Noise$ is the standard deviation of the concatenated residuals (spike waveforms minus average spike waveform) (Joshua et al. 2007). Isolation score is an estimate of the probability that a given individual spike waveform (typically 66 samples, e.g. 1.5 ms, long) belongs to the assigned spike cluster rather than the noise cluster (Joshua et al. 2007). Clusters for each candidate single

unit and for noise (all other waveforms from the same recording) were defined in the first two dimensions of a principal components analysis (Plexon Offline Sorter). Our measure of the similarity of waveforms within a cluster was based on the Euclidean distance, $d(X,Y)$, between raw waveforms X and Y, both from the same cluster:

$$\text{Similarity}(\text{X},\text{Y}) = \exp(-d(X,Y)(\lambda/d_0)) \text{ ,} \quad \text{(Equation 6)}$$

normalized according to the average Euclidean distance between spikes in the spike cluster, d_0, and a gain constant, λ (equal to 10 (Joshua et al. 2007)). That similarity index was then normalized according to the mean similarity between within-cluster waveforms X and all other waveforms Z (e.g., waveforms from other spikes and noise):

$$P_X(Y) = \frac{\exp(-d(X,Y)(\lambda/d_0))}{\sum_{Z\neq X}\exp(-d(X,Z)(\lambda/d_0))} \text{ ,} \quad \text{(Equation 7)}$$

Importantly, in order to consistently characterize this quantity across units, we chose to modify the method by Joshua and colleagues by selecting an equal number of waveforms in the spike and noise clusters for each unit whenever possible. Thus, if a sort resulted in a greater number of noise waveforms then spike waveforms, the noise cluster was estimated by randomly subsampling noise waveforms to match the number of spike waveforms (random subsampling was performed using MATLAB function randperm, using a uniform distribution). If, on the other hand, the number of spike waveforms (N_{spike}) was greater then the number of noise waveforms (N_{noise}), the normalization term in the similarity index was adjusted to weight spike and noise waveforms equally:

$$P_X(Y) = \frac{\exp(-d(X,Y)(\lambda/d_0))}{\sum_{Z\in Spike, Z\neq X}\exp\left(-d(X,Z)\left(\frac{\lambda}{d_0}\right)\right)+\frac{N_{spike}}{N_{noise}}\sum_{Z\in Noise, Z\neq X}\exp\left(-d(X,Z)\left(\frac{\lambda}{d_0}\right)\right)} \text{ .} \quad \text{(Equation 8)}$$

Summing the similarity index over all waveforms in the spike cluster results in a measure of how close waveform X is to the spike cluster compared to the noise cluster:

$$P(X) = \sum_{Y\in Spike} P_X(Y) \text{ .} \quad \text{(Equation 9)}$$

Equal weighting of the normalization term in equations 7 and 8 thus ensures that a P(X) value of 0.5 indicates that waveform X is equidistant from the spike and noise clusters. Finally, the isolation score is computed by taking the mean value of the above measure in the spike cluster:

$$IS = \frac{1}{N_{spike}}\sum_{X\in Spike} P(X) \text{ .} \quad \text{(Equation 10)}$$

Signal-to-noise ratios and isolation scores were computed for all single- and multi- units, including all spikes during speech task performance. In order to further assess spike stability during speech, these measures were then calculated separately for spikes recorded during the baseline epoch and speech epoch (1 s following production onset). For each unit, baseline and speech epoch spikes were pooled across trials. We then used a permutation testing procedure to determine whether the difference between baseline and speech measures of signal-to-noise ratio and isolation score was greater then expected by chance. In order to determine the null distribution of the test statistic – the difference between baseline and speech measures of isolation – we generated 1000 surrogate statistics by randomly selecting "baseline" waveforms and "speech" waveforms from all waveforms detected during the baseline and speech epochs.

To assess the degree to which quantitative measures of isolation predicted unit type (single- vs. multi- unit) and unit sort quality (A vs. B sort) linear discriminant analysis was used to classify units based on signal-to-noise ratio and isolation score measures (MATLAB function fitcdiscr). 10-fold cross validation was used to estimate classification accuracy.

RESULTS

Subject demographics are summarized in Tables 1 and 2. Twelve subjects each performed between 1 and 4 blocks of 60 trials during single unit recording sessions (median 2.5 blocks, mean 160 trials). An average 6.5 ± 1.9% of trials were excluded from analysis due to incorrect responses. Across subjects, the mean latency to the onset of a production was 1.10 ± 0.31s, and the mean duration of speech was 0.605 ± 0.175s. A subject's mean production latency correlated significantly with the subject's speech UPDRS sub-score (Spearman rho = 0.72, p = 0.02). This correlation failed to reach significance for speech duration (Spearman rho = -0.09, p = 0.8) or the fraction of trials with incorrect responses (Spearman rho = -0.62, p = 0.06). The subjects' total UPDRS score was not correlated with any of these task measures (production latency: Spearman rho = 0.22, p = 0.5; speech duration: Spearman rho = -0.23, p = 0.5; percent correct: Spearman rho = 0.22, p = 0.5).

A total of 45 neuronal recordings met the criteria for A-sorts (22 single-unit, 23 multi-unit recordings). Thirty-four additional recordings met criteria for B-sorts (3 single-unit, 31 multi-unit recordings). The mean baseline

firing rates were not significantly different between A- and B-sort single units (21.8 ± 3.2 spikes/s vs. 27.3 ± 7.1 spikes/s; mean ± standard error, p = 0.55, unpaired t-test), and were consistent with data reported previously from the human STN (Rodriguez-Oroz et al., 2001; Abosch et al., 2002; Starr et al., 2003; Theodosopoulos et al., 2003; Romanelli et al., 2004; Schrock et al., 2009).

Spike sort quality was quantified for all units using signal-to-noise ratio and isolation score measures. Isolation scores were significantly different between single- and multi- units (single-unit median = 0.97, inter-quartile range (IQR) = 0.06; multi-unit median = 0.86 IQR = 0.15; $p = 8.6 \times 10^{-9}$, Wilcoxon rank sum test), and between A and B sorts (A sort median = 0.93, IQR = 0.11; B sort median = 0.86 IQR = 0.19; $p = 3.1 \times 10^{-5}$, Wilcoxon rank sum test). Similarly, signal-to-noise ratios were significantly different between single- and multi- units (single-unit median = 9.8, IQR = 2.2; multi-unit median = 5.4 IQR = 1.5; $p = 6.3 \times 10^{-11}$, Wilcoxon rank sum test), and between A and B sorts (A sort median = 7.7, IQR = 4.3; B sort median = 5.3 IQR = 1.7; $p = 8.5 \times 10^{-5}$, Wilcoxon rank sum test). Based on these two measures, a linear discriminant analysis classifier could distinguish between single- and multi- units with 85.0 ± 0.5% accuracy (significantly greater then chance, 66.6%, $p = 6.6 \times 10^{-16}$, unpaired t-test), and between A and B sorts with 67.2 ± 4.4% accuracy (significantly greater then chance, 54.3%, $p = 6.5 \times 10^{-6}$, unpaired t-test).

Overall, a high percentage of units demonstrated a speech-related change in firing. 22 units exhibited significant increases in firing rate, 13 units showed significant decreases, and 7 units showed a mixed increase/decrease response during the production epoch. The proportion of units exhibiting these speech-related changes did not depend on sort quality (A- or B-sorts) or on unit type (single or multi unit; Table 3). Figure 3A-C shows examples of these unit response categories. While there was an overall significant difference in the proportions of neurons in the four response categories (increase, decrease, mixed, and non-response, $\chi^2 = 25.8608$, $p = 1.02 \times 10^{-5}$), there was no significant difference between the proportion of increase-type and decrease-type units ($\chi^2 = 2.3$ $p = 0.13$). The prevalence of speech-responsive units did not relate to the subjects' symptom severity, and the proportion of units recorded from each subject that showed increase, decrease, cue-locking or speech-locking response types was not correlated with the speech sub-

score or total UPDRS score (Table 4). Increase- and decrease-type single-units were not differentiated statistically by baseline firing rates (increase-type firing rate = 20.6 ± 6.4 spikes/s; decrease-type firing rate = 34.4 ± 10.5 spikes/s; p = 0.27, unpaired t-test). The mean latency of neuronal responses (defined as onset of the first significant change relative to production onset, see Figure 3D-F) also was similar between increase and decrease response types (-0.23 ± 0.07s and -0.20 ± 0.14s, respectively, p = 0.87, unpaired t-test). In the one participant with right STN recordings (2 multi-units), no speech-related responses were found.

Response types were observed to be differentially associated with speech onset- and cue- locking. Among 29 units with significant increases in firing rate during the production epoch, the responses were preferentially time-locked to production (41%), with a minority time-locked to cue onset (7%) or to both cue and production onset (7%) (Figure 4). In contrast, among 20 units with significant decreases in firing rate, 40% were time-locked to cue onset, while only 15% had responses time-locked to production onset, and no responses were time-locked to both cue and production onset (Figure 5). Again, the proportion of responses showing speech- versus cue- locking firing changes did not depend on sort quality or on unit type (Table 5). A Chi-square test was used to verify that increase-type neural responses were more likely to be time-locked to the production onset than were decreases (χ^2 = 3.89, p = 0.049), whereas decrease-type responses were more likely to be time-locked to cue onset (χ^2 = 7.99, p = 0.0047; Table 6). The mean latency of neuronal responses (defined as the mean neuronal response latency across trials) was shorter for cue-locked responses (0.76 ± 0.12 s) than for speech-locked responses (1.08 ± 0.08 s; p = 0.039, unpaired t-test). The mean neuronal response to production interval (defined as the mean neuronal response to production interval across trials) was also greater in magnitude for cue-locked responses (-0.48 ± 0.12 s) than for speech-locked responses (-0.15 ± 0.6 s; p = 0.011, unpaired t-test).

Encoding of speech duration was not prevalent in recorded STN units. The duration of the neural response had a significant correlation with the duration of speech production (Spearman correlation, p < 0.05) in only 2 of 29 units with increase-type responses, and in only 1 of 20 neurons with decrease-type responses. These proportions were not significantly different from zero (Fisher's exact test, p=0.49 for increase-type responses,

p=1.0 for decrease-type responses), indicating that they are too small to be estimated statistically from this experiment.

Evidence for encoding the volume of speech was found in a small number of decrease-type STN units. When firing rate during speech was examined for each trial, none of the 29 increase-type responses and 1 of 23 decrease-type responses showed a significant correlation (rho = -0.27, p = 0.04) with relative speech volume. Similarly, when mean burst or pause firing rate was examined, none of the 29 increase-type responses and 2 of 23 decrease-type responses (2 subjects) showed a significant correlation (rho = -0.42, -0.30; p = 0.020, 0.025, respectively) with relative speech volume across trials.

We did not find evidence for topographical organization of response types. Unit recording locations were analyzed based on the recording trajectory (center, 23 units, average span 4.7 ± 0.5 mm; posterior, 29 units, average span 5.2 ± 0.5 mm; or medial, 27 units, average span 5.6 ± 0.6 mm), and the recording depth, relative to the microelectrode recording-defined boundaries of the STN within each trajectory. There was no significant difference in STN recording depth between speech response types (excitatory, inhibitory, mixed, no response; see Figure 6 (Kruskal-Wallis test, central trajectory $\chi^2 = 7.2$, $p = 0.066$; posterior trajectory $\chi^2 = 6.2$, $p = 0.10$; medial trajectory $\chi^2 = 7.2$, $p = 0.066$). There was also no significant difference in STN recording depth between locking response types (production onset-locked, cue-locked, locked to both events, no locking) in any of the recording trajectories (Kruskal-Wallis test, central trajectory $\chi^2 = 5.6$, $p = 0.13$; posterior trajectory $\chi^2 = 3.9$, $p = 0.27$; medial trajectory $\chi^2 = 2.1$, $p = 0.35$). Collapsing the recording depths across trajectories did not reveal significant differences between response types. Microelectrode recording locations additionally were normalized to MNI space, allowing for group-level analysis within a common coordinate system (Figure 7). Linear discriminant analysis was used to model speech-related response types and locking types of units based on their MNI coordinates, in order to test whether speech-related responses are anatomically segregated within the sampled region of the STN. The classification accuracy of this model was not higher than expected by chance.

Finally, we tested for potential influences of recording stability by comparing single-unit isolation between baseline and speech epochs, for each unit. Overall, 23/79 units showed a small but significant change between baseline and speech isolation scores (7 decreases, 16 increases, 3.8 ± 0.7% mean magnitude change from baseline; $p < 0.05$, permutation testing). Similarly, a significant change between baseline and speech signal-to-noise ratios was observed in 25/79 units (14 decreases, 11 increases, 9.5 ± 1.0% mean magnitude change from baseline; $p < 0.05$, permutation testing). However, the specific change in spike isolation measure was not consistently related to the speech-related modulation in firing. Specifically, among 22 units with increase-type responses, 4 showed decreases, and 5 showed increases between baseline and speech isolation scores (4.1 ± 1.3% mean magnitude change from baseline; $p < 0.05$, permutation testing); while 7 showed decreases, and 3 showed increases between baseline and speech signal-to-noise ratios (6.6 ± 1.2% mean magnitude change from baseline; $p < 0.05$, permutation testing). Similarly, among 13 units with decrease-type responses, 2 showed decreases, and 2 showed increases between baseline and speech isolation scores (4.6 ± 2.6% mean magnitude change from baseline; $p < 0.05$, permutation testing); while none showed an decrease, and 1 showed an increase between baseline and speech signal-to-noise ratios (18% mean change from baseline; $p < 0.05$, permutation testing). Among 7 units with mixed-type responses, none showed a decrease, and 3 showed increases between baseline and speech isolation scores (2.5 ± 1.3% mean change from baseline; $p < 0.05$, permutation testing); while none showed an decrease, and 1 showed an increase between baseline and speech signal-to-noise ratios (8% mean change from baseline; $p < 0.05$, permutation testing).

DISCUSSION

We found that both phasic increases and decreases in the discharge rate of STN neurons accompany the production of speech. In this study, subjects read aloud syllables presented on a computer screen, a behavioral paradigm that requires a series of neural events beginning from processing the visual cue to activating motor commands for the vocal organ. Neural events that occur early in this series, such as processing of the visual cue and forming a phonological plan, might be expected to be time-locked to cue presentation. Events that occur later in the series, such as forming and executing the motor speech plan, might

425 be expected to be time-locked to speech output. We showed that decrease-type responses are predominantly
426 locked to cue presentation and increase-type STN responses are predominantly locked to the onset of speech.
427 These findings suggest that STN inhibition may be associated with early, cognitive aspects of speech
428 production, while STN excitation may be associated with later, motor aspects of speech production.
429
430 The extent to which speech-related activity in the STN may reflect lower-order movement-related activity, akin
431 to results from studies involving simple limb movements, versus higher order functions has important
432 implications. Although kinematic aspects of speech production often improve following DBS (Pinto et al., 2004;
433 De Gaspari et al., 2006; Parsons et al., 2006; Mikos et al., 2011), a decrease in verbal fluency is the most
434 common cognitive side effect of STN-DBS, with specific deficits in lexical and grammatical processing having
435 been observed, albeit inconsistently across studies (Phillips et al., 2012). The observation of increases in firing
436 rates associated with speech onset are expected, in the context of previous studies of limb movement-related
437 activity. In STN recordings from both human subjects and non-human primates, firing rate increases comprise
438 75-93% of movement-related responses during active and passive limb movements (Wichmann et al., 1994;
439 Rodriguez-Oroz et al., 2001; Abosch et al., 2002; Starr et al., 2003; Theodosopoulos et al., 2003; Romanelli et
440 al., 2004; Schrock et al., 2009). We found that nearly half of increase-type responses in our study were locked
441 to the onset of speech, indicating that motor aspects of speech production are encoded in STN activity. A
442 significantly smaller proportion of increase-type responses was locked to cue presentation (7%) and to both
443 cue presentation and speech production onset (7%), with remaining responses not clearly associated with
444 either event.
445
446 In contrast, we observed that early stages of speech production may involve the inhibition of STN neurons. We
447 found that a large proportion (40%) of decrease-type responses were locked to cue presentation, with cue-
448 locked responses occurred at significantly lower latencies relative to cue presentation, compared with speech-
449 locked responses. A smaller proportion (15%) of decrease-type responses were locked to speech production
450 onset, with remaining responses not clearly associated with either event. Although minority populations of
451 neurons with movement-related firing-rate decreases have been reported previously (Wichmann et al., 1994;

Schrock et al., 2009; Lipski et al., 2017), and active movements have been associated with a higher proportion of decrease-type responses in the STN (Lipski et al., 2017), it is remarkable that such a high percentage of decrease-type responses were observed in the present study. Interestingly, and in contrast to our results, a marked reduction of STN activity was reported to be associated with the onset of speech production in the only previous report of STN unit activity recorded during speech production (Watson and Montgomery, 2006), although that study was largely descriptive in nature, limiting comparisons to our data. Although other investigators have shown correlations of STN single unit firing rates and rhythms to premotor functions, such as the encoding of difficulty level of a choice task (Zaghloul et al., 2012; Zavala et al., 2016), cue-locked decreases in firing were not reported. Our data do suggest that, in comparison to limb movement, speech may involve a different balance of activation and suppression in the STN, and that modulation of this balance may occur at the single neuron level prior to speech onset.

This study was not designed to determine whether early, cue-locked STN modulation of activity reflects responses to the presented stimulus (i.e. reading) versus other aspects of preparing to speak. Although it is important to note that cortical activation of motor commands, as well as adjustments in the chest wall, laryngeal and articulatory musculature, occur well before the acoustic signal is realized, and in a time-locked manner, we relied upon the acoustic output as a simple and non-invasive landmark for exploring timing relationships (Bouchard et al., 2013). Direct measurements of respiratory or articulatory kinematics, however, are indicated for futures studies, to more clearly understand behavioral correlates of STN speech-related activity. Whether similar STN responses would be observed with non-speech related engagement of the same musculature also is an open question. Notably, our findings are based on data collected in patients with PD, and it was not possible to determine the extent to which the dynamics of cortico-subthalamic coupling described reflect physiological versus pathophysiological basal ganglia function. Nonetheless, the prevalence of speech-responsive units did not relate to the subjects' symptom severity, as measured by UPDRS.

The STN functions within the basal ganglia thalamocortical circuit primarily by way of glutamatergic inputs to the GABAergic output neurons of the globus pallidus internus and substantia nigra pars reticulata. The firing

rate model of basal ganglia function posits that increases in STN activity may have a suppressive effect on basal ganglia-recipient circuits while decreases may be facilitatory. This balance of basal ganglia-mediated activation and suppression has been understood most frequently in terms of either selecting and focusing motor actions (Mink, 1996; Redgrave et al., 2010), or modulating their gain over time (Alexander and Crutcher, 1990b; Nambu et al., 2000, 2002; Nambu, 2005; Turner and Desmurget, 2010; Thura and Cisek, 2017). Proponents of an action selection hypothesis have proposed that the STN participates in a response inhibition function to reduce premature action when multiple competing responses are possible (Frank, 2006). Our findings of suppressed STN firing locked to speech cues and increased STN firing locked to speech production, however, are not consistent with action selection-related functions of the STN. Similarly, Zeigler and Ackerman (Ziegler and Ackermann, 2017) recently compiled extensive evidence in support of the idea that, for well-learned adult speech, basal ganglia circuits play key roles in the emotional/motivational modulation of speech (i.e., in prosody) but not in the selection and sequencing of articulatory gestures.

Speech-related phasic increases in the STN likely are a result of excitatory inputs and decreases likely a result of inhibitory inputs. The major excitatory input into the STN comes from the neocortex via the basal ganglia hyper-direct pathway (Nambu et al., 2002) which forms glutamatergic synapses onto distal dendrites of STN projection neurons (Künzle, 1978; Romansky et al., 1979; Kitai and Deniau, 1981; Romansky and Usunoff, 1987). The primate STN receives direct projections from broadly distributed cortical areas including primary motor cortex, pre-motor cortex, supplementary motor area, dorsolateral prefrontal, anterior cingulate, and inferior frontal cortex (Afsharpour, 1985; Parent and Hazrati, 1995; Nambu et al., 1997; Haynes and Haber, 2013). A primary form of inhibitory drive arises from GABAergic projections to the STN from the external segment of the globus pallidus, via the indirect basal ganglia pathway (Nauta and Mehler, 1966; Romansky and Usunoff, 1987; Bell et al., 1995; Sato et al., 2000). Thus, it is possible that speech onset-locked increased firing rate responses (STN excitation) could be mediated via the hyper-direct pathway, while cue-locked inhibitory responses during speech could be mediated via the indirect pathway. These findings also can be interpreted in the context of the GODIVA model (Bohland et al., 2010) of speech production. This model posits a dual role for the basal ganglia, participating in two processes that may be correlated with cue presentation

and speech production in our task: (1) a planning loop that is involved in generating a phonological sequence corresponding to the target word, and (2) a motor loop that releases the planned speech sounds for motor execution.

This study did not examine inter-hemispheric differences in speech-related STN activity, as recordings were performed in the left STN in 11/12 subjects. Similarly, language laterality was not specifically assessed, and the current cohort is skewed towards right-handed individuals (See Table 1). There are good reasons to expect both the left and right STN will exhibit speech-related responses, since speech-related potentials are represented in a bilateral fashion (Grözinger et al., 1980) and functional neuroimaging studies have demonstrated robust activation of the precentral and postcentral gyri bilaterally during overt speech production (Turkeltaub et al., 2002; Guenther and Hickok, 2015). Interestingly, clinical outcome studies on speech and STN DBS have suggested that left STN stimulation has a greater impact on speech production compared to right sided stimulation (Aldridge et al., 2016), thus future experiments designed to examine bilateral responses in individual patients are needed to address questions of the impact of language laterality.

It is important to consider that respiratory kinematics and articulatory movements may change intracranial pressure, potentially transiently affecting unit recording quality. In order to examine the possibility of these transient changes affecting our assessment of speech-related physiological modulation of STN neuron firing, we tested unit isolation measures following the onset of speech relative to baseline. We found that signal-to-noise ratio and isolation score were significantly altered in 25 and 23 of 79 units, respectively. Importantly, however, the magnitude of change in isolation measures was small, and the direction of change was not predictive of speech-related response type. Given that intraoperatively recorded human single-unit activity seldom is completely stable across time, the isolation measure difference between baseline and speech likely reflects ongoing fluctuations in isolation rather than specific effects of speech. Small changes in isolation during speech may also be attributed to modulation of background population activity during speech, which affects isolation of sorted units.

In summary, our results demonstrate that STN neurons comprise separate functional populations whose activity during speech production can be differentiated by the timing and direction of firing rate changes. The extent to which these functional groupings may be specific to speech versus common to complex motor function is an important question for future work, in light of conflicting theories of the role of the STN, and that of the basal ganglia as a whole, in motor behaviors. Our ongoing studies aim to examine the granularity of STN functional encoding in and to verify the specificity of these findings to speech production.

FIGURE LEGENDS

Figure 1. Speech task and representative spoken and neural responses. (A) Intraoperative syllable speech task. Subjects were asked to read aloud words presented on a computer screen. Each trial consisted of a sequence beginning with the fixation cross turning green for 250 ms, followed by a variable delay black screen (500-1000 ms), followed by a unique CVC syllable cue appearing on the screen until the response was recorded. A white fixation cross appeared during the inter-trial interval. (B) An example audio spectrogram time-aligned to the onset of a subject's utterance of the syllable "loath." The time (in s) of cue presentation is indicated by the solid vertical line, and the response onset and offsets are indicated by dotted lines. (C) A single unit recording from the subject's STN, showing an increase in firing during speech. Red hash marks indicate timing of detected spike waveforms from the background activity. (D) Overlay of 50 spike waveforms from the single unit shown in (C). Scatterplots of the first two principal components (E; principal component1 and principal component2), as well as the first principal component and spike timestamp (F), showing clear separation of single unit spike waveforms (red) corresponding to the example shown in (C) from background (blue).

Figure 2. Schematic illustrating cue- and speech production-locking neuronal response types. (A) Hypothetical example of cue-aligned trials, illustrating a constant neuronal response latency with varying speech production latencies. (B) Corresponding correlation schematic showing that a significant correlation between neuronal response to production onset interval and speech production latency indicates cue-locking. (C) Hypothetical example of cue-aligned trials, illustrating a constant neuronal response to production onset interval with varying speech production latencies. (D) Corresponding correlation schematic showing that a significant correlation between the neuronal response latency and speech production latency indicates speech-locking.

Figure 3. STN neuronal firing is modulated during speech. Examples of A-sort single unit neuronal responses during speech showing (A) increases, (B) decreases and (C) mixed responses in firing rate, aligned to production onset (t=0). Spike rasters across trials are shown on top in panels A-C, and mean firing rate (A,

C) or mean inter-spike interval (ISI; B) is shown on the bottom. Diamonds labeled with a “c” indicate mean time of cue presentation; diamonds labeled with an “e” indicate mean speech end; dashed error bars indicate the corresponding standard deviations. (D-F) Raster plots illustrating the timing of firing rate responses across the population of unit recordings. Each row represents a unit’s significant changes relative to baseline, during a time segment surrounding production onset. The time scale is normalized across units from 0.5 s before the mean cue onset until 0.5 s after the mean end of speech.

Figure 4. STN neuronal firing increases are primarily speech-locked. (A) An example of an A-sort single unit whose firing rate increase is locked to production onset. Spike raster (top) and mean firing rate (bottom) aligned to cue presentation. Significant spike bursts are shaded for each trial according to their Poisson Surprise index. Trials are sorted by speech production latency; speech production onset for each trial is indicated in green. (B) The time interval between cue presentation and burst onset (neuronal response latency) and between burst onset and production onset (neuronal response to production interval) for each trial is correlated against production latency. (C) Summary of correlation analyses for all unit recordings with increase-type responses, showing 12/29 responses locked to production onset (red circles), 2/29 responses locked to cue presentation (blue circles), and 2/29 responses locked to both cue and production onset (black circles). Open circles in (C) and indicate B sorts.

Figure 5. STN neuronal firing decreases are primarily cue-locked. (A) An example of an A-sort multi-unit whose firing rate decrease is locked to cue presentation. Spike raster (top) and mean firing rate (bottom) aligned to cue presentation. Significant decreases in firing rate (pauses) are shaded for each trial according to their Poisson Surprise index. Trials are sorted by speech production latency; speech production onset for each trial is indicated in green. (B) The time interval between cue presentation and pause onset (neuronal response latency) and between pause onset and production onset (neuronal response to production interval) for each trial is correlated against production latency. (C) Summary of correlation analyses for all unit recordings with inhibitory responses, showing 3/20 responses were locked to production onset (red circles), 8/20 units were

locked to cue presentation (blue circles), and none locked to both cue presentation and production onset (black circles). Open circles in (C) and indicate B sorts.

Figure 6. Anatomical distribution of speech responses in the STN. Unit locations are represented according to the recording trajectory and recording depth relative to electrophysiology-defined STN boundaries (0% corresponds to the ventral STN border and 100% corresponds to the dorsal STN border. Box plots represent the median and inter-quartile range of recording depths within each response category.

Figure 7. Anatomical distribution of STN microelectrode unit recordings in Montreal Neurological Institute space. (A) Speech-related unit response types and (B) locking types were not segregated in normalized anatomical coordinates.

TABLE LEGENDS

Table 1. Subject characteristics. Demographic, recording and speech performance characteristics. NR = not recorded, s.e. = standard error.

Table 2. Additional subject characteristics. Side of tremor dominance and presence of voice or hearing complaints.

Table 3. Unit type and sort quality do not determine response type.

Table 4. Subject symptom severity is not correlated with unit speech response types.

Table 5. Unit type and sort quality do not determine locking type.

Table 6. Dissociation between cue-locking decreases and speech-locking increases of firing.

REFERENCES

Abosch A, Hutchison WD, Saint-Cyr JA, Dostrovsky JO, Lozano AM (2002) Movement-related neurons of the subthalamic nucleus in patients with Parkinson disease. J Neurosurg 97:1167–1172 Available at: http://www.ncbi.nlm.nih.gov/pubmed/12450039 [Accessed February 9, 2015].

Afsharpour S (1985) Topographical projections of the cerebral cortex to the subthalamic nucleus. J Comp Neurol 236:14–28 Available at: http://www.ncbi.nlm.nih.gov/pubmed/2414329 [Accessed February 10, 2015].

Aldridge D, Theodoros D, Angwin A, Vogel AP (2016) Speech outcomes in Parkinson's disease after subthalamic nucleus deep brain stimulation: A systematic review. Parkinsonism Relat Disord 33:3–11 Available at: http://linkinghub.elsevier.com/retrieve/pii/S1353802016303807 [Accessed March 21, 2018].

Alexander GE, Crutcher MD (1990a) Preparation for movement: neural representations of intended direction in three motor areas of the monkey. J Neurophysiol 64:133–150 Available at: http://www.ncbi.nlm.nih.gov/pubmed/2388061 [Accessed September 20, 2017].

Alexander GE, Crutcher MD (1990b) Functional architecture of basal ganglia circuits: neural substrates of parallel processing. Trends Neurosci 13:266–271 Available at: http://www.ncbi.nlm.nih.gov/pubmed/1695401 [Accessed September 2, 2017].

Alm PA (2004) Stuttering and the basal ganglia circuits: a critical review of possible relations. J Commun Disord 37:325–369 Available at: http://linkinghub.elsevier.com/retrieve/pii/S0021992404000280 [Accessed September 20, 2017].

Anderson ME, Turner RS (1991) A quantitative analysis of pallidal discharge during targeted reaching movement in the monkey. Exp brain Res 86:623–632 Available at: http://www.ncbi.nlm.nih.gov/pubmed/1761096 [Accessed October 27, 2017].

Bell K, Churchill L, Kalivas PW (1995) GABAergic projection from the ventral pallidum and globus pallidus to the subthalamic nucleus. Synapse 20:10–18 Available at: http://doi.wiley.com/10.1002/syn.890200103 [Accessed November 16, 2017].

Bohland JW, Bullock D, Guenther FH (2010) Neural representations and mechanisms for the performance of simple speech sequences. J Cogn Neurosci 22:1504–1529 Available at:

http://www.mitpressjournals.org/doi/10.1162/jocn.2009.21306 [Accessed September 1, 2017].

Bouchard KE, Mesgarani N, Johnson K, Chang EF (2013) Functional organization of human sensorimotor cortex for speech articulation. Nature 495:327–332 Available at: http://www.ncbi.nlm.nih.gov/pubmed/23426266 [Accessed September 20, 2017].

Canteras NS, Shammah-Lagnado SJ, Silva BA, Ricardo JA (1988) Somatosensory inputs to the subthalamic nucleus: a combined retrograde and anterograde horseradish peroxidase study in the rat. Brain Res 458:53–64 Available at: http://www.ncbi.nlm.nih.gov/pubmed/2463044 [Accessed November 16, 2017].

De Gaspari D, Siri C, Di Gioia M, Antonini A, Isella V, Pizzolato A, Landi A, Vergani F, Gaini SM, Appollonio IM, Pezzoli G (2006) Clinical correlates and cognitive underpinnings of verbal fluency impairment after chronic subthalamic stimulation in Parkinson's disease. Parkinsonism Relat Disord 12:289–295 Available at: http://linkinghub.elsevier.com/retrieve/pii/S1353802006000071 [Accessed October 4, 2017].

Desmurget M, Turner RS (2010) Motor Sequences and the Basal Ganglia: Kinematics, Not Habits. J Neurosci 30:7685–7690 Available at: http://www.ncbi.nlm.nih.gov/pubmed/20519543 [Accessed November 29, 2017].

DiCarlo JJ, Maunsell JHR (2005) Using neuronal latency to determine sensory-motor processing pathways in reaction time tasks. J Neurophysiol 93:2974–2986 Available at: http://jn.physiology.org/cgi/doi/10.1152/jn.00508.2004 [Accessed September 20, 2017].

Doupe AJ, Kuhl PK (1999) Birdsong and human speech: common themes and mechanisms. Annu Rev Neurosci 22:567–631 Available at: http://www.annualreviews.org/doi/10.1146/annurev.neuro.22.1.567 [Accessed September 20, 2017].

Frank MJ (2006) Hold your horses: A dynamic computational role for the subthalamic nucleus in decision making. Neural Networks 19:1120–1136 Available at: http://www.ncbi.nlm.nih.gov/pubmed/16945502 [Accessed December 1, 2017].

Giraud A-L, Neumann K, Bachoud-Levi A-C, von Gudenberg AW, Euler HA, Lanfermann H, Preibisch C (2008) Severity of dysfluency correlates with basal ganglia activity in persistent developmental stuttering. Brain Lang 104:190–199 Available at: http://linkinghub.elsevier.com/retrieve/pii/S0093934X07000697 [Accessed September 20, 2017].

674 Grözinger B, Kornhuber HH, Kriebel J, Szirtes J, Westphal KT (1980) The Bereitschaftspotential preceding the
675 act of speaking. Also an analysis of artifacts. Prog Brain Res 54:798–804 Available at:
676 http://linkinghub.elsevier.com/retrieve/pii/S0079612308617057 [Accessed March 21, 2018].
677 Guenther FH, Hickok G (2015) Role of the auditory system in speech production. In: Handbook of clinical
678 neurology, pp 161–175 Available at: http://www.ncbi.nlm.nih.gov/pubmed/25726268 [Accessed March 21,
679 2018].
680 Haynes WIA, Haber SN (2013) The organization of prefrontal-subthalamic inputs in primates provides an
681 anatomical substrate for both functional specificity and integration: implications for Basal Ganglia models
682 and deep brain stimulation. J Neurosci 33:4804–4814 Available at:
683 http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3755746&tool=pmcentrez&rendertype=abstract
684 [Accessed January 21, 2015].
685 Hickok G (2012) Computational neuroanatomy of speech production. Nat Rev Neurosci 13:135–145 Available
686 at: http://www.ncbi.nlm.nih.gov/pubmed/22218206 [Accessed September 20, 2017].
687 Horn A, Kühn AA (2015) Lead-DBS: a toolbox for deep brain stimulation electrode localizations and
688 visualizations. Neuroimage 107:127–135 Available at:
689 http://linkinghub.elsevier.com/retrieve/pii/S1053811914009938 [Accessed March 19, 2018].
690 Jiao Y, Medina L, Veenman CL, Toledo C, Puelles L, Reiner A (2000) Identification of the anterior nucleus of
691 the ansa lenticularis in birds as the homolog of the mammalian subthalamic nucleus. J Neurosci 20:6998–
692 7010 Available at: http://www.ncbi.nlm.nih.gov/pubmed/10995845 [Accessed September 20, 2017].
693 Joshua M, Elias S, Levine O, Bergman H (2007) Quantifying the isolation quality of extracellularly recorded
694 action potentials. J Neurosci Methods 163:267–282 Available at:
695 http://linkinghub.elsevier.com/retrieve/pii/S0165027007001409 [Accessed March 14, 2018].
696 Kitai ST, Deniau JM (1981) Cortical inputs to the subthalamus: intracellular analysis. Brain Res 214:411–415
697 Available at: http://www.ncbi.nlm.nih.gov/pubmed/7237177 [Accessed November 16, 2017].
698 Künzle H (1978) An autoradiographic analysis of the efferent connections from premotor and adjacent
699 prefrontal regions (areas 6 and 9) in macaca fascicularis. Brain Behav Evol 15:185–234 Available at:
700 http://www.ncbi.nlm.nih.gov/pubmed/99205 [Accessed November 16, 2017].

Lipski WJ, Wozny TA, Alhourani A, Kondylis ED, Turner RS, Crammond DJ, Richardson RM (2017) Dynamics of human subthalamic neuron phase-locking to motor and sensory cortical oscillations during movement. J Neurophysiol 118.

Mikos A, Bowers D, Noecker AM, McIntyre CC, Won M, Chaturvedi A, Foote KD, Okun MS (2011) Patient-specific analysis of the relationship between the volume of tissue activated during DBS and verbal fluency. Neuroimage 54:S238–S246 Available at: http://linkinghub.elsevier.com/retrieve/pii/S1053811910003551 [Accessed October 4, 2017].

Mink JW (1996) The basal ganglia: focused selection and inhibition of competing motor programs. Prog Neurobiol 50:381–425 Available at: http://www.ncbi.nlm.nih.gov/pubmed/9004351 [Accessed November 16, 2017].

Nambu A (2005) A new approach to understand the pathophysiology of Parkinson's disease. J Neurol 252:iv1-iv4 Available at: http://www.ncbi.nlm.nih.gov/pubmed/16222431 [Accessed September 2, 2017].

Nambu A, Tokuno H, Hamada I, Kita H, Imanishi M, Akazawa T, Ikeuchi Y, Hasegawa N (2000) Excitatory cortical inputs to pallidal neurons via the subthalamic nucleus in the monkey. J Neurophysiol 84:289–300 Available at: http://www.ncbi.nlm.nih.gov/pubmed/10899204 [Accessed September 2, 2017].

Nambu A, Tokuno H, Inase M, Takada M (1997) Corticosubthalamic input zones from forelimb representations of the dorsal and ventral divisions of the premotor cortex in the macaque monkey: comparison with the input zones from the primary motor cortex and the supplementary motor area. Neurosci Lett 239:13–16 Available at: http://www.ncbi.nlm.nih.gov/pubmed/9547160 [Accessed February 10, 2015].

Nambu A, Tokuno H, Takada M (2002) Functional significance of the cortico-subthalamo-pallidal "hyperdirect" pathway. Neurosci Res 43:111–117 Available at: http://www.ncbi.nlm.nih.gov/pubmed/12067746 [Accessed September 2, 2017].

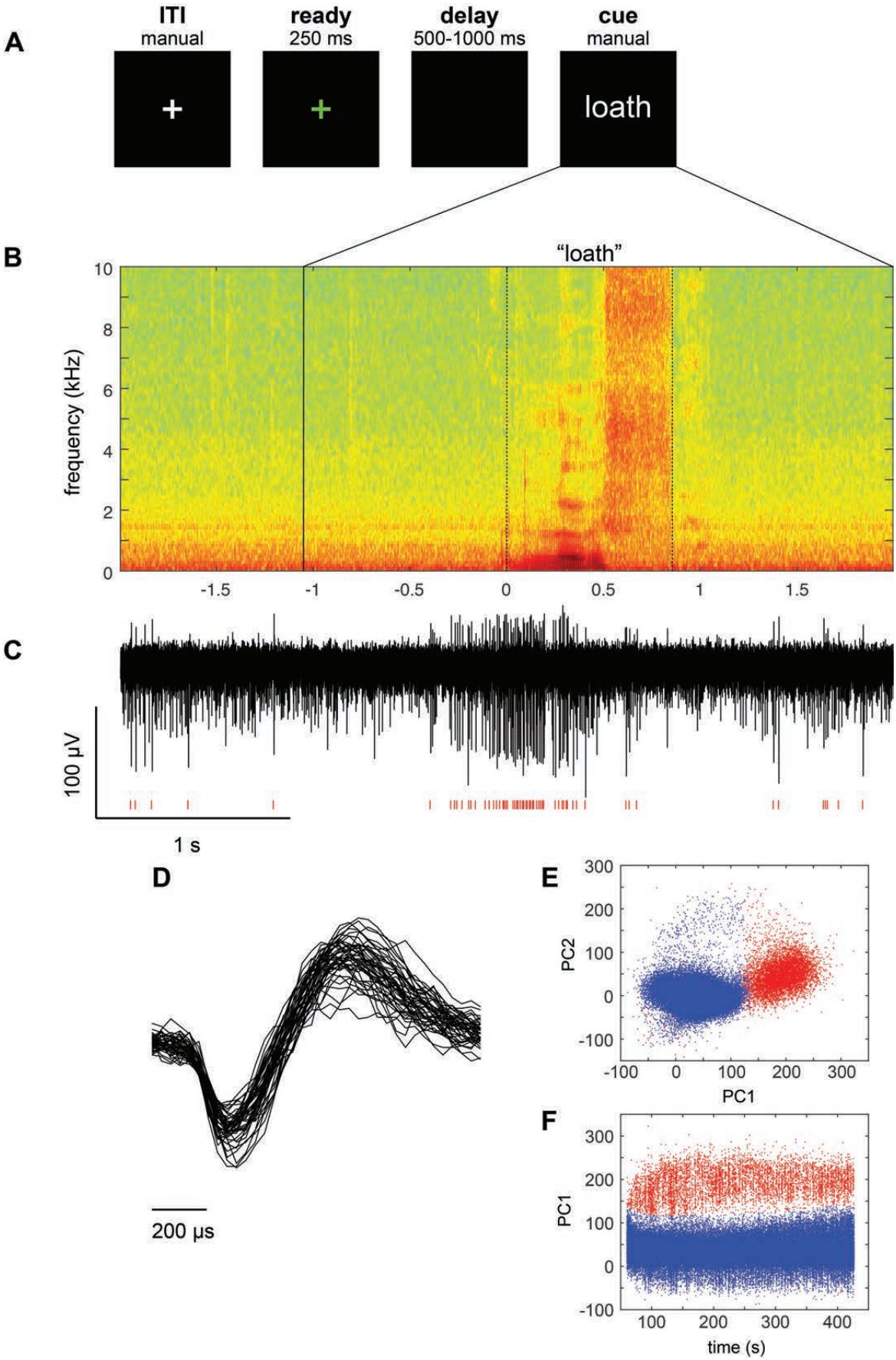
Nauta WJ, Mehler WR (1966) Projections of the lentiform nucleus in the monkey. Brain Res 1:3–42 Available at: http://www.ncbi.nlm.nih.gov/pubmed/4956247 [Accessed November 16, 2017].

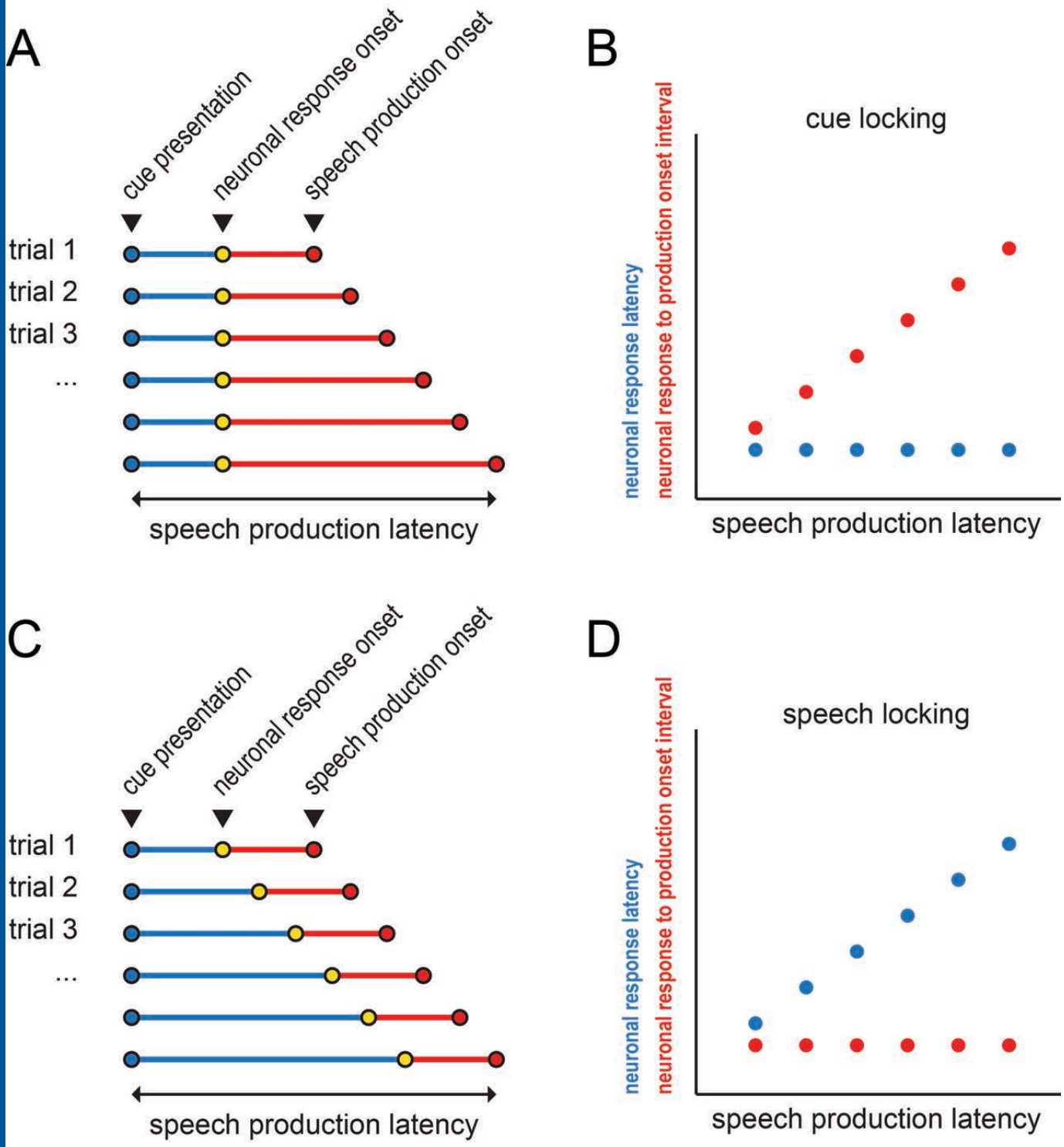
Parent A, Hazrati LN (1995) Functional anatomy of the basal ganglia. II. The place of subthalamic nucleus and external pallidum in basal ganglia circuitry. Brain Res Brain Res Rev 20:128–154 Available at: http://www.ncbi.nlm.nih.gov/pubmed/7711765 [Accessed February 10, 2015].

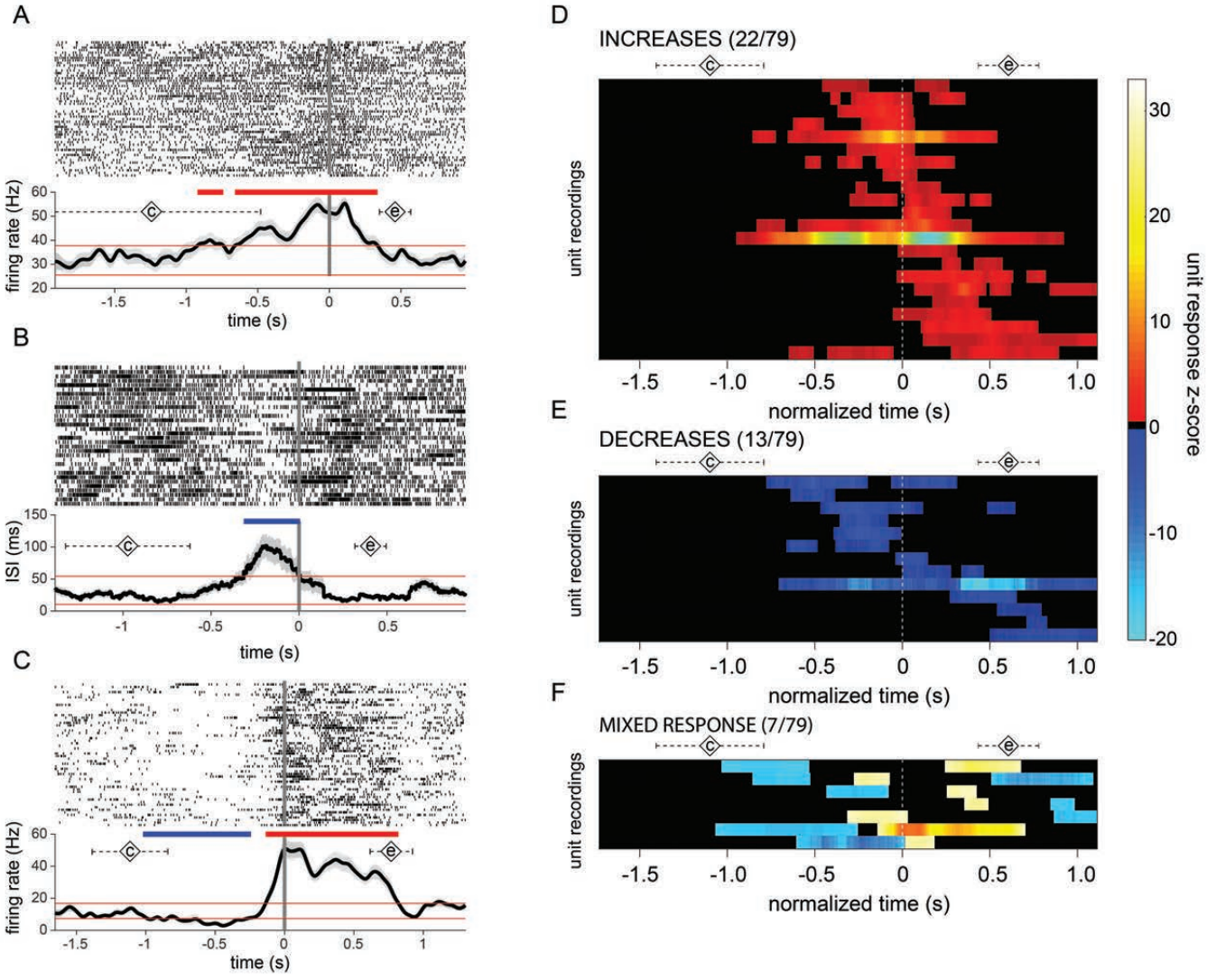
728 Parsons TD, Rogers SA, Braaten AJ, Woods SP, Tröster AI (2006) Cognitive sequelae of subthalamic nucleus
729 deep brain stimulation in Parkinson's disease: a meta-analysis. Lancet Neurol 5:578–588 Available at:
730 http://www.ncbi.nlm.nih.gov/pubmed/16781988 [Accessed October 4, 2017].
731 Phillips L, Litcofsky KA, Pelster M, Gelfand M, Ullman MT, Charles PD (2012) Subthalamic nucleus deep brain
732 stimulation impacts language in early Parkinson's disease. Sgambato-Faure V, ed. PLoS One 7:e42829
733 Available at: http://dx.plos.org/10.1371/journal.pone.0042829 [Accessed October 4, 2017].
734 Pinto S, Ozsancak C, Tripoliti E, Thobois S, Limousin-Dowsey P, Auzou P (2004) Treatments for dysarthria in
735 Parkinson's disease. Lancet Neurol 3:547–556 Available at:
736 http://linkinghub.elsevier.com/retrieve/pii/S1474442204008543 [Accessed October 4, 2017].
737 Redgrave P, Rodriguez M, Smith Y, Rodriguez-Oroz MC, Lehericy S, Bergman H, Agid Y, DeLong MR, Obeso
738 JA (2010) Goal-directed and habitual control in the basal ganglia: implications for Parkinson's disease.
739 Nat Rev Neurosci 11:760–772 Available at: http://www.ncbi.nlm.nih.gov/pubmed/20944662 [Accessed
740 November 16, 2017].
741 Rodriguez-Oroz MC, Rodriguez M, Guridi J, Mewes K, Chockkman V, Vitek J, DeLong MR, Obeso JA (2001)
742 The subthalamic nucleus in Parkinson's disease: somatotopic organization and physiological
743 characteristics. Brain 124:1777–1790 Available at: http://www.ncbi.nlm.nih.gov/pubmed/11522580
744 [Accessed February 9, 2015].
745 Romanelli P, Heit G, Hill BC, Kraus A, Hastie T, Brontë-Stewart HM (2004) Microelectrode recording revealing
746 a somatotopic body map in the subthalamic nucleus in humans with Parkinson disease. J Neurosurg
747 100:611–618 Available at: http://www.ncbi.nlm.nih.gov/pubmed/15070113 [Accessed February 9, 2015].
748 Romansky K V, Usunoff KG (1987) The fine structure of the subthalamic nucleus in the cat. II. Synaptic
749 organization. Comparisons with the synaptology and afferent connections of the pallidal complex and the
750 substantia nigra. J Hirnforsch 28:407–433 Available at: http://www.ncbi.nlm.nih.gov/pubmed/3655332
751 [Accessed November 16, 2017].
752 Romansky K V, Usunoff KG, Ivanov DP, Galabov GP (1979) Corticosubthalamic projection in the cat: an
753 electron microscopic study. Brain Res 163:319–322 Available at:
754 http://www.ncbi.nlm.nih.gov/pubmed/427548 [Accessed November 16, 2017].

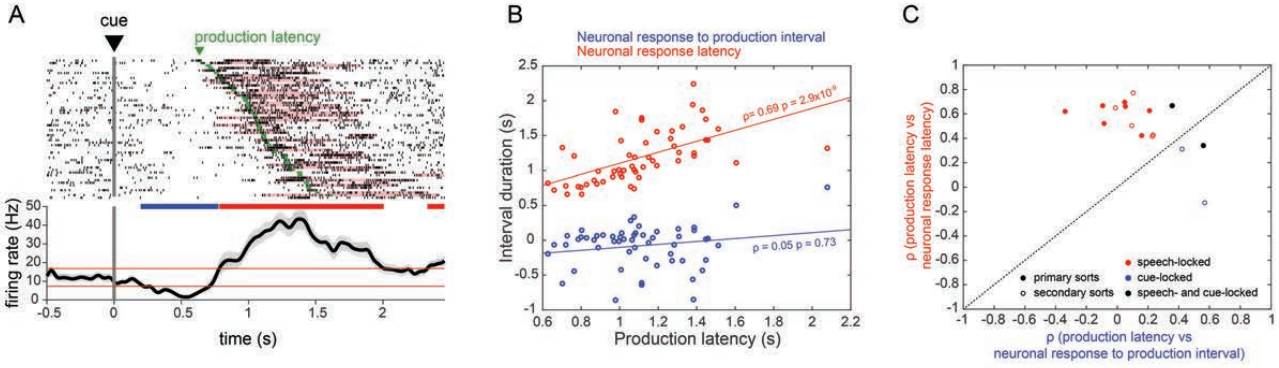
755 Sato F, Lavallée P, Lévesque M, Parent A (2000) Single-axon tracing study of neurons of the external segment
756 of the globus pallidus in primate. J Comp Neurol 417:17–31 Available at:
757 http://www.ncbi.nlm.nih.gov/pubmed/10660885 [Accessed November 16, 2017].
758 Schrock LE, Ostrem JL, Turner RS, Shimamoto SA, Starr PA (2009) The subthalamic nucleus in primary
759 dystonia: single-unit discharge characteristics. J Neurophysiol 102:3740–3752 Available at:
760 http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4073906&tool=pmcentrez&rendertype=abstract
761 [Accessed February 9, 2015].
762 Seal J, Commenges D (1985) A quantitative analysis of stimulus- and movement-related responses in the
763 posterior parietal cortex of the monkey. Exp brain Res 58:144–153 Available at:
764 http://www.ncbi.nlm.nih.gov/pubmed/3987845 [Accessed October 27, 2017].
765 Starr PA, Theodosopoulos P V, Turner R (2003) Surgery of the subthalamic nucleus: use of movement-related
766 neuronal activity for surgical navigation. Neurosurgery 53:1146–9; discussion 1149 Available at:
767 http://www.ncbi.nlm.nih.gov/pubmed/14580281 [Accessed February 9, 2015].
768 Theodosopoulos P V, Marks WJ, Christine C, Starr PA (2003) Locations of movement-related cells in the
769 human subthalamic nucleus in Parkinson's disease. Mov Disord 18:791–798 Available at:
770 http://www.ncbi.nlm.nih.gov/pubmed/12815658 [Accessed February 9, 2015].
771 Thura D, Cisek P (2017) The Basal Ganglia Do Not Select Reach Targets but Control the Urgency of
772 Commitment. Neuron 95:1160–1170.e5 Available at: http://www.ncbi.nlm.nih.gov/pubmed/28823728
773 [Accessed November 16, 2017].
774 Toyomura A, Fujii T, Kuriki S (2015) Effect of an 8-week practice of externally triggered speech on basal
775 ganglia activity of stuttering and fluent speakers. Neuroimage 109:458–468 Available at:
776 http://linkinghub.elsevier.com/retrieve/pii/S1053811915000385 [Accessed September 20, 2017].
777 Turkeltaub PE, Eden GF, Jones KM, Zeffiro TA (2002) Meta-analysis of the functional neuroanatomy of single-
778 word reading: method and validation. Neuroimage 16:765–780 Available at:
779 http://www.ncbi.nlm.nih.gov/pubmed/12169260 [Accessed March 21, 2018].
780 Turner RS, Desmurget M (2010) Basal ganglia contributions to motor control: a vigorous tutor. Curr Opin
781 Neurobiol 20:704–716.

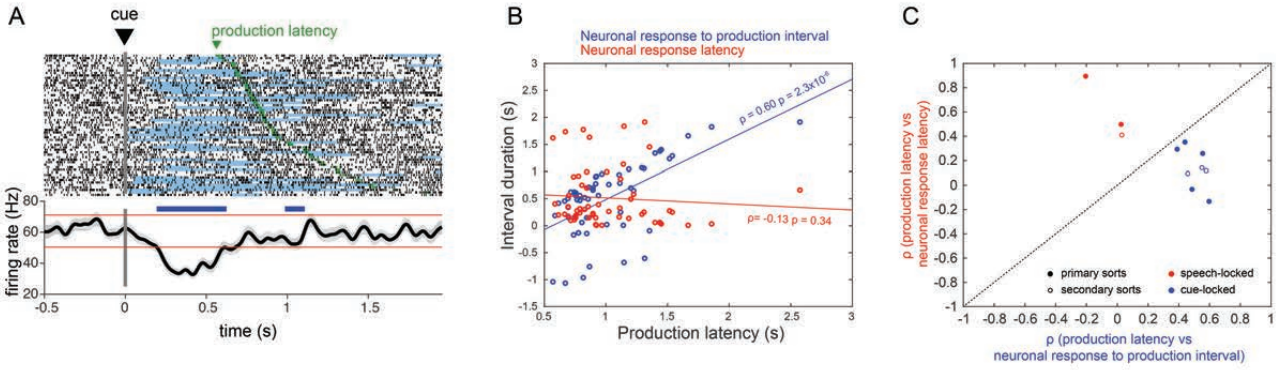
782 Watson P, Montgomery EB (2006) The relationship of neuronal activity within the sensori-motor region of the
783 subthalamic nucleus to speech. Brain Lang 97:233–240 Available at:
784 http://linkinghub.elsevier.com/retrieve/pii/S0093934X05003354 [Accessed October 4, 2017].
785 Wichmann T, Bergman H, DeLong MR (1994) The primate subthalamic nucleus. I. Functional properties in
786 intact animals. J Neurophysiol 72:494–506 Available at: http://www.ncbi.nlm.nih.gov/pubmed/7983514
787 [Accessed June 18, 2015].
788 Zaghloul KA, Weidemann CT, Lega BC, Jaggi JL, Baltuch GH, Kahana MJ (2012) Neuronal activity in the
789 human subthalamic nucleus encodes decision conflict during action selection. J Neurosci 32:2453–2460
790 Available at: http://www.ncbi.nlm.nih.gov/pubmed/22396419 [Accessed September 20, 2016].
791 Zavala B, Tan H, Ashkan K, Foltynie T, Limousin P, Zrinzo L, Zaghloul K, Brown P (2016) Human subthalamic
792 nucleus-medial frontal cortex theta phase coherence is involved in conflict and error related cortical
793 monitoring. Neuroimage 137:178–187 Available at: http://www.ncbi.nlm.nih.gov/pubmed/27181763
794 [Accessed September 20, 2016].
795 Ziegler W, Ackermann H (2017) Subcortical Contributions to Motor Speech: Phylogenetic, Developmental,
796 Clinical. Trends Neurosci 40:458–468 Available at: http://www.ncbi.nlm.nih.gov/pubmed/28712469
797 [Accessed October 27, 2017].
798

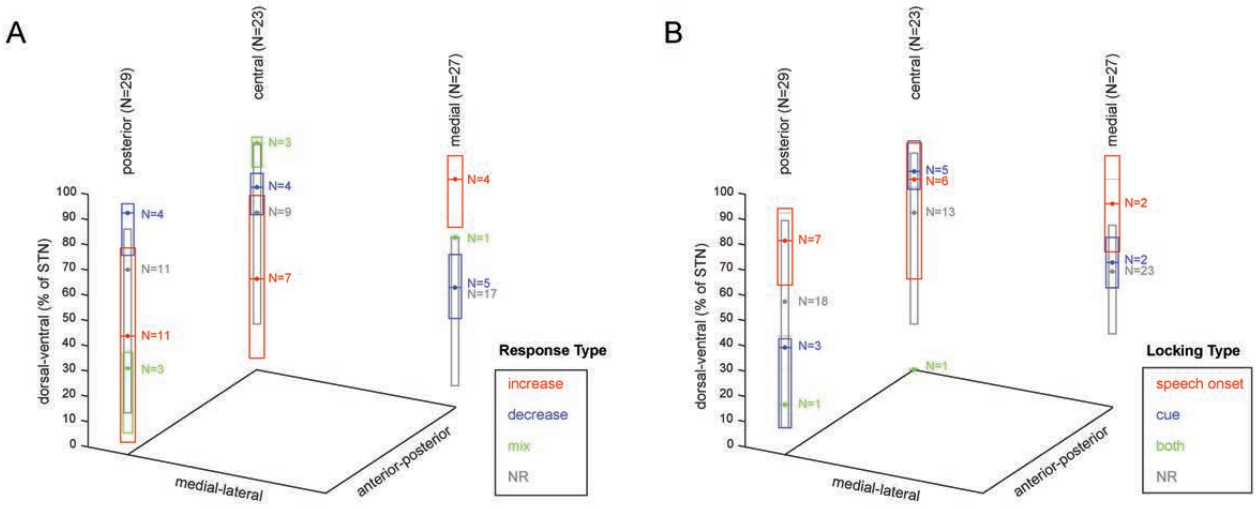




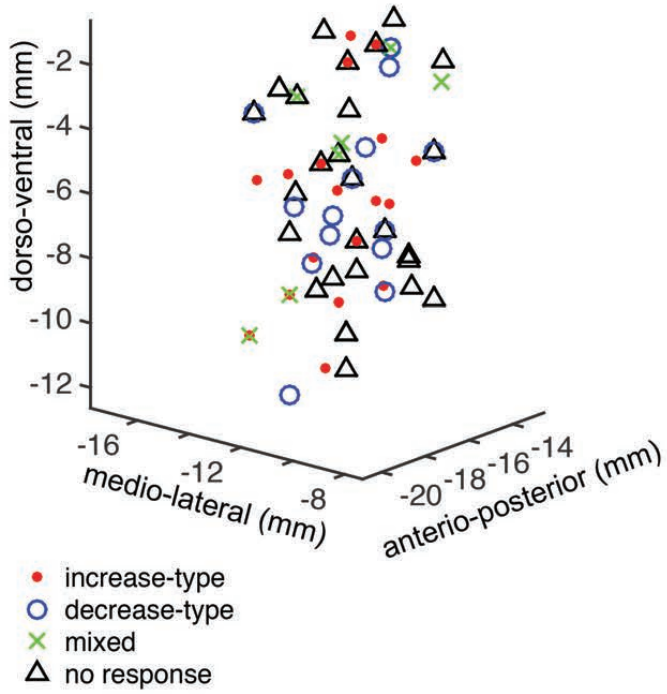








A



B

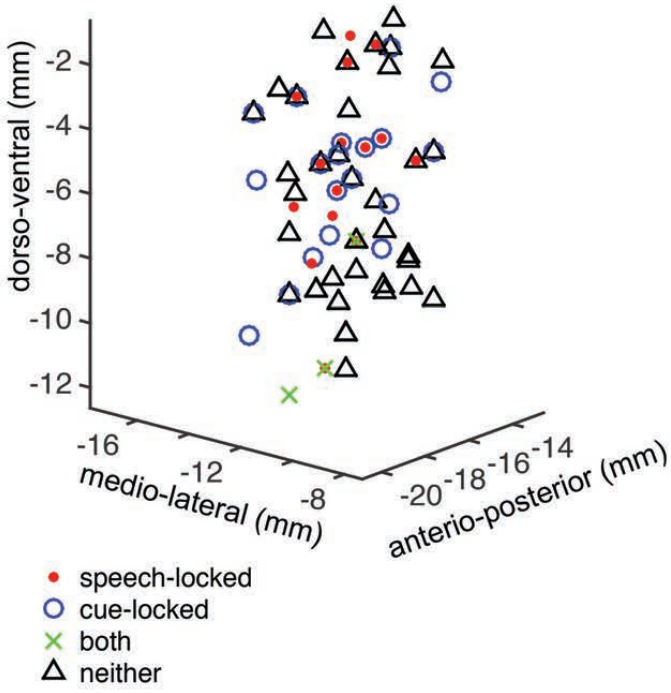


Table 1. Subject Characteristics.

Subject	Age	sex	handedness	UPDRS III off score		recorded hemisphere	# units recorded	# sessions	mean production latency (s)	production latency S.E. (s)	mean speech duration (s)	speech duration S. E. (s)	% correct trials
				speech	total								
1	60	male	R	NR	53	L	6	2	1.5	0.005	0.777	0.005	89%
2	68	male	R	1	47	L	15	4	1.01	0.016	0.578	0.016	95%
3	47	female	R	NR	NR	R	2	3	0.91	0.010	0.623	0.01	100%
4	60	male	R	0	31	L	1	2	0.77	0.020	0.642	0.02	98%
5	68	male	L	1	50	L	6	2	0.75	0.039	0.592	0.039	97%
6	56	male	R	1	46	L	5	2	0.86	0.009	0.467	0.009	98%
7	82	male	R	2	36	L	11	3	1.99	0.007	0.442	0.007	76%
8	66	male	R	0	46	L	8	4	0.67	0.007	0.61	0.007	97%
9	66	male	R	2	45	L	8	2	1.13	0.023	0.745	0.023	93%
10	71	female	R	1	24	L	6	3	1.26	0.016	0.796	0.016	86%
11	77	male	R	1	27	L	2	2	1.33	0.012	0.539	0.012	96%
12	59	male	R	1	40	L	9	3	1.04	0.004	0.452	0.004	96%
mean ± s.e.	65.0 ± 2.7	-	-	1.0 ± 0.2	40.4 ± 2.9	-	6.6 ± 1.2	2.7 ± 0.2	1.10 ± 0.11	0.014 ± 0.003	0.605 ± 0.035	0.014 ± 0.003	93 ± 2%

Table 2. Additional Subject Characteristics.

Subject	tremor dominance	voice or respiratory complaints	hearing complaints
1	Bilateral (greater on L)	none	none
2	NR	none	none
3	L	none	none
4	NR	none	none
5	Bilateral (greater on R)	none	none
6	L	none	none
7	NR	none	none
8	NR	dysphonia; atrophy of the bilateral true vocal fold; hypophonic speech related to Parkinsonism and atrophy.	none
9	no	NR	yes
10	NR	vocal fold atrophy, dysphonia, dysphagia	none
11	yes	vocal fold atrophy, dysphonia	bilateral hearing aids
12	no	none	none

Table 3. Unit type and sort quality do not determine response type.

	total	no response	increase	decrease	mixed
A sort units	45	19	13	10	3
B sort units	34	18	9	3	4
χ^2	-	0.89	0.06	2.53	0.62
p	-	0.34	0.81	0.11	0.43
Single units	25	10	6	5	4
Multi units	54	27	16	8	3
χ^2	-	0.69	0.27	0.33	2.3
p	-	0.41	0.60	0.56	0.13
total	**79**	**37**	**22**	**13**	**7**
% of baseline firing rate (mean ± s.e.)			**178 ± 15%**	**68 ± 3%**	**203 ± 30%** **62 ± 7%**

Table 4. Subject symptom severity is not correlated with unit speech response types.

Correlation with Speech UPDRS	proportion of units by response type			
	increase-type	decrease-type	speech-locked	cue-locked
Spearman ρ	0.28	0.08	0.39	0.45
p-value	0.44	0.82	0.27	0.20
Correlation with Total UPDRS	proportion of units by response type			
	increase-type	decrease-type	speech-locked	cue-locked
Spearman ρ	0.04	-0.03	-0.51	-0.30
p-value	0.92	0.92	0.11	0.36

Table 5. Unit type and sort quality do not determine locking type.

	total	locked to cue	locked to production onset	locked to both
A sort units	29	5	9	2
B sort units	20	5	6	0
χ^2	-	0.44	0.006	1.44
p	-	0.51	0.94	0.23
Single units	19	3	5	1
Multi units	30	7	10	1
χ^2	-	0.41	0.27	0.11
p	-	0.52	0.60	0.74
total	**49**	**10**	**15**	**2**

Table 6. Dissociation between cue-locking decreases and speech-locking increases of firing.

	total	locked to cue	locked to production onset	locked to both
Increase-type responses	29	2 (7%)	12 (41%)	2 (7%)
Decrease-type responses	20	8 (40%)	3 (15%)	0 (0%)
χ^2	-	7.99	3.89	1.44
p	-	0.0047*	0.049*	0.23
total	**49**	**10**	**15**	**2**