Spatiotemporally-specific cortical-subthalamic coupling differentiates aspects of speech performance

Matteo Vissani[1,2,9], Alan Bush[1,2], Witold J. Lipski[3], Petra Fischer[4], Clemens Neudorfer[1,2], Lori L. Holt[5], Julie A. Fiez[6], Robert S. Turner[3], R. Mark Richardson[1,2,+]

Affiliations

[1]Department of Neurosurgery, Massachusetts General Hospital, Boston, MA, 02114, USA

[2]Harvard Medical School, Boston, MA, 02115, USA

[3]Department of Neurobiology, Systems Neuroscience Center and Center for Neuroscience, University of Pittsburgh School of Medicine, Pittsburgh, PA, 15213, USA

[4]School of Physiology, Pharmacology & Neuroscience, University of Bristol, University Walk, BS8 1TD Bristol, United Kingdom

[5]Department of Psychology, The University of Texas at Austin, Austin, TX 78712 USA

[6]Department of Psychology, University of Pittsburgh, Pittsburgh 15260, PA, USA

[+]Corresponding author
mark.richardson@mgh.harvard.edu

Word counts

Number of words in the abstract: 150 / 150

Number of display items: 8/ 8

Summary

Speech provides a rich context for exploring human cortical-basal ganglia circuit function, but direct intracranial recordings are rare. We recorded electrocorticographic signals in the cortex synchronously with single units in the subthalamic nucleus (STN), a basal ganglia node that receives direct input from widespread cortical regions, while participants performed a syllable repetition task during deep brain stimulation (DBS) surgery. We discovered that STN neurons exhibited spike-phase coupling (SPC) events with distinct combinations of frequency, location, and timing that indexed specific aspects of speech. The strength of SPC to posterior perisylvian cortex predicted phoneme production accuracy, while that of SPC to perirolandic cortex predicted time taken for articulation Thus, STN-cortical interactions are coordinated via transient bursts of behavior-specific synchronization that involves multiple neuronal populations and timescales. These results both suggest mechanisms that support auditory-sensorimotor integration during speech and explain why firing-rate based models are insufficient for explaining basal ganglia circuit behavior.

Keywords

Parkinson's disease, Subthalamic nucleus, Speech production, Superior temporal gyrus, Spike-phase coupling, Deep brain stimulation, Cortico-subcortical connectivity.

Introduction

Effective communication during normal conversations necessitates the seamless understanding and production of fluent speech. Speaking entails the precise coordination and sequential movement of multiple articulators in the order of milliseconds [1,2]. Lips, jaw, larynx, and tongue work in rapid harmony, molding and reshaping our vocal tract to generate and combine discrete phonological units (phonemes or syllables) into meaningful words and sentences. Subserving the coordination of articulation is an intricate cortico-subcortical network, whose disruption can lead to significant deficits in speech processing and production, severely compromising an individual's ability to communicate effectively. Impairments within the basal ganglia-thalamocortical loop have been linked to speech-related challenges in Parkinson's Disease (PD), Dystonia, Huntington's Disease and Multiple System Atrophy.

Despite the significant impact and prevalence of these disorders, the role of the basal ganglia in speech processing remains under-explored. Much of the existing literature has emphasized expansive cortical networks and their dynamics, while often minimizing the role of the basal ganglia[2–4]. The limitations of functional magnetic resonance imaging (fMRI) have further contributed to this oversight, as paradigms for measuring the activity of individual basal ganglia nuclei have often precluded the explicit integration of subcortical nodes within current theoretical frameworks [1,5–7], despite the fact that the basal ganglia modulate all types of human behavior, including speech. The scarcity of invasive subcortical data during speech production presents a knowledge gap that hinders our broader understanding of basal ganglia function in physiological and pathological contexts and limits the expansion of current cortico-centric models of speech production (e.g., DIVA [5] and SFC [6]).

Deep Brain Stimulation (DBS) of the subthalamic nucleus (STN) offers a unique means to explore the physiological mechanisms of speech. The STN occupies a privileged role in this network, receiving monosynaptic cortical projections through the hyperdirect pathway [8–10], in addition to striatal input through the indirect pathway, positioning it as a hub for the temporal integration and spatial compression of information originating from multiple cortical areas [11,12]. The discovery of single unit and population level activity in the subthalamic nucleus that tracks multiple aspects of speech production [13–15] and the emerging evidence of anatomical [9] and functional connectivity [16] between the STN and sensorimotor and auditory cortical areas, raises the question of how STN and cortex interact to mediate speech-related behavior. Despite critical insights, however, the exact substrates and dynamics underlying the regulation of speech in the basal ganglia remain elusive [17–19].

To understand how different areas of speech-related cortex influence STN activity, we employed high-density electrocorticography (ECoG) to record local field potentials across cortical areas known to participate in speech perception, planning and, production, simultaneous with recording STN single-unit activity during DBS surgery. This unique approach provides the necessary resolution to investigate the spatiotemporal organization of cortico-subcortical interactions with millisecond resolution during speech

production. We examined (1) whether spike-phase coupling (SPC) – the relationship between STN spike timing and cortical oscillation phase – was modulated during a syllable triplet repetition task, (2) the spatial and behavioral (temporal) specificity of these relationships and (3) if SPC reflected participants' ability to utter the desired phoneme and the time required for speech production.

Results

We studied intracranial recordings in 24 participants (see Table S1 for demographic details) undergoing DBS surgery for the treatment of Parkinson's disease. High-density field potentials (ECoG) across the left ventral sensorimotor cortex, superior temporal gyrus, and inferior frontal regions were recorded simultaneously with single neuron activity from the STN (Figure 1A). Following the presentation of an auditory cue of a syllable triplet comprised of three unique consonant-vowel (CV) pairs, participants were instructed to repeat the syllable triplet (speech production) at their own pace into an omnidirectional microphone (64 recording sessions, 2.67 ± 0.62 sessions, 379.88 ± 99.49 trials on average across participants) (Figure 1B). Participants produced the CV-CV-CV sequences in 1.35 ± 0.41 s with a phonetic accuracy of 93.92 ± 14.99 %.

Cortical potentials and subthalamic neurons exhibit complex dynamics

We employed Wavelet decomposition to extract the spectral components (4-140 Hz) of cortical field potentials. Expected cortical evoked activity was observed during both listening (locked to auditory cue onset) and speech production (locked to speech onset)(Figure 1) [20–22]. Figure S1 illustrates different patterns of evoked activity in five representative electrodes from four participants. Spectrograms demonstrated consistent neural suppression in lower α–β frequencies (8-30 Hz) as well as elevation in the gamma γ range (50-150 Hz) during auditory cue presentation and speech production. A large fraction of STG electrodes displayed either briefly increased or sustained γ–activity in response to auditory cues[23]. As expected, premotor and postcentral gyrus electrodes showed β suppression and γ–elevation preceding the speech onset and during speech production. The same channels demonstrated above-baseline β increase (i.e., rebound) after the speech offset. Our data also revealed more complex response profiles, such as γ–activation of STG electrodes during speech (Figure S1), consistent with the role of STG during auditory feedback [21,24].

During the speech production task, we recorded 245 neurons in the STN, out of which 211 were stable and isolated (Figure 1D shows recording density, on average 3.28 ± 1.25 neurons per session). Spike sorting and quality metrics were conducted as previously described [14,25]. Our analysis demonstrated a rich and complex speech-related modulation of the instantaneous firing rate (IFR) (Figure 1D). Specifically, N = 84/211 neurons (39.8%) exhibited a significant increase in their firing rate in a window around the speech onset. Other neurons (N = 37/211, 17.5%) displayed a decrease in their firing rate.

Interestingly, N = 23/211 neurons (10.9%) showed mixed behavior with both increased and decreased firing rates. The remaining neurons (N = 67/211, 31.8%, No IFR modulation) did not exhibit a significant modulation of firing rate during the speech production task. For a comprehensive description of firing rate modulation in this dataset, see Lipski et al [26].

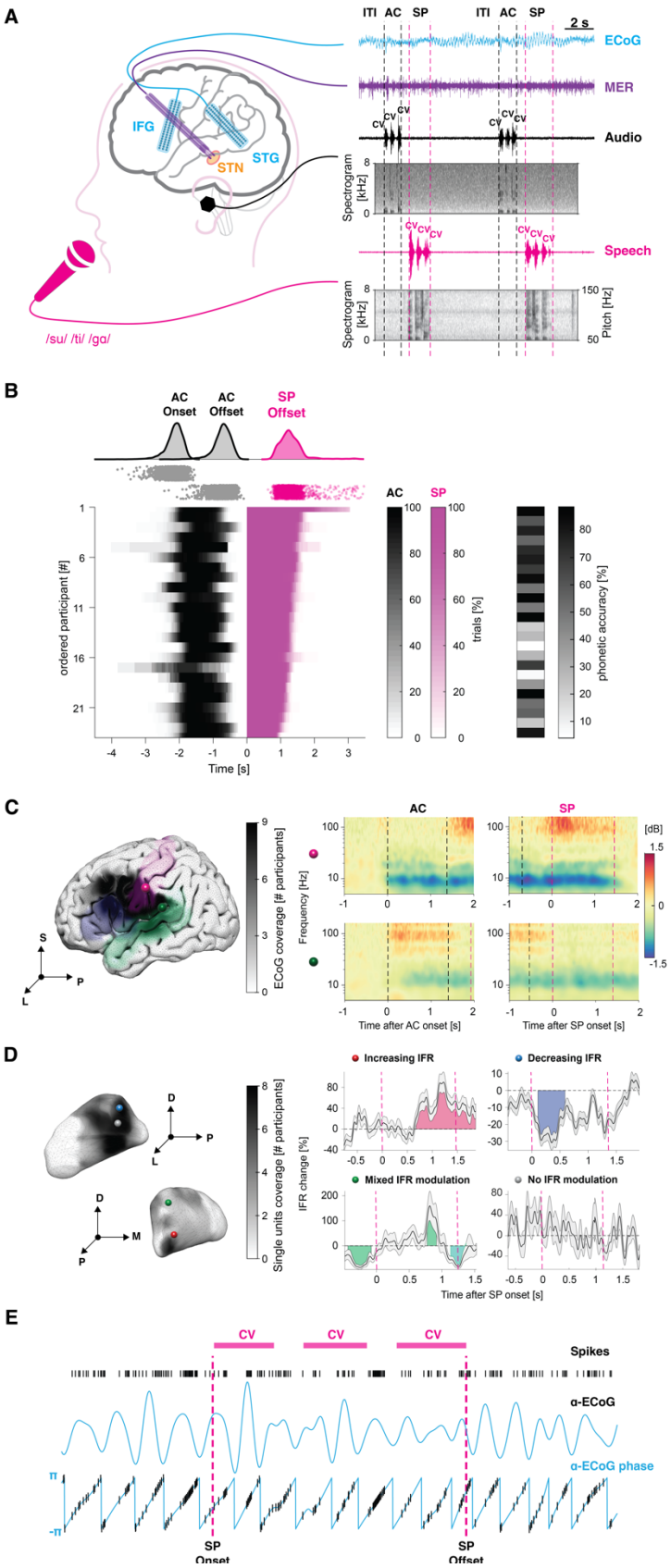

Figure 1: Syllable triplet repetition task and electrophysiological recordings. (**A**) Illustration of the syllable triplet repetition task. Participants are instructed to repeat unique consonant-vowel ("CV") syllable triplets (magenta). The auditory stimuli are presented through earphones (black). Electrocorticography (ECoG) local field potentials are recorded using high-density subdural ECoG strips placed in auditory and sensorimotor areas through the burr hole (cyan). Microelectrode recordings are acquired in the subthalamic nucleus during functional mapping (purple). Spectrograms of the audio signals are shown. (**B**) Timing of behavioral events, relative to speech-onset. Heatmap of the duration of the auditory cue (AC) and speech production (SP) windows expressed as percentage across trials for each participant. Average phonetic accuracy of the produced syllables for each participant are shown on the right. (**C**) ECoG strips localizations. The coverage of the ECoG strips across participants is superimposed on three different target areas from the Destrieux atlas[26]: postcentral gyrus (purple), inferior frontal gyrus (blue) and superior temporal gyrus (green). Exemplary auditory-locked and speech-locked spectrograms of activity in the postcentral gyrus (purple sphere) and superior temporal gyrus (green sphere) after normalization with respect to the baseline are displayed. (**D**) MER localization. Coverage of single units across participants is depicted. Spheres denote the location of four exemplary neurons with different categories of IFR modulation (expressed as Mean ± SEM percentage change for visualization): Increasing IFR (red), Decreasing IFR (blue), Mixed IFR modulation (green) and No IFR modulation (gray). (**E**) Exemplary transient spike-phase coupling in the α range during the speech production window (magenta dashed line). Spike timestamps, α oscillations, and instantaneous phase are illustrated. Magenta bars delineate the duration of the syllable triplet. List of abbreviations: ITI (inter-trial interval), AC (auditory cue), SP (speech production) and IFR (instantaneous firing rate).

STN neuronal spiking locks transiently with cortical field potentials

To investigate whether the subthalamic nucleus exhibits cortical phase-of-firing coding that is time-aligned with key events in speech production (Figure 1D), we employed a variable-window width spike-phase coupling (SPC) estimation procedure[27]. This method provides an unbiased, time-resolved estimate of the strength of spike-phase coupling across multiple frequencies (4-140 Hz) over the entire duration of the task, overcoming the limitations of traditional event-locked analyses. To account for variability in the number of spikes during low and high firing rate periods, the method adjusts the window width centered around each computational bin, resulting in more accurate and less biased SPC estimates. The average window width was 0.15 ± 0.01 s with an average number of 350 ± 169 spikes across trials per window (see Methods for details). We obtained 19755 time-frequency SPC maps between 4-140 Hz over the duration of the trial (from 0.75 s before the auditory cue onset to 0.75s after speech production termination), 92.78 ± 33.28 maps for each neuron versus signals from different ECoG contacts. Cluster-based permutation tests revealed that approximately 11% (2148/19755) of the time-frequency maps displayed significant SPC (see methods for details). The distribution of the percentage of significant SPC pairs across participants is illustrated in Figure S2. When averaging only significant SPC maps (N = 2148), we observed tonic (θ-α)-SPC around 4-12 Hz during the speech production interval, and a broad suppression of beta-SPC around 13-21 Hz from auditory cue presentation through speech production (Figure 2A). Interestingly, we found an increase in β SPC, consistent with movement-offset β rebound [28], around the time of speech production offset. Notably, these results were consistent whether we averaged all SPC maps (N = 19755) or the most significant SPC map for each unit (N = 211) (Figure S3).

We sought to investigate the extent to which single-pair SPC maps accurately reflect the group-level SPC patterns observed during prolonged SPC changes. Strikingly, examples of single-pair SPC maps showed that STN neurons preferentially entrain cortical phases only during brief and transient episodes (Figure 2B), which we termed t-SPC (N = 2987). Although most significant SPC maps exhibit only one t-SPC (77%, 1652/2148), we also found examples with multiple (up to seven) t-SPC events, which can occur during different key events of the task and in different frequency bands (Figure 2B-C). Thus, the observed difference between group-level SPC and single-pair SPC highlights that periods of increased or suppressed group-level SPC signal the most likely type of t-SPC occurrence.

We then characterized the task-related timing and frequency centroid of t-SPC events to test whether STN neurons display speech-related frequency-specific SPC. Our analysis revealed that t-SPC events had a median duration of 0.268 s and the majority were observed in the beta range (~16 Hz) (Figure 2D). We observed a mild negative correlation between t-SPC duration and t-SPC frequency centroid (R^2 = 0.12 (ρ = - 0.35), $p < 0.001$), suggesting that STN spiking is more likely to entrain for longer periods of time to lower frequencies. Figure 2E lists all the t-SPC events ordered by frequency band. When we grouped the t-SPC occurrence by frequency band, we found that (θ-α) t-SPC events significantly

aggregate during speech production ($p_{perm} < 0.05$, permutation test) (Figure 2F). Notably, α t-SPC events occurred throughout the entire speech production duration, whereas θ t-SPC events signaled preferentially the final part of the utterance. Moreover, α t-SPC events decreased during the auditory cue presentation ($p_{perm} < 0.05$, permutation test). As expected, β t-SPC events were more prominent during the inter-trial interval (ITI), dipped during both auditory cue presentation and speech production, and rebounded above baseline levels after the natural termination of the utterance ($p_{perm} < 0.05$, permutation test). Neither γ_L nor γ_H t-SPC event occurrence showed prominent deviation from uniform distribution during the task.

Speech orchestrates frequency-specific cortical phase coupling to subthalamic spikes

To further understand the relationship between STN SPC and speech production, we tested whether STN neurons share preferred phase of firing or if they tune to different preferential phases. We quantified the t-SPC phase and compared the consistency across t-SPC events over time. The polarity standardization procedure implemented in our pipeline allowed the comparison of phases across recordings (see Methods). We found that only α t-SPC events are locked around the same phase of firing (108°, during the decay after the peak of oscillation) during speech production ($p< 0.05$, Hodges-Ajne test) (Figure 2G). In contrast, β t-SPC events are uniformly locked around the trough (-90°) of the oscillation before and after the β t-SPC rebound ($p< 0.05$, Hodges-Ajne test), consistent with previous studies that employed electroencephalography or low-density EEG strips. These results suggest that cortical beta rebound entrains STN activity in a different manner than occurs for baseline coupling.

We then utilized the t-SPC events in the 5-40Hz range to derive time information from the relationship between unwrapped t-SPC phase and t-SPC centroid frequency (see Methods). A positive slope indicates that STN spikes occur at a consistent time lag relative to the peak of the rhythmic ECoG activity, at a latency given by the slope divided by 2π. Grouping all data together, the gradient of the positive slope translated to a time lag of 40.92 ms between STN spikes and cortical activity ($R^2 = 0.76$, $p_{perm} < 0.001$) (Figure 2H). When we performed the same analysis over time, we found that this relationship was particularly consistent during the β t-SPC rebound (39.81 ms, black dots in Figure 2H bottom). Interestingly, STN spiking transiently led ECoG activity during speech production (-32 ms), i.e., the sign of the slope switched, mirrored the dominance of (θ-α) t-SPC over β t-SPC events (green curve in Figure 2H bottom). This result suggests that (θ-α) t-SPC and β t-SPC observed here may reflect information flow from STN back to cortex.

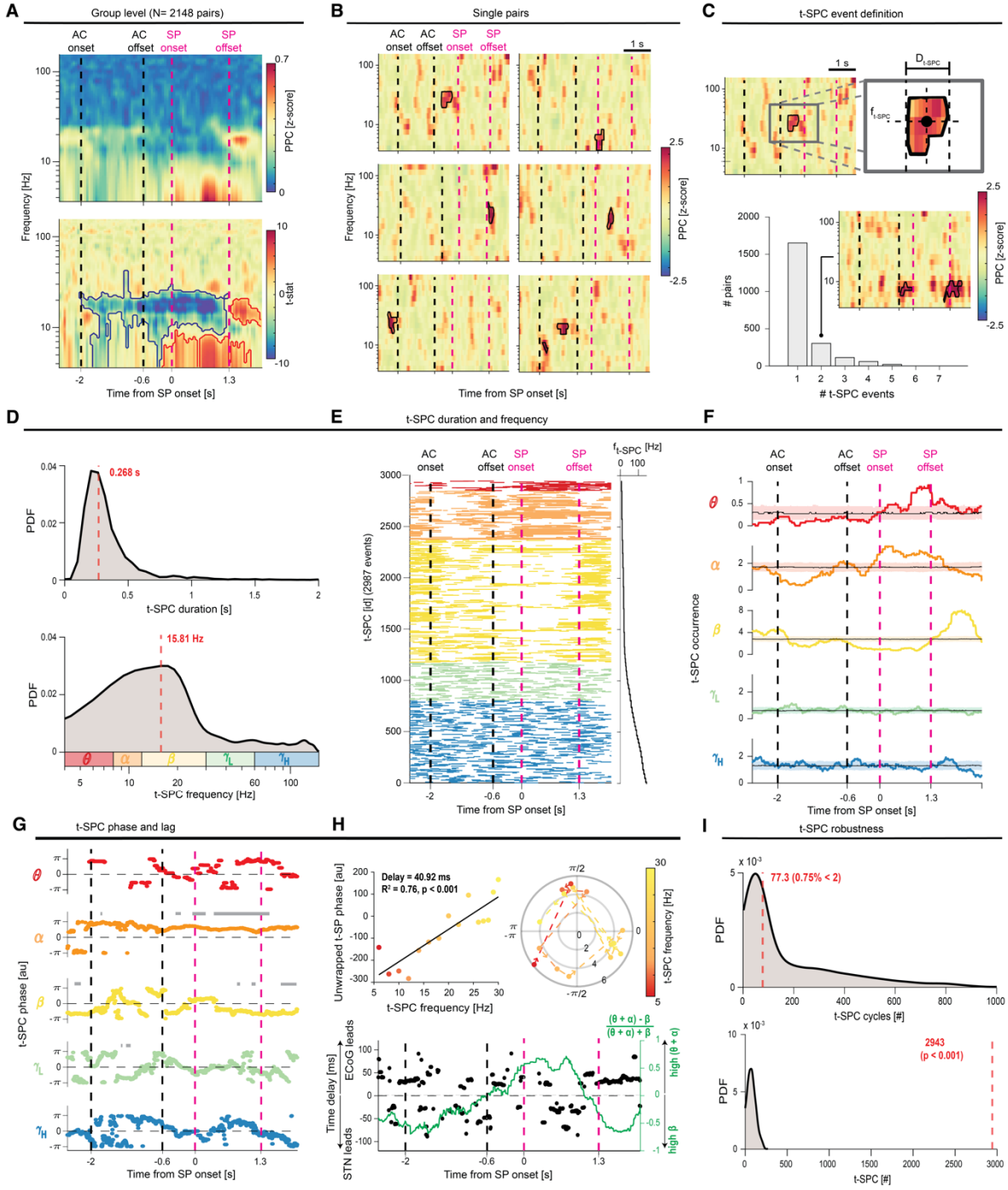
A
Group level (N= 2148 pairs)
AC onset
AC offset
SP onset
SP offset
Frequency [Hz]
PPC [z-score]
t-stat
Time from SP onset [s]
B
Single pairs
1 s
C
t-SPC event definition
pairs
t-SPC events
D
0.268 s
PDF
t-SPC duration [s]
15.81 Hz
t-SPC frequency [Hz]
E
t-SPC duration and frequency
t-SPC [id] (2987 events)
F
t-SPC occurrence
G
t-SPC phase and lag
t-SPC phase [au]
H
Delay = 40.92 ms
R² = 0.76, p < 0.001
Unwrapped t-SP phase [au]
t-SPC frequency [Hz]
Time delay [ms]
ECoG leads
STN leads
high (θ + α)
high β
I
t-SPC robustness
77.3 (0.75% < 2)
t-SPC cycles [#]
2943
(p < 0.001)
t-SPC [#]

Figure 2: STN spike-to-ECoG phase coupling during speech production. (**A**) Average of the spike-phase coupling (SPC) maps with significant spike-phase coupling (N = 2148 pairs). The pairwise-phase consistency (PPC) index is compared to the permutation distribution and expressed as z-score. Group-level statistical test (t-stat) of the significance of the z-score PPC with respect to the baseline across all the significant pairs. Red and blue lines contour regions of significant SPC increase or decrease, respectively. (**B**) Examples of single-pair SPC maps showing that STN neurons preferentially entrain cortical phases only during brief and transitory episodes. (**C)** Definition of the transient SPC event (t-SPC event) in a single-pair SPC map. We calculated the onset and offset times, temporal duration, and the frequency centroid for each t-SPC event. Most of pairs exhibit only one t-SPC event, as shown by the barplot. The inset plot depicts an exemplary SPC map with two t-SPC events in the same frequency band. (**D**) Distribution of the t-SPC duration and t-SPC frequency centroid. To augment the readability of the t-SPC frequency distribution, we adopted a logarithmic scale. (**E**) List of the t-SPC events (N = 2987) ordered by frequency centroid. (**F**) t-SPC events occurrence grouped by frequency band. Shaded areas illustrate the 5th and 95th percentiles of the permutation distribution for the aggregation test. (**G**) Average preferred phase of firing during the t-SPC events across frequency bands. Gray bar depicts time bins with significant non-uniformity of the preferred phase across all pairs. (**H**) Unwrapped phase values show their progression on a linear scale. The time delay derived from the gradient of the linear fit was 40.92 ms. A permutation test confirmed the significance of the relationship between unwrapped phases and frequency. The sign of the time delays defines the leading activity: positive (ECoG leads STN) and negative (STN leads ECoG). Circular plot shows mean angle (position on circle) and strength of SPC (PPC) of pairs that are significantly locked at frequencies from 5 to 30 Hz (as indicated by the color scale). Time-resolved estimation of the time delay superimposed with the relative occurrence of (θ-α) and β t-SPC events. The sign of this quantity determines the leading frequency band: positive ((θ-α) dominates) and negative (β dominated). (**I**) Robustness analysis of the t-SPC event identification. Distribution of the number of cycles that span the duration of the t-SPC events. Distribution of the number of t-SPC events observed in the permuted SPC maps. Red dashed line depicts the median of the distribution. Black and purple dashed lines denote auditory cue and speech production windows. θ (red), α (dark orange), β (yellow), γ_L (green) and γ_H (blue).

Transient spike-phase coupling events are reliably observed

We confirmed that t-SPC events observed in our data were not a by-product of the natural cluster tendency arising in small samples of random distribution. To this end, we evaluated the number of cycles of oscillations spanned by t-SPC events and the count of potential t-SPC events observed in the permuted SPC maps, previously used to convert the SPC maps into z-scores (see Methods). On average, t-SPC events encompassed 77.3 oscillation cycles (only 0.75% with less than 2 cycles) and significantly outnumbered the t-SPC events observed in the permuted SPC maps ($p_{perm} < 0.001$, permutation test) (Figure 2I). This control analysis suggests that t-SPC events reflect genuine SPC mechanisms.

STN-cortical spike-phase coupling during speech is temporally and topographically organized

Previous studies have shown that neural activity in the cortex and STN feature frequency-specific topographies during resting state[29,30] and movement execution[31,32]. We sought to delineate the cortical and STN spatio-frequency distribution of t-SPC events during syllable repetition (Figure 3).

At the cortical level across all frequency bands, most t-SPC events occurred in the postcentral gyrus (PostCG, 17.13% significant pairs, t-SPC density = 25.57, $p_{perm} < 0.05$, permutation test) and supramarginal gyrus (SMG, 13.18% significant pairs, t-SPC density = 18.45, $p_{perm} < 0.05$, permutation

test). Significant numbers of t-SPC events were found also in the superior temporal gyrus (STG, 13.74% significant pairs, t-SPC density = 15.06), precentral gyrus (PreCG, 10.74% significant pairs, t-SPC density = 14.09) and subcentral gyrus (SCG, 11.34% significant pairs, t-SPC density = 15.94), but t-SPC density was less frequent in the middle and inferior frontal areas (Figure 3A).

We observed aggregation (i.e., increased occurrence) of θ t-SPC events during speech production in SMG and STG ($p_{perm} < 0.05$). Similarly, α t-SPC events dispersed (i.e., reduced occurrence) during auditory cue presentation and aggregated during speech production in SMG and STG regions, in addition to other areas such as PreCG and PostCG ($p_{perm} < 0.05$). Most of the α t-SPC events during the ITI were observed in the MFG. Interestingly, α t-SPC events were not significantly spatially clustered around their centroid (x = -62.23 mm, y = -13.04 mm, z = 30.4 mm, $p_{perm} = 1$, Figure S4). β t-SPC events occurred in the ITI and later dispersed temporally from auditory cue presentation through speech production in the PostCG and SCG ($p_{perm} < 0.05$). Interestingly, the β t-SPC rebound was a more wide-spread phenomenon, observed in PostCG and SCG, and also in cortical regions like PreCG, SMG and STG which did not exhibit t-SPC during the ITI ($p_{perm} < 0.05$). β t-SPC events were spatially clustered around their centroid (x = -65.27 mm, y = -10.14 mm, z = 28.41 mm, $p_{perm} < 0.05$, Figure S4). γ_L and γ_H t-SPC events showed no preferential spatio-temporal distribution. We also compared the t-SPC event duration and centroid frequency across different region of interests. Longer t-SPC events with a lower frequency centroid occurred in the PostCG (~320 ms, ~20 Hz) and SMG (~310 ms, ~18 Hz) ($p_{perm} < 0.05$, permutation test) (Figure S5). These results indicate that different epochs of speech perception and production are accompanied by frequency-specific STN-cortical SPC signatures, whose variance was greatest in the antero-posterior plane.

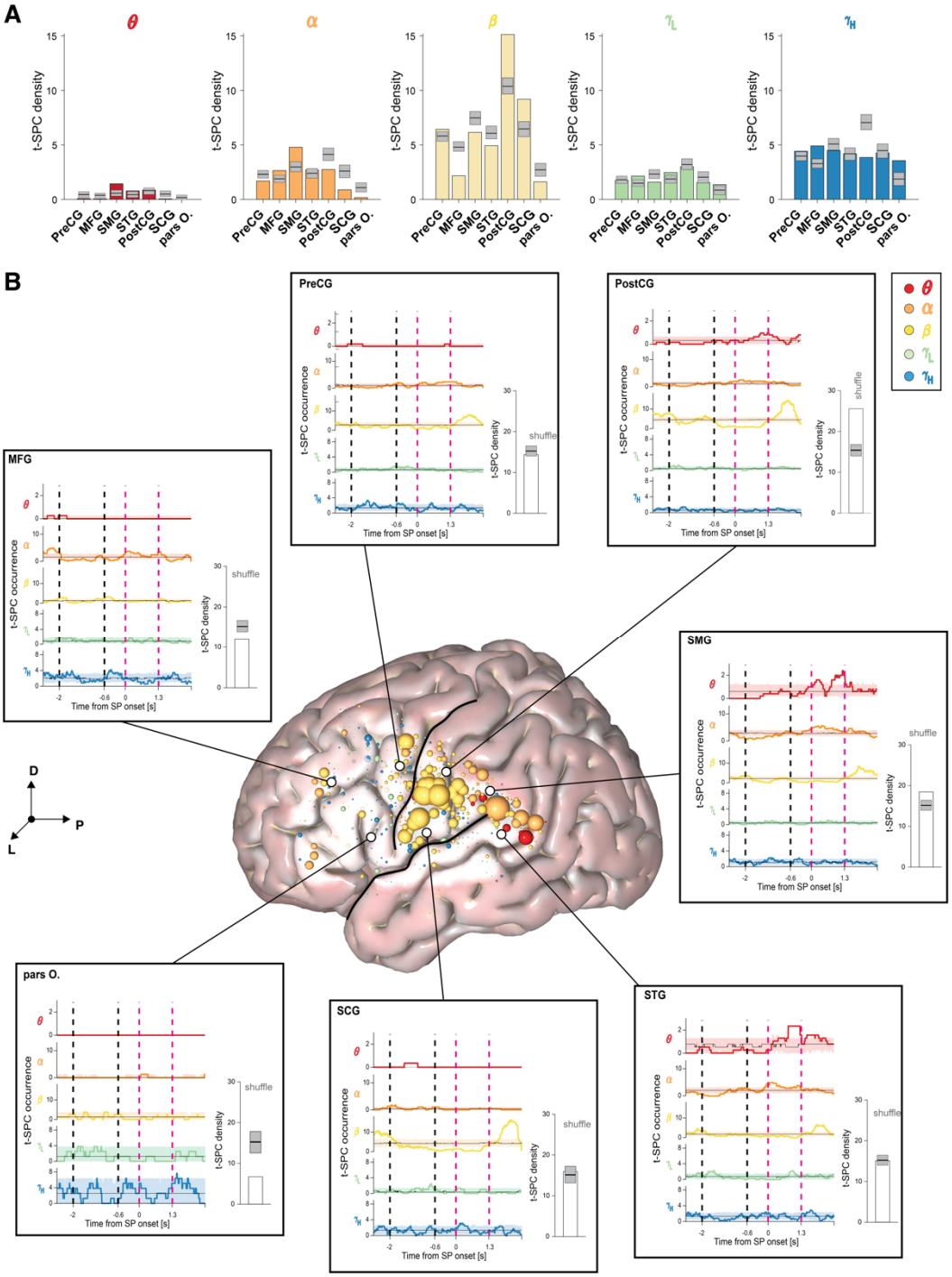
A
θ
α
β
ηL
ηH
t-SPC density
PreCG
MFG
SMG
STG
PostCG
SCG
pars O.
B
PreCG
PostCG
MFG
SMG
pars O.
SCG
STG
t-SPC occurrence
t-SPC density
shuffle
Time from SP onset [s]
D
P
L

Figure 3: Cortical location of the spike-phase coupling. (**A**) Spatial density of the t-SPC events across frequency bands in seven region of interests, as derived from the Destrieux atlas[26]. (**B**) Cortical spatial density map (radius of 2 mm) across frequency bands. The size of the spheres represents the degree to which t-SPC events are localized in a 2mm radius around the center of the spheres. Inset plots illustrate the overall t-SPC spatial density (white barplots) and t-SPC event occurrence in each region of interest. Shaded areas illustrate the 5th and 95th percentiles of the permutation distribution for the aggregation test. Dark gray boxes indicate the 5th and 95th percentile of the permutation distribution for the spatial preference test. Regions with spatial density higher or lower than the permutation distribution are labelled as high or low spatial preference. Black and purple dashed lines denote auditory cue and speech production windows. θ (red), α (dark orange), β (yellow), γ_L (green) and γ_H (blue). Black lines on the cortical surface delineate two anatomical landmarks: the Sylvain fissure (SF), which divides the temporal from the frontal and parietal lobes, and the Central sulcus (CS), which separates the Precentral gyrus (PreCG) anteriorly from the Postcentral gyrus (PostCG) posteriorly. List of cortical regions of interests: Precentral gyrus (PreCG), Postcentral gyrus (PostCG), Supramarginal gyrus (SMG), Subcentral gyrus (SCG), Superior temporal gyrus (STG), Middle frontal gyrus (MFG), and the orbital part of the inferior frontal gyrus (pars O.) The anatomical reference of frame shows the dorsal (D), lateral (L) and posterior (P) directions.

With regard to location of phase-locked STN spiking, θ t-SPC events were significantly aggregated in the posterior-medial region of the STN ($p_{perm} < 0.05$, θ_1 in Figure 4 and Figure S7). Moreover, θ t-SPC events were localized significantly superior (higher MNI z-coordinate) compared to t-SPC events in other frequency bands (Figure S7). Two α t-SPC hotspots ($p_{perm} < 0.05$, Figure 4 and Figure S7) were identified in the posterior-dorsal (α_1) and posterior-ventral (α_2) region of the STN. Of note, MFG SPC exclusively contributed to the posterior-ventral hotspot. Overall, α t-SPC events were localized significantly inferior/ventrally (lower MNI y-coordinate and PC2 coordinate) compared to t-SPC events in other frequency bands (Figure S7). Spatial density analysis in the β range demonstrated a more frequent emergence of β-SPC in the dorsolateral part of the STN during the ITI and the rebound phase ($p_{perm} < 0.05$, β_1 in Figure 4 and Figure S6). β-SPC density appeared more focalized during rebound than during the ITI. A transient increase in β-SPC events during auditory cue presentation was observed in the centro-medial region of the STN.

Table S2 summarizes centroids of t-SPC event location and peak of t-SPC spatial density for each frequency band on the cortex and STN. These results indicate that frequency-specific SPC of speech production are represented in topographically distinct regions within STN.

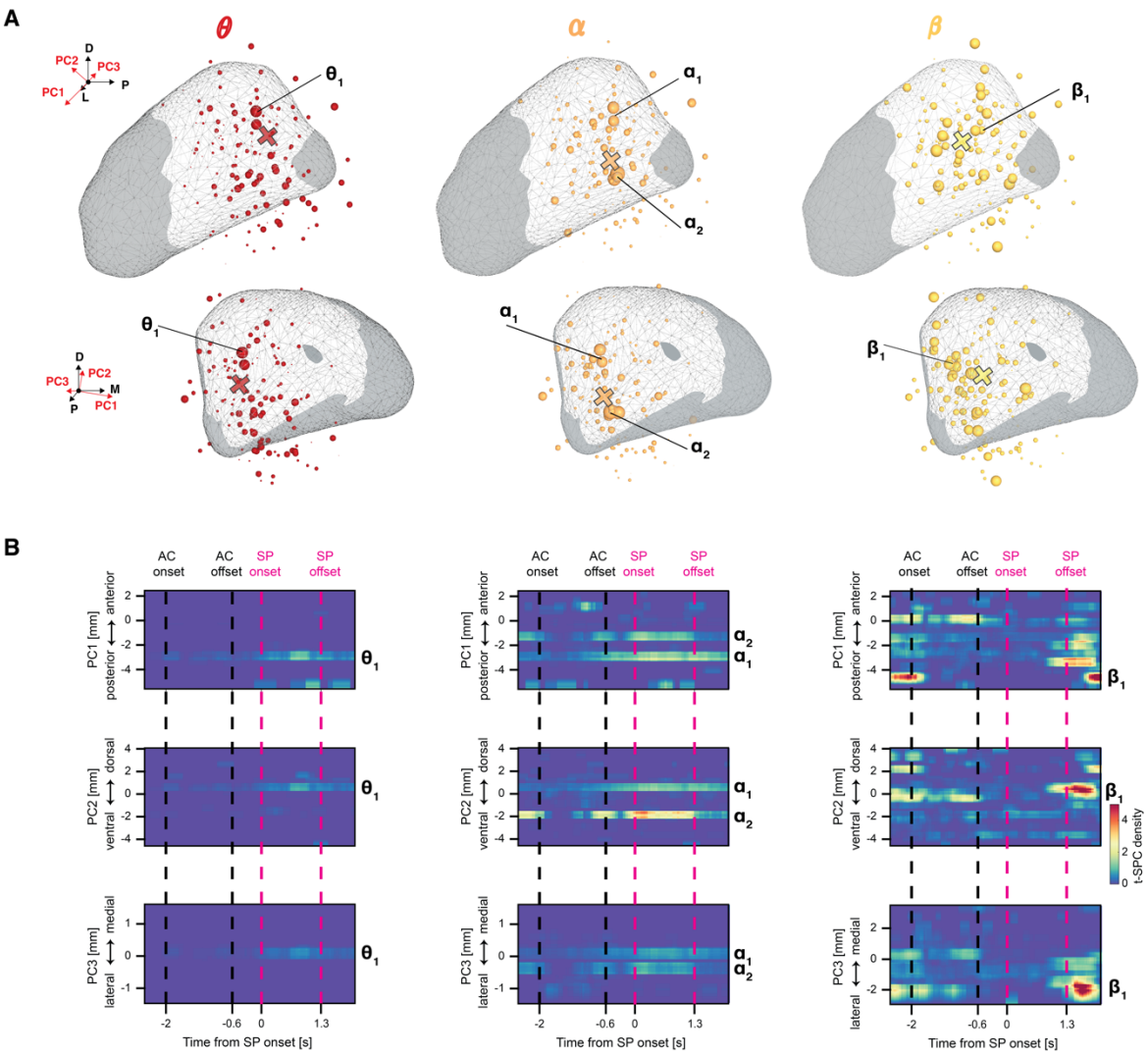

Figure 4: Subthalamic location of the spike-phase coupling. (**A**) Subthalamic spatial density maps (radius of 1 mm) *across frequency bands.* STN regions sampled by microelectrode recordings are depicted as white overlay. Size of the spheres represents the degree to which t-SPC events are localized in a 2 mm radius around the center of each sphere. To augment the readability of the visualization, we adopted the logarithmic scale for the spatial density. (**B**) Spatial density of the t-SPC events mapped along the three principal component axes. The anatomical reference of frame shows the relative orientation between the dorsal (D), lateral (L) and posterior (P) directions and the first three principal components directions (PC1: anterior-posterior axis, PC2: dorso-ventral axis and PC3: medio-lateral). θ (red), α (dark orange) and β (light orange). Black and purple dashed lines denote auditory cue and speech production windows. Cross indicates the spatial centroid of the t-SPC event locations. θ_1, α_1, α_2 and β_1 depict the location of peaks of the t-SPC spatial density. Note that principal component scores represent actual physical distances in mm. θ (red), α (dark orange), β (yellow).

Firing rate modulation predicts preferred spike-phase coupling frequency

We used a variable-window width and pairwise-phase consistency (PPC) correction to ensure that changes in coupling strength were not merely a result of firing rate modulation. However, these two

neural phenomena may represent distinct, yet overlapping, mechanisms of modulation[33,34]. To explore this possibility, we conducted a series of correlation analyses between t-SPC and IFR across neurons (Figure S8). Importantly, we found no significant correlation between average IFR and either average t-SPC density (R^2 = 0.006, p_{perm} = 0.23) or t-SPC centroid frequency (R^2 = 0.006, p_{perm} = 0.27).

To characterize the extent of overlap between changes in t-SPC density and IFR modulation, we plotted SPC density changes (expressed as the difference) against IFR modulation (expressed as z-score) with respect to the ITI (peak IFR modulation) during speech production (see Methods) (Figure S8). We observed (θ-α) t-SPC changes (increase: 31/211 units, 14.7% and decrease: 2/211 units, 0.9%) only in neurons exhibiting low or negative IFR changes (< 5 z-score). Among the 27 neurons that displayed β t-SPC events during the ITI, 25 (92.6%) neurons significantly reduced their β t-SPC during speech production, either completely (22/25) or partially (3/25). Interestingly, a fraction of neurons (16/211, 7.6%) exhibited a slight increase in β t-SPC during speech production, suggesting a partial maintenance of the β SPC at the single-unit level. Importantly, we observed no significant differences in IFR changes between neurons that decreased β t-SPC density during speech production or increased β t-SPC density after speech termination and neurons with no changes in β t-SPC density (Figure S8). Subsequently, we investigated whether individual neurons displayed a preference for a specific SPC frequency range or exhibited no preference at all. Intriguingly, our findings revealed that SPC was highly specific in a frequency band at the single neuron level. Indeed, neurons that increased (θ-α) t-SPC did not exhibit any modulation of β SPC, and vice versa (Figure S9).

Comparing SPC across the different IFR categories, we found that neurons whose firing rates were modulated by speech (18.51%) displayed a higher average t-SPC density compared to neurons without speech-related firing rate modulation (10.86%) (p_{perm} < 0.01, Figure S10). Among the speech-related neurons, those with decreased IFR exhibited the highest average t-SPC density (23.43%), surpassing both increased IFR (14.52%) and mixed IFR modulation neurons (17.59%) (p_{perm} < 0.01). Both group-level SPC and t-SPC analyses indicated that only decreased IFR neurons contributed significantly to the aggregation of θ t-SPC events during speech production (Figure 5). On the other hand, both decreased and increased IFR neurons featured similar α t-SPC event occurrence profiles, occurring less frequently during auditory cue presentation and more frequently during speech production. Notably, increased, and mixed IFR modulation neurons primarily contributed to the aggregation of β t-SPC events during the rebound phase, but only mixed IFR modulation neurons displayed significant aggregation of β t-SPC events during the ITI. In contrast, no discernible pattern of t-SPC coupling emerged in neurons with no IFR modulation across all levels of analysis (Figure 5).

We then examined changes in the centroid frequency and duration of t-SPC events across IFR categories. Longer t-SPC events were observed in mixed IFR neurons (~0.31 s) and decreased IFR

neurons (~0.31 s) with median centroid frequencies of 18 Hz and 20 Hz, respectively (Figure S10). When analyzed by IFR category, only decreased IFR neurons demonstrated a spatial preference for spike-phase coupling, which occurred with the SMG and STG, likely driven by t-SPC in the θ-α range. Together, these results suggest that while firing rate modulation alone does not fully explain the dynamics of SPC, the direction of firing rate change is correlated with distinctive patterns of SPC characterized by spectral, spatiotemporal, and anatomical features.

Distinct physiological fingerprints of pre- and post- speech production β spike-phase coupling

Prior studies have demonstrated that sensorimotor β power preceding and following a movement exhibits distinct spatiotemporal and anatomical characteristics [35]. Building upon this knowledge, we sought to determine whether β t-SPC events during the ITI and the rebound phase during speech similarly exhibit distinct functional roles (Figure 6). First, we found that the proportion of neurons displaying overlapping β t-SPC during the ITI and the rebound phase was significantly lower than what would be expected by chance across all IFR neuron categories (Figure 6A-C, all $p_{perm} > 0.19$). This result implies that neurons that engage in β t-SPC rebound at speech termination tended not to have significant β t-SPC during the ITI. Second, at the cortical level we observed that the reduction of β t-SPC during speech production was primarily localized more anterior and ventral compared to the rebound of β t-SPC after speech termination (Figure 6D). Similarly, a comparable pattern of segregation was observed in the STN, where the reduction of β t-SPC during speech production was localized more medially than the rebound of β t-SPC after speech termination in (Figure 6E). These results describe fingerprints of β t-SPC preceding and following speech production that are functionally distinct and anatomically segregated processes.

Spike-phase coupling strength predicts phonetic accuracy and speech reaction time

To assess the behavioral implications of our findings, we investigated the relationship between phonetic accuracy and strength of SPC. We categorized trials into correct and error trials based on phonetic accuracy and computed SPC for each subset (see Figure 7 A-B). Our analysis unveiled that error trials exhibited lower θ-α SPC preceding speech production, followed by an increase in θ-α SPC after the speech termination (Figure 7A-B). This trend held true whether we considered the SPC map (Figure 7A) or the t-SPC occurrence (Figure 7B) as a measure of SPC strength ($p_{perm} < 0.05$, permutation test). We further assessed the predictive power of θ-α SPC before speech production on phonetic accuracy using a linear mixed-effect model. A likelihood ratio test demonstrated a significant positive covariation between θ-α SPC within [-1.5 to 0 s] before speech onset and phonetic accuracy on a pair-by-pair basis ($\chi^2 = 10.49$, $p_{perm} < 0.001$) (Figure 7 E).

Additionally, we considered the reaction time for each trial of triplet syllables, reflecting the processing time required for articulating the intended speech command. Analysis revealed that slow trials exhibited a delayed desynchronization of β SPC as evidenced by an increase of SPC strength during the ITI and

early stages of auditory cue presentation (Figure 7C-D). A likelihood ratio test showed a significant positive correlation between β SPC within [-2.5 to -1 s] before speech onset and reaction time on a pair-by-pair basis ($\chi^2 = 6.21$, $p_{perm} < 0.001$) (Figure 7 F).

In summary, the SPC strength provided frequency-specific insights into speech behavior. Notably, this effect diminished during speech production, indicating comparable SPC strength regardless of phonetic accuracy or reaction times leading up to speech production.

Control analyses

To ensure the robustness and validity of our SPC estimates, we conducted a comprehensive set of control analyses (Figure S11-S12). First, we examined the impact of removing the event-related component of the ECoG signal before the computation of the SPC metric. After eliminating this component, the profile of the SPC maps remained largely unchanged, even at lower frequencies (Figure S11 A-B). Indeed, we found no evidence of phase-reset of oscillations at the onset of auditory cue and speech production (Figure S11C). These findings suggest that the event-locked components (trial-averaged speech-locked signals) did not significantly influence the observed SPC pattern. Second, we assessed the influence of periods with high (top 10th percentile) or low (low 10th percentile) oscillation amplitude. When excluding these specific periods, the results remained comparable (Figure S12). These control analyses strengthen the validity of our findings and reinforce the significance of the observed SPC patterns.

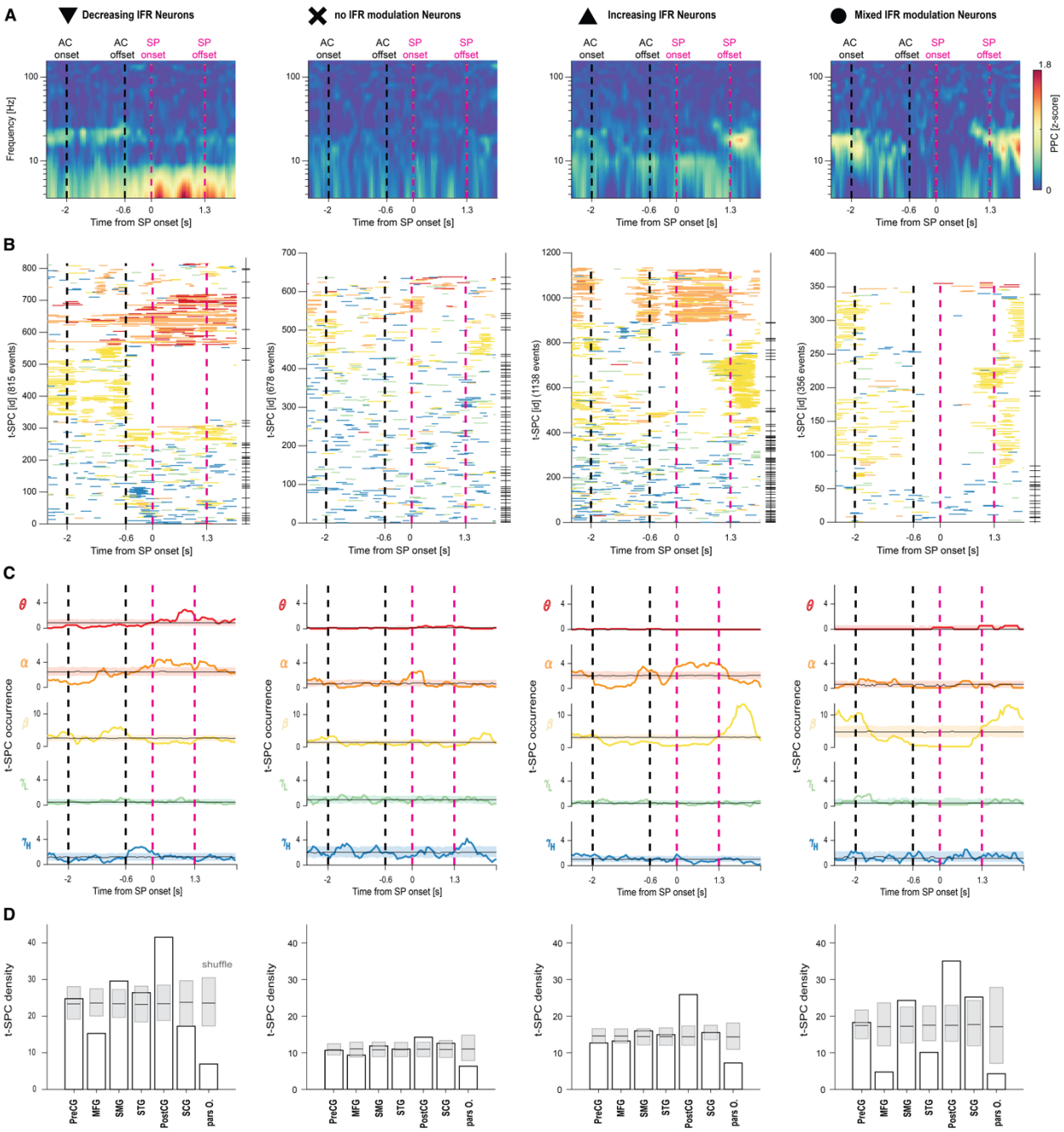

Figure 5: Spike-phase coupling changes in different instantaneous firing rate categories of neurons. (**A**) Average of the spike-phase coupling (SPC) maps for significant spike-phase coupling pairs (N = 2148). The pairwise-phase consistency (PPC) index is compared to the permutation distribution and expressed as z-score. (**B**) List of the t-SPC events (N = 2987) sorted by increasing frequency centroid. Horizontal ticks indicate different neurons. (**C**) t-SPC event occurrence grouped by frequency band. Shaded areas illustrate the 5th and 95th percentiles of the permutation distribution for the aggregation test. (**D**) Overall spatial density of the t-SPC events in seven regions of interest, as derived from the Destrieux atlas[26]. Dark gray boxes indicate the 5th and 95th percentile of the permutation distribution for the spatial preference test. Regions with spatial density higher or lower than the permutation distribution are labelled as high or low spatial preference. Black and purple dashed lines denote auditory cue and speech production windows. θ (red), α (dark orange), β (yellow), γ_L (green) and γ_H (blue).

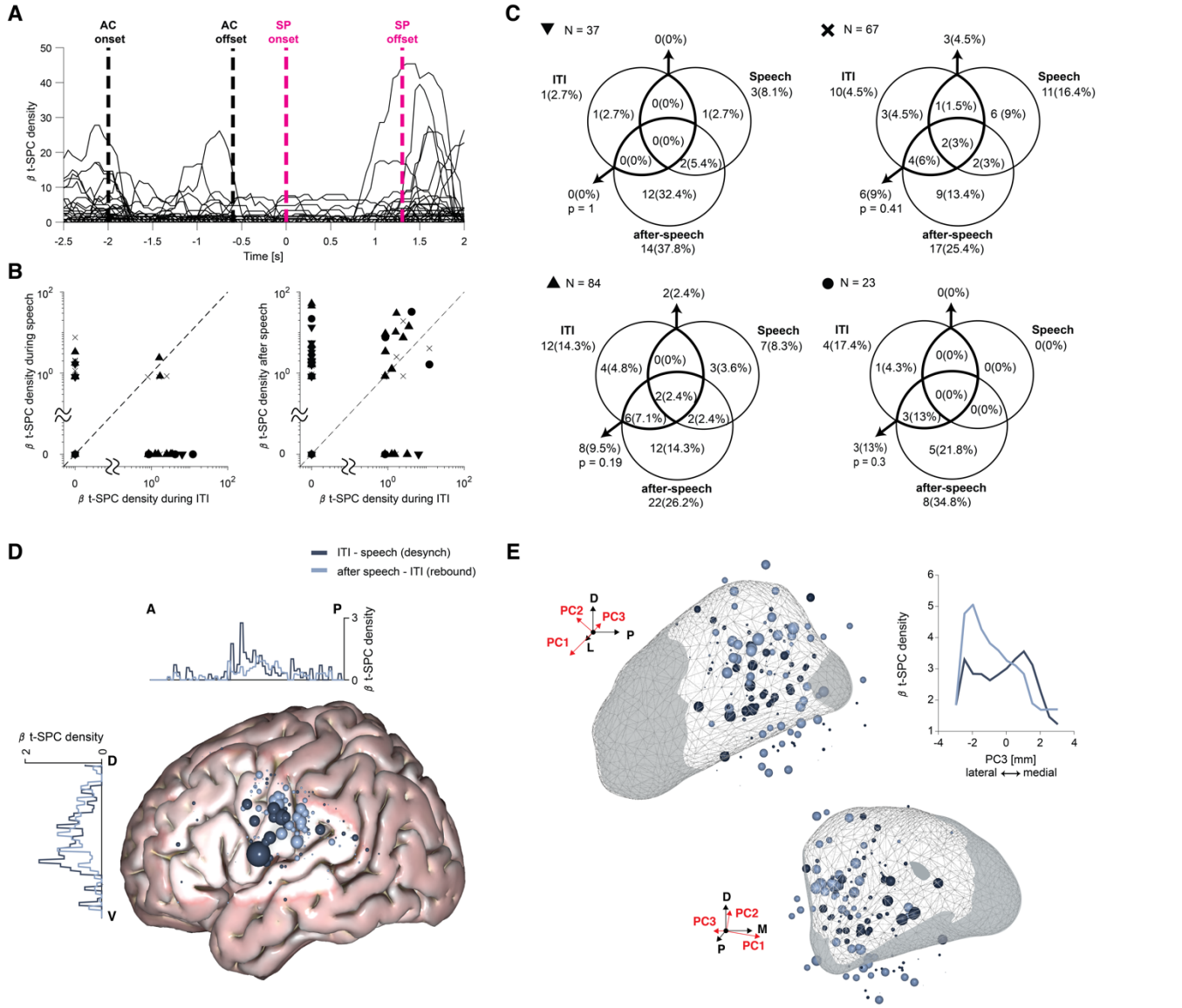

Figure 6: Pre- and post- speech production β spike-phase coupling show distinct fingerprints. (**A**) β-tSPC spatial density of each neuron over time. Each line represents a neuron (N = 211). (**B**) Comparison between β-tSPC spatial density during ITI, speech production and after speech termination. Marker shape indicates the four categories of neurons: Decreasing IFR neurons (downward triangle), no IFR modulation neurons (cross), Increasing IFR neurons (upward triangle) and Mixed IFR modulation neurons (circle). Black dashed line depicts the identity line. (**C**) Venn diagrams showing overlap of β-tSPC during ITI, speech production and after speech termination at the neuron level. The numbers within the circles indicate neurons showing one or more t-SPC events during the behavioral windows. The number outside of the circles indicate the total number of neurons in each category. P-value of the permutation test assesses whether overlap between neurons with ITI and after speech termination t-SPC events is significant. (**D**) Cortical map of the decrease (dark gray) in the β-tSPC spatial density during speech production and increase (rebound) in the β-tSPC spatial density after speech termination with respect to the baseline. Inset plots show the spatial density changes mapped along the dorso (D) – ventral (V) and anterio (A) – posterior (P) axes. (**E**) Subthalamic map of the decrease (dark gray) in the β-tSPC spatial density during speech production and increase (rebound) in the β-tSPC spatial density after speech termination with respect to the baseline. Inset plots show the spatial density changes mapped along the third principal component (PC3). The anatomical reference of frame shows the relative orientation between the dorsal (D), lateral (L) and posterior (P) directions and the first three principal components directions (PC1: anterior-posterior axis, PC2: dorso-ventral axis and PC3: medio-lateral). Note that principal component scores represent actual physical distances in mm. To augment the readability of the visualization, we adopted the logarithmic scale for the spatial density in panels (**B**) and (**E**).

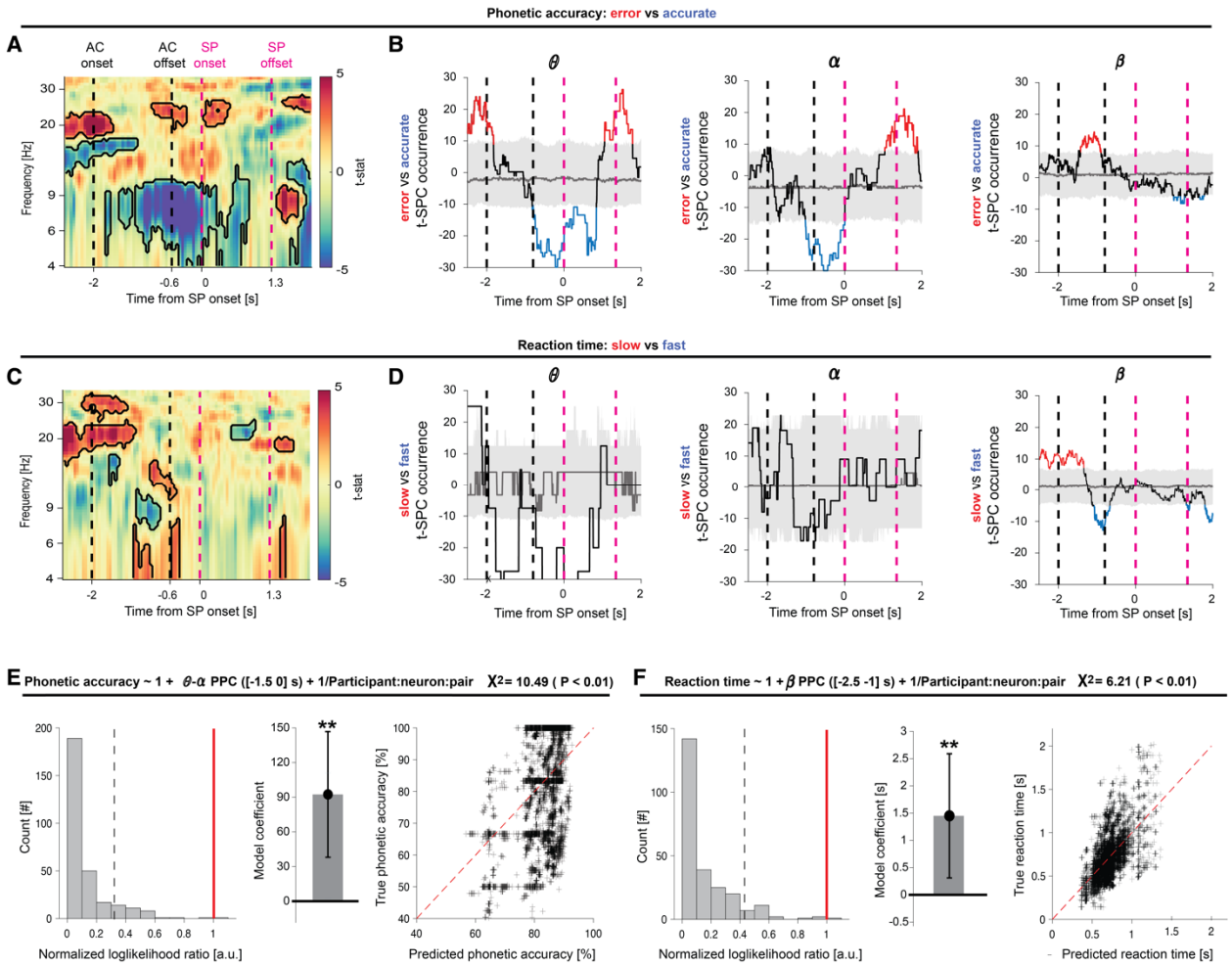

Figure 7: Spike-phase coupling predicts speech-related behavior. (**A**) Comparison of the spike-phase coupling (SPC) maps between trials with and without phonetic errors (see Methods). In error trials, θ spike-phase coupling before speech production onset was significantly lower than in correct trials. (**B**) Relative t-SPC events occurrence in trials with and without phonetic errors grouped by frequency band. An aggregation test identifies when t-SPC events are statistically more aggregated in error (positive, red) or correct (negative, blue) trials. Shaded areas illustrate the 5th and 95th percentiles of the permutation distribution for the aggregation test. (**C**) Comparison of SPC maps between trials with slow and fast reaction times, as determined by median splitting. In slow trials, β SPC during baseline and auditory cue presentation was significantly higher than in faster trials. (**D**) Same as (**B**) between slow and fast reaction times. An aggregation test identifies when t-SPC events are statistically more aggregated in slow (positive, red) or fast (negative, blue) trials. Black and purple dashed lines denote auditory cue (AC) and speech production (SP) windows. (**E**) A linear mixed model between θ-α SPC preceding speech production ([-1.5; 0] s window) and phonetic accuracy shows a significant relationship, as assessed by a permutation test of likelihood ratio (χ^2 = 10.49, p_{perm} < 0.001). (Left) Histogram of the normalized loglikelihood ratio values in the permutation set. Dashed vertical line depicts the 95th percentile of the distribution. All likelihood ratio values are normalized with respect to the likelihood ratio of the model (red vertical line at 1 by definition). (Middle) Regression coefficients for the fixed effect of PPC. Error bars represent 95% confidence interval. (Right) Phonetic accuracy predicted by the model is plotted against the true phonetic accuracy. Each cross represents a pair. Red dashed line depicts the identity line. (**F**) Same as (**E**) with a linear mixed model between β SPC preceding speech production ([-2.5 1] s window) and reaction time. A permutation test of likelihood ratio (χ^2 = 6.21, p_{perm} < 0.001) identifies a significant positive relationship between PPC and reaction time.

Discussion

In this study, we found robust evidence of transient (~268 ms) spike-phase coupling events that occur between STN neurons and broad cortical regions during a syllable triplet repetition task. We found that cortical potentials and subthalamic neurons exhibit complex dynamics during speech, where STN neuronal firing becomes phase-coupled to cortex in a frequency-specific manner that is organized both, temporally and topographically (Figure 8). Interestingly, θ and α SPC events dominated within STG and SMG during speech, while distinct physiological fingerprints of pre- and post- speech production β were observed in SPC. While changes in coupling strength did not arise from modulation of the firing rate, we determined that the direction of firing rate change (increase or decrease) during speech predicted the preferred spike-phase coupling frequency with cortex. The strength of SPC was predictive of participants' ability in producing the desired phoneme and the time taken for articulation.

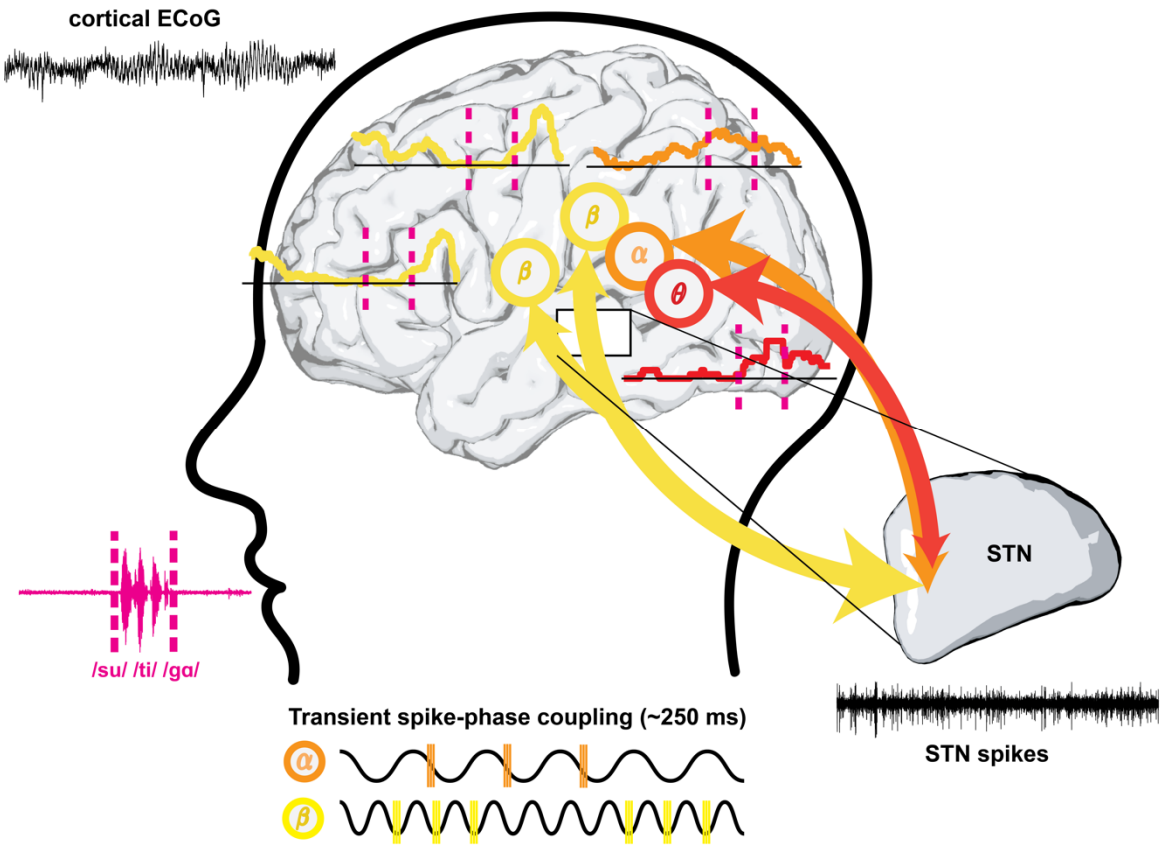

Figure 8. Proposed spike-phase coupling mechanism of speech production. STN engages in transient events of spike-phase coupling with broad cortical regions during the production of syllable triplets. During speech production, β spike-phase coupling is suppressed in the subcentral gyrus and rebounds after speech termination in the postcentral gyrus. θ and α spike-phase coupling in the supramarginal gyrus and superior temporal gyrus is elevated during speech production. The smaller arrowheads indicate which location leads the coupling. θ (red), α (dark orange), β (yellow), γ_L (green) and γ_H (blue).

Our results align with the notion of the cortico-basal ganglia thalamic loop serving as a temporal-integration site for modulating motor control [36–38], consistent with precious evidence of SPC between STN neurons and cortical field potentials during limb movement [25,33,34,39,40]. Similarly, cortical oscillations manifest as transient bursts, whose onset is preceded by an increase of spike-phase coupling with the STN[41]. These transient periods might represent “open windows” for effective communication between STN and cortex. The duration of these windows may be constrained by a slower subcortical neural timescale and requirements of a given motor instantiation[41,42].

Our observations suggest that θ-SPC, driven by the STN in the STG and SMG, may facilitate the online monitoring of auditory signals for speech production. This mirrors the dynamics seen in overt speech perception[43], where the phase alignment of STG theta oscillations efficiently encodes acoustic features, parses syllabic boundaries, dictates the syllabic rhythm and integrates multi-level phonemic encoding [43–45]. It is plausible that STN neurons may synchronize the timing of STG and SMG theta oscillations to organize phonological and prosodic representations, aligning them with the current auditory target during speech production. This theta SPC might represent a complementary mechanism adding to previous reports in the literature that observed gamma amplitude changes at the level of STG and SMG [21,46]. Our results further corroborate the growing body of evidence highlighting the interaction between STG and STN [9,16]. Event-related causality analyses have shown that STN and STG may also be involved in the feedforward control of speech, as evidenced by the propagation of a lexical-encoding gamma signal from STG to STN prior to speech production[16].

While alpha and theta SPC exhibited similar dynamics (Figure 2-3), we identified alpha SPC in other cortical regions, including the PostCG and PreCG gyrus, in addition to the STG and SMG. This hints at the role of alpha SPC in integrating information from both sensorimotor and auditory processing. We therefore posit that strong phase alignment across neurons in this range could support this mechanism. Supporting this hypothesis, theta-alpha SPC was observed preceding speech onset when participants accurately pronounced the target phoneme, contrasting with trials involving phonetic errors, and predicted the level of phonetic accuracy. Both theta and alpha SPC localize to cortical regions connected through the superior longitudinal fasciculus, also referred to as the “dorsal stream”, which is implicated in feedback processing during speech production[47].

The role of beta SPC in speech is intriguing. Beta oscillations in the sensorimotor cortex desynchronize before and during movement (event-related desynchronization) and resynchronize after its termination (post-movement rebound) [48]. These patterns suggest a role for beta as an active process that promotes the existing state “status quo” of the sensorimotor cortex. In contrast, beta desynchronization is thought to disrupt the idle state, with the subsequent rebound mediating the return-to-baseline following sensorimotor engagement [49]. Other research posits that the temporal integration of somatosensory afferent and motor efferent signals plays a key role in the genesis of the beta rebound signal [50]. Our

findings support a similar interpretation, albeit focusing on beta SPC instead of beta oscillations. STN neurons may gate sensorimotor information by intermittently synchronizing with beta cortical oscillations[51], and hence, delaying the time required for the articulation. The level of beta SPC during ITI and auditory cue presentation was found to be predictive of the reaction time. Moreover, our results suggest a functional segregation of these two beta phenomena implemented at the single-unit level. Cortical regions with reduction of beta SPC during speech production were found more anteriorly and ventrally than those that show rebound in beta-SPC. We found a similar gradient in the medio-lateral axis of the STN. These variations suggest the presence of two distinct functional beta networks that control and gate the temporal integration of sensorimotor information in the cortex [35,51,52].

At the neuron level, changes in spike-phase coupling during speech production were not correlated with speech-related changes in the instantaneous firing rate or the baseline instantaneous firing rate. Other studies have reported similar decoupling between SPC and firing rate in STN neurons during movement [25,39]. Moreover, we found that neurons fired at comparable intensity, irrespective of their significant SPC display, except for theta SPC events. These were predominantly driven by neurons that decreased their firing rate in the early stages of speech production. Interestingly, we previously found that neurons often reducing their firing rates during speech did so in association with the onset of the auditory cue [14]. Our findings underscore the notion that information can traverse the cortico-basal ganglia loop either though changes in the firing rate activity or spike timing. We also found that the pronounced overall reduction of beta SPC observed at the population level during speech production did not reflect a uniform reduction at the single-unit level. A small subset of neurons (7.6%) increased their beta SPC during speech production, suggesting a partial maintained beta SPC and the presence of a distinct functional beta SPC network. This aligns with two other studies that reported similar subpopulations of STN neurons, which increased their beta SPC during motor activity [25,33]. The functional relevance of this partially maintained beta SPC during speech production remains uncertain.

Consistent with other studies, we demonstrated that cortical oscillations lead in phase over firing activity in the STN [25,33,39], especially in the beta range post-speech. We found a delay of ~41 ms aligning closely with other reports [25,33]. Notably, the STN led the cortex exclusively when SPC in theta and alpha was more pronounced during speech production. While this delay does not necessarily signify synaptic transmission delay, it is consistent with the transmission of information through the indirect cortico-basal ganglia loop. In light of our recent findings suggesting the presence of monosynaptic connections between non-motor regions of the cortex (sensorimotor and auditory areas) and the STN [9], the SPC between STN and non-motor cortical regions may be anatomically rooted in the hyperdirect pathway.

Theoretical models have posited a spatial scaling role for basal ganglia [11], which mediates information processing and spatial integration [53]. The deep placement of the STN represents a suitable integration hub to process multimodal cortical information. Our results, which highlight preserved frequency-wise SPC topographies from cortex to STN align well with this theory and confirm a spatial scaling of the

information flow. However, caution is necessary when interpreting SPC results, as the presence of SPC between STN neurons and narrowband cortical oscillations does not imply that STN neurons generate and resonate coherent rhythms with the cortex [54]. For example, neurons displaying significant theta SPC with the STG do not necessarily oscillate in the theta-rhythm at the population level. Therefore, our frequency-wise STN SPC topographies during speech production would not necessarily align with STN power-based topographies based on local field potentials recorded at rest [30,55].

Our results have important implications for theoretical frameworks of both basal ganglia function and speech production. Current models, such as DIVA[5] and SFC[6], are predominantly cortico-centric and do not explicitly incorporate basal ganglia nodes. In the context of the DIVA model, our analysis suggests a role for the STN in online monitoring of both auditory and sensory-motor states during speech production. Two feedback controllers, the auditory controller and the somatosensory controller, are included in the DIVA model. Theta SPC might participate in the auditory controller (auditory state, auditory target, and auditory error maps) in the STG, while beta SPC could mediate the sensorimotor controller (sensorimotor state, sensorimotor target and sensorimotor error maps) in the PCG and SMG. Interestingly, alpha SPC might subserve the integration of the two controllers in the SMG. More broadly, these findings clearly illustrate how classic firing-rate based models are insufficient for explaining basal ganglia circuit behavior.

Beyond theoretical frameworks, our results may have important implications for clinical therapies. Although many Parkinsonian motor symptoms can often be satisfactorily controlled by STN DBS, stimulation-induced effects on the speech motor system can be heterogeneous [56]. How can stimulating the same target nucleus consistently ameliorate some parkinsonian symptoms yet have mixed and variable effects on the speech motor system? Our results align with the notion that variability in DBS lead placement can explain most of the reported variance of outcomes in the literature [29]. Relative to the optimal therapeutic target defined by Caire et al [57](x = -12.6 mm, y = -13.4 mm, z = -5.9 mm), the spatial centroid of our speech-related STN SPC topographies is at least 2.5 mm distant and overall located more posterior and ventral (Table S2). This aligns with studies that found detrimental effects on speech outcomes when stimulating more posteriorly [18,58–60] and ventrally[61]. However, all these studies simply compare the speech outcomes between DBS ON and DBS OFF conditions without considering the stimulation amplitude and the spread of the stimulation volume towards neighboring regions. One hypothesis for future investigation is that stimulating the STN in areas of SPC density peaks would alter the STN activity and hamper integration of sensorimotor and auditory signals at the cortical level, leading to speech effects.

Our findings should be interpreted in the light of several limitations. First, a recent study suggested that spike-phase coupling does not mean that STN neurons generate and resonate rhythms coherent with the cortical synaptic input activity or that cortico-subcortical coherence has a functional effect on information transmission gain [54]. Consequently, spike-phase coupling may simply reflect the fact that STN neurons

integrate cortical inputs to some extent. Analysis of spike-phase coupling in both directions and causal investigations are necessary to circumvent these theoretical limitations [54,62]. Second, phase-based time delays do not necessarily reflect synaptic transmission delays [63], but likely reflect the combination of cortical and subcortical inputs to the STN that become transiently stable [33,40]. Third, ECoG coverage varied across participants and spanned a limited region of the cortical surface. However, we were able to perform group-level analysis by considering only regions of interest with adequate coverage. We cannot rule out any other interaction of the STN with other cortical regions. Fourth, we were unable to record single-unit activity across the entire volume of the STN due to unequal spatial sampling. Most microelectrode trajectories traversed the dorsolateral part of the STN, the clinical target for PD DBS [29]. Hence, sampling of the ventro-medial region of the STN is limited. Differences in cortical coverage and subcortical sampling confounded correlative analyses between neural signals and clinical scales. Fifth, our findings are based on Parkinson's Disease patients and caution must be exercised when interpretations of human neurophysiology are drawn from observations collected in a pathological state. Specifically, differences in the STN baseline firing rate [64] and abnormal subcortical beta oscillations [65–68] that characterize the Parkinsonian state may confound distinction of whether our observations generalize to speech in individuals without PD. Sixth, warping participant-specific coordinates into the MNI template space for group-level visualization and statistics may have introduced slight anatomical discrepancies.

In summary, we discovered evidence that the STN participates in the auditory-sensorimotor processing of speech production. We identified topographical changes of spike-phase coupling as a function of frequency at the cortical level, which were compressed and preserved in the STN. These insights provide a deeper understanding of how different types of information are processed in the basal ganglia during speech and have significant implications for understanding basal ganglia function related to sensorimotor integration in human behavior[69].

Methods

Participants

Electrophysiological signals were recorded intraoperatively from 24 participants (20 males and 4 females, age: 65.4 ± 7.1 years) with Parkinson's Disease undergoing awake stereotactic neurosurgery for implantation of DBS electrodes in the subthalamic nucleus (Table S1 for clinical details). Participants performed up to 4 sessions of the task, leading to a total of 64 sessions, after overnight dopaminergic medication withdrawal. All procedures were approved by the University of Pittsburgh Institutional Review Board (IRB Protocol #PRO13110420) and all participants provided informed consent to participate in the study.

Method details

Speech production task – Participants were tasked to intraoperatively repeat aloud consonant-vowel (CV) syllable triplets. The stimuli were presented auditorily via earphones (Etymotic ER-4 with ER38-14F Foam Eartips) and were delivered at either low (~50 dB SPL) or high (~70 dB SPL) volume using BCI2000 as stimulus presentation software. The absolute intensity was tailored to each participant's comfort level, keeping fixed the difference between high and low conditions at 25 dB SPL. The experiment utilized a set of phonemes consisting of four consonants (/v/, /t/, /s/, /g/) with different manners of articulation and three cardinal vowels (/i/, /a/, /u/) with distinctive acoustic properties. We created a unique set of 120 triplets of CV syllables, forbidding CV repetition within the triplet and balancing syllables and phoneme occurrence, and CV position within the triplet across a run of the task. The audio produced by the participant was recorded with an PRM1 Microphone (PreSonus Audio Electronics Inc., Baton Rouge, LA, USA) at 96kHz using the Zoom-H6 portable audio recorder (Zoom Corp., Hauppauge, NY, USA).

Neural recordings – As part of the standard DBS clinical procedure, functional mapping of the STN was performed using microelectrode recordings (MER) acquired with the Neuro-Omega recording system (Alpha-Omega Engineering, Nof HaGalil, Israel) using parylene insulated tungsten microelectrodes (25 µm in diameter, 100 µm in length). The microelectrodes were oriented using three trajectories (Central, Posterior and Medial) of a standard cross-shaped Ben-Gun array with a 2 mm center-to-center shaping. MER signals were referenced to the metal screw holding one of the guide cannulas used to carry the microelectrodes and recorded at 44 KHz. Prior to STN mapping, participants were temporarily implanted with two high-density subdural electrocorticography (ECoG) strips consisting of 54 or 63 contacts, respectively (PMT Cortact). These strips were placed through the standard burr hole, targeting the left ventral sensorimotor cortex, and left inferior frontal gyrus. Signals from ECoG contacts were referenced to a sterile stainless-steel subdermal needle electrode placed on the scalp and acquired at 30kHz with a Grapevine Neural Interface Processor equipped with Micro2 Front Ends (Ripple LLC, Salt Lake City, UT, USA).

Electrode localization – We localized the ECoG strips and DBS leads using well-established pipelines in the literature. For ECoG strips, contact locations were determined using the Randazzo localization method[70] that utilizes a preoperative T1 weighted MRI scan, an intraoperative fluoroscopy and a postoperative CT scan (github.com/Brain-Modulation-Lab/ECoG_localization). CT and MRI were coregistered using SPM (https://www.fil.ion.ucl.ac.uk/spm/) and then rendered into a 3D skull and brain using Osirix (www.osirix-viewer.com) and Freesurfer (https://surfer.nmr.mgh.harvard.edu) software. The position of the frame's tips on the skull and the implanted DBS leads were used as fiducial markers, which were coregistered and aligned with the projection observed in the fluoroscopy. The position of the contacts in the ECoG strip were manually marked on the fluoroscopy image and then projected to the convex hull of the cortical surface. To extract native coordinates of individual contacts, we leveraged the known layout of the ECoG strip. All coordinates were then transformed into the ICBM MNI152 Non-Linear Asymmetric 2009b space, employing the Symmetric Diffeomorphism algorithm implemented in Advanced Normalization Tools (ATNs). For DBS lead reconstruction, we used the Lead-DBS localization pipeline[71]. Briefly, the process involved coregistering the MRI and CT scans, and manually identifying the position of individual contacts based on the CT artifact, constrained by the geometry of the DBS lead used. The coordinates for the leads in each participant's native space were rendered after this process. Custom Matlab scripts (github.com/Brain-Modulation-Lab/Lead_MER) were then used to calculate the position of the micro- and macro-recordings from the functional mapping based on the position of the lead, the known depth, and tract along which the lead was implanted in each hemisphere. Anatomical labels were assigned to each contact based on the Destrieux atlas[72] for cortical contacts, and the DISTAL atlas[73] for subcortical contacts.

Quantification and statistical analysis

Phonetic coding – To extract phoneme characteristics from the produced speech signals such as onset and offset times, IPA code and accuracy, we employed a custom Matlab GUI (github.com/Brain-Modulation-Lab/SpeechCodingApp). Phonetic coding of each produced phoneme was performed by a trained team of speech pathology students using Praat (https://www.fon.hum.uva.nl/praat/).

Behavioral events – For each trial, we defined four different behavioral epochs: baseline epoch as a 500 ms time window between -550 ms and -50 ms prior to the auditory cue onset, auditory cue presentation as the window during which syllable triplets were presented auditorily (~1.5 s duration), speech production as the variable time window during which participants repeated aloud the syllable triplet (~1.6 s duration on average) and post-speech as the 500 ms time window after the speech offset. We defined the reaction time as the time interval between auditory cue offset and speech production onset.

Electrophysiological data preprocessing – ECoG preprocessing was performed using custom code based on the Fieldtrip toolbox[74] implemented in Matlab, available at (github.com/Brain-Modulation-

Lab/bml). To temporarily align recordings from the Ripple, Neuro-Omega and Zoom-H6 systems, we employed a linear time-warping algorithm based on the stimulus and produced audio channels. The alignment was continuous throughout the entire recording session with sub-millimetric precision. Data was low pass filtered at 250Hz using a 4th order Butterworth filter, downsampled to 1KHz and stored as a Fieldtrip object. Metadata such as descriptions of each session, phonetic coding, event times and electrode locations were stored in annotation tables. We applied a 5th order high-pass Butterworth filter at 1 Hz to remove drifts and low-frequency components. Segments with conspicuous high-power artifacts were identified using an automatic data cleaning procedure[75], based on a power-based threshold. Specifically, we extracted power at frequencies in different canonical bands (3 Hz for δ, 6 Hz for θ, 10 Hz for α, 21 Hz for β, 45 Hz for γ_L and 160 Hz for γ_H) by convolving ECoG signals with a 9-cycles Morlet wavelet. A time bin was classified as artifactual if its log-transformed power in any band exceeded a threshold defined as the mean $\pm$ 2.5 std (~10-fold higher the mean). Trials with time segments flagged as artifactual were discarded and channels with more than 30% of artifactual time bins were not included in the analysis.

Spike sorting – Spike sorting was performed using Plexon (https://plexon.com/products/offline-sorter/) as previously described[14]. We used a 4th order Butterworth high-pass filter with a cutoff frequency at 200 Hz and set a manual threshold to extract putative waveforms. Single units were discriminated and graded based on factors such as cluster isolation in the principal component, the spike sorting's stability over time, a refractory period of at least 3 ms in the inter-spike interval distribution, and the shape of the waveform.

Instantaneous firing rate – To analyze changes in spike rate activity, we followed the procedure described in [14]. We sought elevated and reduced firing activity by computing the instantaneous firing rate (gaussian kernel, σ = 25 ms) and the inter-spike interval (smoothing window 25 ms), which scales with the reciprocal of the instantaneous firing rate, respectively. We aligned these quantities with speech production onset and analyzed time bins from auditory cue onset through speech production offset. A neuron was considered as Decreasing IFR neuron if the inter-spike interval exceeded for at least 100 ms the upper 5% of a normal distribution with mean and standard deviation calculated during the baseline period. Similarly, a neuron was considered as Increasing IFR neuron if the instantaneous firing rate exceeded for at least 100 ms the upper 5% of a normal distribution with mean and standard deviation calculated during the baseline period. Neurons that exhibit both modulations were named as Mixed IFR modulation neurons, while neurons that did not exhibit any significant speech-related firing changes were labelled as No IFR modulation neurons. For a comprehensive description of the firing rate modulation, please refer to Lipski et al [26].

Time-frequency decomposition – Time-varying power and phase were obtained by applying the Hilbert Transform to the band-pass filtered ECoG signal. The signal was bandpass filtered using a 4th order Butterworth filter, with the following frequency ranges: 5–8 Hz for theta, 8–12 Hz for alpha, 12–20 Hz for low beta, 20– 30 Hz for high beta, and center frequencies ranging from 40 to 150 Hz with a bin width of 10 Hz, incrementing by 10 Hz.

Spike-phase coupling implementation – To calculate spike-phase coupling, we considered each possible pair of neurons and ECoG signals that were synchronously recorded. We enforced the following criterion for determining the eligibility of pairs (N = 19755) for subsequent analysis: a minimum 10 trials with stable firing rate and clean ECoG signal. The strength of the spike-phase coupling was quantified by the phase-locking value (PLV), which represents the magnitude of the circular average of unit complex vectors corresponding to the ECoG phase at the time of each spike φ_t, as follows:

$$\mathrm{PLV} = \left| \frac{\sum_{t=1}^{N} e^{j*\varphi_t}}{N} \right|$$

where N is the number of spikes included in the window. PLV is bounded between 0 and 1, indicating lack or perfect spike-phase coupling, respectively. Importantly, PLV is inflated toward 1 when N is low. When N is sufficiently large (N > 50), the pairwise phase consistency (PPC) yields an unbiased estimator of spike-phase coupling[76], as follows:

$$PPC = \frac{N}{N-1}(PLV^2 - \frac{1}{N})$$

In the absence of spike-phase coupling, PPC is expected to be centered around zero, including negative value, when N is finite. As N increases towards infinity, the PPC tends to PLV^2. To ensure that changes in spike-phase coupling depicted genuine and comparable neural correlates, methodological considerations must be discussed. First, the presence of speech-related fluctuations in the instantaneous firing rate poses a challenge in selecting a fixed window size for calculating the phase-locking value (PLV) or PPC over time. This is because variable N can result in uncontrollable and variables biases. Furthermore, low N can lead to noisy estimates of PPC. Second, intra- and inter-participant variability in speech production onset and duration makes event-locked analysis less accurate for the alignment of data around all key task events and not just for one single event (the one used for locking). To overcome all these limitations, we employed a variable-window width SPC estimation procedure developed by Fischer and colleagues[27]. First, we defined five intervals: from 0.75 before auditory cue to auditory cue onset, from auditory cue onset to offset, from auditory cue offset to speech production onset, from speech production onset to offset, and from speech production offset to 0.75 s after. Second, we subdivided these intervals into 8 equidistant anchor points, resulting in 36 anchor points for each trial. Third, we scaled the width of the

window centered at each anchor point such that the sum of spikes (N) across all trials would match a target number as closely as possible. The target number was defined as the average number of spikes in a window of 0.15 s, but always greater than 25 to avoid fewer representative samples. This process allowed to enlarge/shrink the computational window during reduced/increased firing rate periods, ensuring that N remained constant over time. Note that we allowed variable number of spikes across participants to reduce variability in the window width. Finally, each window was placed symmetrically around each anchor point, and we subsequently calculated the PLV metric and applied the PPC correction. The resulting 19755 SPC maps (PPC values of size 16 frequency bins x 36 time points) were smoothed using a time-frequency window ([2, 2] size) and rescaled to the average duration of the event intervals. These maps were then event-locked and averaged across pairs and participants.

Phase polarity standardization – When computing the PLV or PPC, information about the preferred phase is not retained. To identify the preferred phase at which spikes are bundled, we calculated the circular mean using the CircStat toolbox[77]. However, it is important to exercise caution when comparing preferred phases across recordings due to the relative orientation between neural sources and electrodes (i.e. source mixing) and the use of different re-referencing schemas, as these factors can obscure the interpretation of the instantaneous absolute phase[78]. For instance, by applying the bipolar schema, the order of subtraction between two electrodes can flip throughs to peaks and peaks to throughs. To ensure that phases were meaningfully computed across recordings, we applied an automatized polarity-standardization procedure[27]. Specifically, we flipped phases (+ π) such that gamma peaks in the 60-80 Hz range consistently coincided with increases in the local high-frequency activity, which served as a polarity-invariant proxy of background unit activity[79,80]. We computed this proxy by high-pass filtering the ECoG signal at 300 Hz, full-wave rectifying it, and low-pass filtering it with a cut-off of 100 Hz. Flipping procedure was required in 7111/19755 pairs (36%).

Spike-phase coupling events – To further correct the SPC maps for any residual bias and identify genuine increases in SPC, we converted PPC values into z-scores relative to a permutation distribution and performed a cluster-based permutation test[81]. We built the permutation distribution by shuffling the trial-association between STN spikes and EcoG phases 500 times. We paired spike timings from the i-th trial with EcoG phases from the j-th trial (where $I \neq j$). Importantly, to be conservative and preserve the natural appearance of clusters, we applied the same randomization across time-frequency bins. Suprathreshold clusters ($p < 0.05$) were identified in both the original SPC map and in each permutation SPC map by computing the z-score relative to the permutation distribution. If the absolute sum of the z-scores within the original suprathreshold clusters exceeded the 95th percentile of the 500 largest absolute sums of z-scores from the permutation distribution, it was considered statistically significant. These significant clusters in the SPC map were referred to as transient-SPC events (t-SPC events). SPC maps that contained at least one t-SPC event were considered significant.

Spike-phase coupling event characteristics – To fully characterize each t-SPC event, we defined a set of characteristics in the time, frequency, and phase domain. In the temporal domain, we calculated the onset and offset times, temporal duration, and the center of the event (i.e. the mean of the onset and offset). For the frequency domain, we calculated the frequency centroid, as follows:

$$f_{t-SPC} = \frac{\sum_{i}^{T}\sum_{j}^{F} PPC(I,j) * j}{\sum_{i}^{T}\sum_{j}^{F} PPC(I,j)}$$

where I and j are the i-th time bin and j-th frequency bin enclosed within the boundaries of the t-SPC event. In the phase domain, we calculated the circular mean of the t-SPC event phases.

Time occurrence – We aimed to quantify the temporal occurrence of t-SPC events, which we defined as the likelihood of observing at least one t-SPC event in a given STN neuron-ECoG contact pair during each time bin and in each frequency band. To achieve this, we transformed each t-SPC event into a binarized vector, where each time-point (at intervals of 5 ms) was labeled either as part of a t-SPC event and assigned a value of 1, or as part of a non-t-SPC period and assigned a value of 0. We then calculated the mean of these binarized vectors at each time bin, expressed as a percentage. Higher values of this quantity indicated a greater temporal aggregation (i.e., overlap) of t-SPC over time, whereas lower values indicated dispersion. To identify significant time windows of aggregation or dispersion, we employed a permutation test, adapting the approach used to calculate significant beta bursts overlap, as described in [82]. We generated a permutation distribution of the time occurrence due to chance by setting a variable break point in the 0s of the binarized vectors (no slicing of t-SPC events), reversing the two segments, and joining them together. We repeated this process 500 times and extracted the permutation distribution over time. We considered t-SPC events to be significantly dispersed when the time occurrence fell below the 5th percentile of the permutation distribution, and significantly aggregated when the time occurrence rose above the 95th percentile of the permutation distribution. By applying this method, we were able to rigorously determine changes in time occurrence even when the number of t-SPC events was low, and to eliminate spurious trends of aggregation when the number of t-SPC events was high.

Spatial density – We quantified the spatial density of t-SPC events by calculating the ratio between the number of t-SPC events and pairs and expressing it as a percentage, both at the cortical and STN level. To calculate the spatial density on the cortical surface, we used two region-of-interest-based methods. In the first method, we identified 7 regions of interest in the Destrieux atlas[72] that satisfied a minimum coverage criterion (>7 participants and >100 pairs): Precentral gyrus (PreCG), Postcentral gyrus (PostCG), Supramarginal gyrus (SMG), Superior temporal gyrus (STG), Middle frontal gyrus (MFG), and the orbital part of the inferior frontal gyrus (pars O.). We calculated the spatial density in each region of

interest and determined whether t-SPC events were preferentially located in any region of the brain or whether they exhibited no spatial preference, both overall and within frequency band. To test for spatial preference, we created a null permutation distribution (spatial uniform distribution) by shuffling the spatial label of each t-SPC event 500 times. We then compared the spatial density of the original data to the 5th-95th percentiles of the spatial density permutation distribution. Regions with spatial density below the 5th percentile and above the 95th percentile were classified as having “low” or “high” spatial preference, respectively. In the second method, we created a cortical spatial density map by calculating the spatial density in a spherical region of interest with a radius of 2 mm centered around the ECoG recording locations in the MNI space. These maps were converted to the “.node” extension and displayed in SurfIce as nodes. For the STN domain, we built an STN spatial density map by locating spheres (1 mm radius) around the STN neuron locations. Again, we used SurfIce for visualization. For each spatial density map, we identified the peak in each frequency band. Additionally, we projected the STN neurons locations onto the three principal directions of the STN extracted from the DISTAL atlas image[73], following the procedure as described in [55]. To preserve the physical meaning, i.e., distance in mm, of the principal component decomposition, we multiplied the principal component scores by the standard deviation of the MNI coordinates. The principal component coordinates (PC1: antero-posterior direction, PC2: dorso-ventral direction and PC3: medio-lateral direction) represents a more suitable reference of frame, as the STN is not fully spatially aligned with the MNI coordinates (Figure S13). Spatial density computation was repeated for each of the three principal axes using a side of 0.8 mm. We also employed the same formula to calculate the spatial density at the STN neuron level (ratio between the number of t-SPC events and pairs in each neuron). We used this level of analysis to compare SPC at different epochs of the task or frequency bands across neurons and to control for the effect of the firing rate (see **Control analysis**). When comparing β t-SPC during baseline and speech production epoch, we included in the computation only those t-SPC events with mean time in the related behavioral epoch.

Spatial aggregation – To extend and further corroborate our findings in the spatial domain, we also conducted a region-of-interest-free analysis, both at the cortical and STN level. MNI and PC coordinates (and their centroid) of t-SPC events were compared across frequency bands using a permutation test. We then investigated whether t-SPC event locations (within each frequency band) were more spatially aggregated around their centroid than expected by chance (uniform distribution). To this end, we computed the average Euclidean distance between t-SPC events locations and their centroid[55], and compared against a null distribution of surrogate average Euclidean distances obtained by randomly sampling recording locations 500 times.

Relationship between STN spike-phase-coupling topographies and DBS anatomical STN targets – To investigate the relationship between the frequency-specific STN topographies and optimal DBS target

for motor symptoms control in PD, we calculated the Euclidean distance between frequency-wise spatial centroids and location of DBS contacts commonly used for therapeutic stimulation[29,57].

Time delay analysis – As STN neurons often lock to cortical signals within a narrow frequency range, power-based estimates of time delay between STN and cortex might be suboptimal[1]. We calculated time delays using the phase-based analysis, as described in [33]. First, we computed the mean preferred phase of units that were significantly locked in each frequency bin (5-30 Hz range). We then averaged these phases to obtain a grand average phase for each frequency band. By analyzing the gradient of these phases, we determined whether the ECoG channel led (positive sign) or lagged (negative sign) relative to STN neuron, and at what latency this occurred. To test the significance of the time delay, we repeated 500 times the computation using randomly selecting mean angles from each frequency bin. To obtain a p-value, we compared the correlation coefficient in the original data and the 5th-95th percentiles of the correlation coefficient permutation distribution.

Relationship between spike-phase coupling and speech behavior– To examine the link between spike-phase coupling and speech behavior, two sets of analyses were conducted on phonetic accuracy and reaction time. We restricted these analyses only in the low-frequency range (4-40 Hz). For phonetic accuracy, trials were categorized into correct (100%) and error (< 100%) groups; for reaction time, into slow and fast using median split. Only pairs with ≥20 trials in each significant condition from the main analysis were included.

To balance the number of trials, we subsampled 20 trials in each condition and run SPC pipeline 50 times. SPC maps were subtracted and averaged across subsamples. We converted the PPC values into z-scores relative to a permutation distribution defined as the difference between the permuted values in the two conditions. Significance was evaluated using the same procedure as above (see **Spike-phase coupling events**). The significant clusters in the difference SPC map were referred as t-SPC events signaling time-frequency bins in which the first condition was either higher or lower than the second one according to the sign of the z-value. For statistical comparison at the group level between the two conditions, we converted the z-scores to t-values and generated 500 permuted samples by randomly permuting the order of subtraction of the two SPC maps. P-values were estimated using the null distribution and corrected using again a cluster-based procedure.

For the second analysis, trials were sorted according to their phonetic accuracy or reaction time. All spikes within a 1.5 s window ([-1.5 0] s before speech onset for phonetic accuracy and [-2.5 -1] s before speech onset for reaction time) and their phases were determined. The resulting array of phases was then subdivided into seven non-overlapping bins that contained the same number of spikes to compute one PPC for each bin. For each pair, we obtained seven values of PPC and behavioral variable. No z-score correction was applied for this calculation. We then employed a linear mixed-effect model using the function *fitglme.m* from Matlab entering the phonetic accuracy or reaction time as the response variable,

the PPC values as fixed effect. To represent the hierarchical structure of our data, we defined the pair ID nested within neuron ID and participant ID (1|participant: neuron: pair) as the random effect. The statistical significance of all mixed models was evaluated by a model comparison approach. We derived the likelihood ratio by comparing the full model and a null model obtained from the full model by removing the fixed effect, leaving the random effects unchanged. To generate an empirical null distribution of likelihood ratio, we permuted the behavioral variable, while keeping the rest of the model unchanged. We then derived the likelihood ratio using the permuted data and extracted a p-value.

Control analysis – To further ensure the reliability of the t-SPC events identified by our cluster-based permutation analysis, we conducted two control analyses. Firstly, we required that a t-SPC event contain at least two cycles of oscillation at the centroid frequency to be classified as reliable, thus ruling out brief and transitory noise-driven clusters. Secondly, we recognized that surrogate SPC maps generated during the permutation procedure may contain surrogate t-SPC events due to natural cluster tendency arising in small samples of random distribution, which can be mistakenly identified as non-random. To this end, we z-scored the surrogate PPC maps and conducted the same cluster-based permutation, defining surrogate t-SPC events as those that met the same criteria as the original t-SPC events. We then compared the number of observed t-SPC events to that of the surrogate t-SPC events. We also carried out several control analyses to rule out confounding factors that might have influenced the SPC changes we observed: difference in firing rates, differences in ECoG power and a phase reset around speech production onset. Although the SPC pipeline is designed to remove any firing rate bias in the SPC estimation, we sought to investigate genuine firing rate effects by plotting firing rate changes against SPC changes across STN neurons. To ensure that phase estimates were not based on unreliable low amplitude oscillation (during beta suppression), we repeated the analysis and discarded instantaneous phase samples in which the instantaneous power fell below the 10th percentile. We also checked whether bouts of oscillatory power (during theta and gamma increase) biased the SPC estimation by discarding instantaneous phase samples in which the instantaneous power rose above the 90th percentile. We examined the impact of phase resetting of brain oscillations, which can generate event-related activity. To this end, we run two complementary analyses. First, we aligned all the trials to the speech production onset, averaged the ECoG signals across trials to obtain evoked activity, and subtracted this component from individual trials before conducting the SPC analysis. Second, we quantified whether auditory cue or speech production onset reset the phase of the ECoG oscillations. We estimated the event-locked SPC the same way as the SPC, except that ECoG segments were aligned at the auditory cue and speech production onset. Each trial thus contributed one spike to the SPC computation.

Statistical analysis – We used the RainCloud library for the visualization of data distributions[83]. Kolmogorov-Smirnov test revealed that the normality assumption of the distribution was rarely satisfied. For this reason, we decided to apply a series of permutation tests (500 permutations) throughout the manuscript whenever definition of a null distribution was methodologically justified. An exception is

represented by circular data (e.g. phases) that required the usage of the CircStat toolbox[77]. All results were assessed at statistical significance of $\alpha = 0.05$.

Lead Contact – Further information and requests for resources should be directed to the Lead Contact, Robert Mark Richardson (mark.richardson@mgh.harvard.edu).

Data and Code Availability – The data of this study is hosted in the Data Archive BRAIN Initiative (DABI, https://dabi.loni.usc.edu/dsi/1U01NS098969) and is available upon request.

Acknowledgments

We would like to thank research participants for their generous contribution of time and effort in the operating room and additional experimenters who acquired and organized the data. This work was funded by the National Institute of Health (BRAIN Initiative), through grants U01NS098969, U01NS117836 and R01NS110424 to R.M.R.

Author contributions

A.B. and W.J.L. wrote experimental code and performed experiments and recorded data. A.B., W.J.L., L.L.H., J.A.F. and R.M.R. designed experiment. R.M.R performed the surgery and supervised the project. P.F. helped to implement the SPC pipeline and contributed to the interpretation of the SPC results. C.N. wrote parts of the discussion and helped to create the 3D visualization of the SPC topography. M.V. analyzed data, prepared figures, and wrote the first draft of the manuscript. All authors discussed results at all stages of the project and revised the manuscript.

Declaration of interests

The authors declare no competing interest.

References

1. Hickok, G. (2012). Computational neuroanatomy of speech production. Nat Rev Neurosci *13*, 135–145. 10.1038/nrn3158.

2. Chartier, J., Anumanchipalli, G.K., Johnson, K., and Chang, E.F. (2018). Encoding of Articulatory Kinematic Trajectories in Human Speech Sensorimotor Cortex. Neuron *98*, 1042-1054.e4. 10.1016/j.neuron.2018.04.031.

3. Guenther, F.H. (2015). Neural control of speech (The MIT Press).

4. Hickok, G., Venezia, J., and Teghipco, A. (2023). Beyond Broca: neural architecture and evolution of a dual motor speech coordination system. Brain *146*, 1775–1790. 10.1093/brain/awac454.

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

6. Houde, J.F., and Nagarajan, S.S. (2011). Speech Production as State Feedback Control. Front. Hum. Neurosci. *5*. 10.3389/fnhum.2011.00082.

7. Bohland, J.W., and Guenther, F.H. (2006). An fMRI investigation of syllable sequence production. Neuroimage *32*, 821–841. 10.1016/j.neuroimage.2006.04.173.

8. Chen, W., de Hemptinne, C., Miller, A.M., Leibbrand, M., Little, S.J., Lim, D.A., Larson, P.S., and Starr, P.A. (2020). Prefrontal-Subthalamic Hyperdirect Pathway Modulates Movement Inhibition in Humans. Neuron *106*, 579-588.e3. 10.1016/j.neuron.2020.02.012.

9. Jorge, A., Lipski, W.J., Wang, D., Crammond, D.J., Turner, R.S., and Richardson, R.M. (2022). Hyperdirect connectivity of opercular speech network to the subthalamic nucleus. Cell Reports *38*, 110477. 10.1016/j.celrep.2022.110477.

10. Bingham, C.S., Petersen, M.V., Parent, M., and McIntyre, C.C. (2023). Evolving characterization of the human hyperdirect pathway. Brain Struct Funct *228*, 353–365. 10.1007/s00429-023-02610-5.

11. Bar-Gad, I., Morris, G., and Bergman, H. (2003). Information processing, dimensionality reduction and reinforcement learning in the basal ganglia. Progress in Neurobiology *71*, 439–473. 10.1016/j.pneurobio.2003.12.001.

12. Haynes, W.I.A., and Haber, S.N. (2013). The organization of prefrontal-subthalamic inputs in primates provides an anatomical substrate for both functional specificity and integration: implications for Basal Ganglia models and deep brain stimulation. J Neurosci *33*, 4804–4814. 10.1523/JNEUROSCI.4674-12.2013.

13. Chrabaszcz, A., Neumann, W.-J., Stretcu, O., Lipski, W.J., Bush, A., Dastolfo-Hromack, C.A., Wang, D., Crammond, D.J., Shaiman, S., Dickey, M.W., et al. (2019). Subthalamic

Nucleus and Sensorimotor Cortex Activity During Speech Production. J Neurosci *39*, 2698–2708. 10.1523/JNEUROSCI.2842-18.2019.

14. Lipski, W.J., Alhourani, A., Pirnia, T., Jones, P.W., Dastolfo-Hromack, C., Helou, L.B., Crammond, D.J., Shaiman, S., Dickey, M.W., Holt, L.L., et al. (2018). Subthalamic Nucleus Neurons Differentially Encode Early and Late Aspects of Speech Production. J. Neurosci. *38*, 5620–5631. 10.1523/JNEUROSCI.3480-17.2018.

15. Dastolfo-Hromack, C., Bush, A., Chrabaszcz, A., Alhourani, A., Lipski, W., Wang, D., Crammond, D.J., Shaiman, S., Dickey, M.W., Holt, L.L., et al. (2022). Articulatory Gain Predicts Motor Cortex and Subthalamic Nucleus Activity During Speech. Cerebral Cortex *32*, 1337–1349. 10.1093/cercor/bhab251.

16. Weiss, A.R., Korzeniewska, A., Chrabaszcz, A., Bush, A., Fiez, J.A., Crone, N.E., and Richardson, R.M. (2023). Lexicality-Modulated Influence of Auditory Cortex on Subthalamic Nucleus During Motor Planning for Speech. Neurobiology of Language *4*, 53–80. 10.1162/nol_a_00086.

17. Baudouin, R., Lechien, J.R., Carpentier, L., Gurruchaga, J., Lisan, Q., and Hans, S. (2023). Deep Brain Stimulation Impact on Voice and Speech Quality in Parkinson's Disease: A Systematic Review. Otolaryngol.--head neck surg. *168*, 307–318. 10.1177/01945998221120189.

18. Tripoliti, E., Zrinzo, L., Martinez-Torres, I., Frost, E., Pinto, S., Foltynie, T., Holl, E., Petersen, E., Roughton, M., Hariz, M.I., et al. (2011). Effects of subthalamic stimulation on speech of consecutive patients with Parkinson disease. Neurology *76*, 80–86. 10.1212/WNL.0b013e318203e7d0.

19. Tsuboi, T., Watanabe, H., Tanaka, Y., Ohdake, R., Yoneyama, N., Hara, K., Nakamura, R., Watanabe, H., Senda, J., Atsuta, N., et al. (2015). Distinct phenotypes of speech and voice disorders in Parkinson's disease after subthalamic nucleus deep brain stimulation. J Neurol Neurosurg Psychiatry *86*, 856–864. 10.1136/jnnp-2014-308043.

20. Crone, N.E., Hao, L., Hart, J., Boatman, D., Lesser, R.P., Irizarry, R., and Gordon, B. (2001). Electrocorticographic gamma activity during word production in spoken and sign language. Neurology *57*, 2045–2053. 10.1212/WNL.57.11.2045.

21. Chang, E.F., Niziolek, C.A., Knight, R.T., Nagarajan, S.S., and Houde, J.F. (2013). Human cortical sensorimotor network underlying feedback control of vocal pitch. Proc. Natl. Acad. Sci. U.S.A. *110*, 2653–2658. 10.1073/pnas.1216827110.

22. Towle, V.L., Yoon, H.-A., Castelle, M., Edgar, J.C., Biassou, N.M., Frim, D.M., Spire, J.-P., and Kohrman, M.H. (2008). ECoG gamma activity during a language task: differentiating expressive and receptive speech areas. Brain *131*, 2013–2027. 10.1093/brain/awn147.

23. Hamilton, L.S., Edwards, E., and Chang, E.F. (2018). A Spatial Map of Onset and Sustained Responses to Speech in the Human Superior Temporal Gyrus. Current Biology *28*, 1860-1871.e4. 10.1016/j.cub.2018.04.033.

24. Tourville, J.A., Reilly, K.J., and Guenther, F.H. (2008). Neural mechanisms underlying auditory feedback control of speech. NeuroImage *39*, 1429–1443. 10.1016/j.neuroimage.2007.09.054.

25. Lipski, W.J., Wozny, T.A., Alhourani, A., Kondylis, E.D., Turner, R.S., Crammond, D.J., and Richardson, R.M. (2017). Dynamics of human subthalamic neuron phase-locking to motor and sensory cortical oscillations during movement. Journal of Neurophysiology *118*, 1472–1487. 10.1152/jn.00964.2016.

26. Lipski, W.J., Bush, A., Chrabaszcz, A., Crammond, D.J., Fiez, J.A., Turner, R.S., and Richardson, R.M. (2023). Subthalamic nucleus neurons encode syllable sequence and phonetic characteristics during speech (Neuroscience) 10.1101/2023.12.11.569290.

27. Fischer, P., Lipski, W.J., Neumann, W.-J., Turner, R.S., Fries, P., Brown, P., and Richardson, R.M. (2020). Movement-related coupling of human subthalamic nucleus spikes to cortical gamma. eLife *9*, e51956. 10.7554/eLife.51956.

28. Heinrichs-Graham, E., Kurz, M.J., Gehringer, J.E., and Wilson, T.W. (2017). The functional role of post-movement beta oscillations in motor termination. Brain Struct Funct *222*, 3075–3086. 10.1007/s00429-017-1387-1.

29. Horn, A., Neumann, W.-J., Degen, K., Schneider, G.-H., and Kühn, A.A. (2017). Toward an electrophysiological “sweet spot” for deep brain stimulation in the subthalamic nucleus. Human Brain Mapping *38*, 3377–3390. https://doi.org/10.1002/hbm.23594.

30. Averna, A., Debove, I., Nowacki, A., Peterman, K., Duchet, B., Sousa, M., Bernasconi, E., Alva, L., Lachenmayer, M.L., Schuepbach, M., et al. (2023). Spectral Topography of the Subthalamic Nucleus to Inform Next□Generation Deep Brain Stimulation. Movement Disorders, mds.29381. 10.1002/mds.29381.

31. Stolk, A., Brinkman, L., Vansteensel, M.J., Aarnoutse, E., Leijten, F.S., Dijkerman, C.H., Knight, R.T., de Lange, F.P., and Toni, I. (2019). Electrocorticographic dissociation of alpha and beta rhythmic activity in the human sensorimotor system. eLife *8*, e48065. 10.7554/eLife.48065.

32. Lofredi, R., Neumann, W.-J., Bock, A., Horn, A., Huebl, J., Siegert, S., Schneider, G.-H., Krauss, J.K., and Kühn, A.A. (2018). Dopamine-dependent scaling of subthalamic gamma bursts with movement velocity in patients with Parkinson’s disease. eLife *7*. 10.7554/eLife.31895.

33. Sharott, A., Gulberti, A., Hamel, W., Köppen, J.A., Münchau, A., Buhmann, C., Pötter-Nerger, M., Westphal, M., Gerloff, C., Moll, C.K.E., et al. (2018). Spatio-temporal dynamics of cortical drive to human subthalamic nucleus neurons in Parkinson’s disease. Neurobiology of Disease *112*, 49–62. 10.1016/j.nbd.2018.01.001.

34. London, D., Fazl, A., Katlowitz, K., Soula, M., Pourfar, M.H., Mogilner, A.Y., and Kiani, R. (2021). Distinct population code for movement kinematics and changes of ongoing movements in human subthalamic nucleus. eLife *10*, e64893. 10.7554/eLife.64893.

35. Zich, C., Quinn, A.J., Bonaiuto, J.J., O'Neill, G., Mardell, L.C., Ward, N.S., and Bestmann, S. (2023). Spatiotemporal organisation of human sensorimotor beta burst activity. eLife *12*, e80160. 10.7554/eLife.80160.

36. Bergman, H. (2021). The hidden life of the basal ganglia: at the base of brain and mind (The MIT Press).

37. DeLong, M., and Wichmann, T. (2010). Changing Views of Basal Ganglia Circuits and Circuit Disorders. Clin EEG Neurosci *41*, 61–67. 10.1177/155005941004100204.

38. Grant, E., Hoerder-Suabedissen, A., and Molnár, Z. (2012). Development of the Corticothalamic Projections. Front. Neurosci. *6*. 10.3389/fnins.2012.00053.

39. Shimamoto, S.A., Ryapolova-Webb, E.S., Ostrem, J.L., Galifianakis, N.B., Miller, K.J., and Starr, P.A. (2013). Subthalamic Nucleus Neurons Are Synchronized to Primary Motor Cortex Local Field Potentials in Parkinson's Disease. J. Neurosci. *33*, 7220–7233. 10.1523/JNEUROSCI.4676-12.2013.

40. Sharott, A., Vinciati, F., Nakamura, K.C., and Magill, P.J. (2017). A Population of Indirect Pathway Striatal Projection Neurons Is Selectively Entrained to Parkinsonian Beta Oscillations. J. Neurosci. *37*, 9977–9998. 10.1523/JNEUROSCI.0658-17.2017.

41. Cagnan, H., Mallet, N., Moll, C.K.E., Gulberti, A., Holt, A.B., Westphal, M., Gerloff, C., Engel, A.K., Hamel, W., Magill, P.J., et al. (2019). Temporal evolution of beta bursts in the parkinsonian cortical and basal ganglia network. Proc Natl Acad Sci USA *116*, 16095–16104. 10.1073/pnas.1819975116.

42. Bush, A., Zou, J., Lipski, W.J., Kokkinos, V., and Richardson, R.M. (2023). Broadband aperiodic components of local field potentials reflect inherent differences between cortical and subcortical activity (Neuroscience) 10.1101/2023.02.08.527719.

43. Giraud, A.-L., and Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. Nat Neurosci *15*, 511–517. 10.1038/nn.3063.

44. Ding, N., Melloni, L., Zhang, H., Tian, X., and Poeppel, D. (2016). Cortical tracking of hierarchical linguistic structures in connected speech. Nat Neurosci *19*, 158–164. 10.1038/nn.4186.

45. Ding, N., and Simon, J.Z. (2012). Neural coding of continuous speech in auditory cortex during monaural and dichotic listening. Journal of Neurophysiology *107*, 78–89. 10.1152/jn.00297.2011.

46. Ozker, M., Doyle, W., Devinsky, O., and Flinker, A. (2022). A cortical network processes auditory error signals during human speech production to maintain fluency. PLoS Biol *20*, e3001493. 10.1371/journal.pbio.3001493.

47. Hickok, G., and Poeppel, D. (2007). The cortical organization of speech processing. Nat Rev Neurosci *8*, 393–402. 10.1038/nrn2113.

48. Kilavik, B.E., Zaepffel, M., Brovelli, A., MacKay, W.A., and Riehle, A. (2013). The ups and downs of beta oscillations in sensorimotor cortex. Experimental Neurology *245*, 15–26. 10.1016/j.expneurol.2012.09.014.

49. Müller, G.R., Neuper, C., Rupp, R., Keinrath, C., Gerner, H.J., and Pfurtscheller, G. (2003). Event-related beta EEG changes during wrist movements induced by functional electrical stimulation of forearm muscles in man. Neuroscience Letters *340*, 143–147. 10.1016/S0304-3940(03)00019-3.

50. Cardellicchio, P., Hilt, P.M., Dolfini, E., Fadiga, L., and D'Ausilio, A. (2020). Beta Rebound as an Index of Temporal Integration of Somatosensory and Motor Signals. Front. Syst. Neurosci. *14*, 63. 10.3389/fnsys.2020.00063.

51. Alhourani, A., Korzeniewska, A., Wozny, T.A., Lipski, W.J., Kondylis, E.D., Ghuman, A.S., Crone, N.E., Crammond, D.J., Turner, R.S., and Richardson, R.M. (2020). Subthalamic Nucleus Activity Influences Sensory and Motor Cortex during Force Transduction. Cerebral Cortex *30*, 2615–2626. 10.1093/cercor/bhz264.

52. Barone, J., and Rossiter, H.E. (2021). Understanding the Role of Sensorimotor Beta Oscillations. Front. Syst. Neurosci. *15*, 655886. 10.3389/fnsys.2021.655886.

53. Hollunder, B., Ostrem, J.L., Sahin, I.A., Rajamani, N., Oxenford, S., Butenko, K., Neudorfer, C., Reinhardt, P., Zvarova, P., Polosan, M., et al. (2024). Mapping dysfunctional circuits in the frontal cortex using deep brain stimulation. Nat Neurosci. 10.1038/s41593-024-01570-1.

54. Schneider, M., Broggini, A.C., Dann, B., Tzanou, A., Uran, C., Sheshadri, S., Scherberger, H., and Vinck, M. (2021). A mechanism for inter-areal coherence through communication based on connectivity and oscillatory power. Neuron *109*, 4050-4067.e12. 10.1016/j.neuron.2021.09.037.

55. van Wijk, B.C.M., Neumann, W.-J., Kroneberg, D., Horn, A., Irmen, F., Sander, T.H., Wang, Q., Litvak, V., and Kühn, A.A. (2022). Functional connectivity maps of theta/alpha and beta coherence within the subthalamic nucleus region. NeuroImage *257*, 119320. 10.1016/j.neuroimage.2022.119320.

56. Aldridge, D., Theodoros, D., Angwin, A., and Vogel, A.P. (2016). Speech outcomes in Parkinson's disease after subthalamic nucleus deep brain stimulation: A systematic review. Parkinsonism & Related Disorders *33*, 3–11. 10.1016/j.parkreldis.2016.09.022.

57. Caire, F., Ranoux, D., Guehl, D., Burbaud, P., and Cuny, E. (2013). A systematic review of studies on anatomical position of electrode contacts used for chronic subthalamic stimulation in Parkinson's disease. Acta Neurochir *155*, 1647–1654. 10.1007/s00701-013-1782-1.

58. Jorge, A., Dastolfo-Hromack, C., Lipski, W.J., Kratter, I.H., Smith, L.J., Gartner-Schmidt, J.L., and Richardson, R.M. (2020). Anterior Sensorimotor Subthalamic Nucleus Stimulation Is Associated With Improved Voice Function. Neurosurg. *87*, 788–795. 10.1093/neuros/nyaa024.

59. Plaha, P. (2006). Stimulation of the caudal zona incerta is superior to stimulation of the subthalamic nucleus in improving contralateral parkinsonism. Brain *129*, 1732–1747. 10.1093/brain/awl127.

60. Åström, M., Tripoliti, E., Hariz, M.I., Zrinzo, L.U., Martinez-Torres, I., Limousin, P., and Wårdell, K. (2010). Patient-Specific Model-Based Investigation of Speech Intelligibility and Movement during Deep Brain Stimulation. Stereotact Funct Neurosurg *88*, 224–233. 10.1159/000314357.

61. Mikos, A., Bowers, D., Noecker, A.M., McIntyre, C.C., Won, M., Chaturvedi, A., Foote, K.D., and Okun, M.S. (2011). Patient-specific analysis of the relationship between the volume of tissue activated during DBS and verbal fluency. NeuroImage *54*, S238–S246. 10.1016/j.neuroimage.2010.03.068.

62. Buzsáki, G., and Schomburg, E.W. (2015). What does gamma coherence tell us about inter-regional neural communication? Nat Neurosci *18*, 484–489. 10.1038/nn.3952.

63. Aumann, T.D., and Prut, Y. (2015). Do sensorimotor β-oscillations maintain muscle synergy representations in primary motor cortex? Trends in Neurosciences *38*, 77–85. 10.1016/j.tins.2014.12.002.

64. Remple, M.S., Bradenham, C.H., Kao, C.C., Charles, P.D., Neimat, J.S., and Konrad, P.E. (2011). Subthalamic nucleus neuronal firing rate increases with Parkinson's disease progression: STN Neurophysiology in Early vs Late-Stage PD. Mov. Disord. *26*, 1657–1662. 10.1002/mds.23708.

65. Brown, P., Oliviero, A., Mazzone, P., Insola, A., Tonali, P., and Di Lazzaro, V. (2001). Dopamine dependency of oscillations between subthalamic nucleus and pallidum in Parkinson's disease. J. Neurosci. *21*, 1033–1038.

66. Hammond, C., Bergman, H., and Brown, P. (2007). Pathological synchronization in Parkinson's disease: networks, models and treatments. Trends Neurosci. *30*, 357–364. 10.1016/j.tins.2007.05.004.

67. Neumann, W.-J., Staub-Bartelt, F., Horn, A., Schanda, J., Schneider, G.-H., Brown, P., and Kühn, A.A. (2017). Long term correlation of subthalamic beta band activity with motor impairment in patients with Parkinson's disease. Clin Neurophysiol *128*, 2286–2291. 10.1016/j.clinph.2017.08.028.

68. Vissani, M., Palmisano, C., Volkmann, J., Pezzoli, G., Micera, S., Isaias, I.U., and Mazzoni, A. (2021). Impaired reach-to-grasp kinematics in parkinsonian patients relates to dopamine-dependent, subthalamic beta bursts. npj Parkinsons Dis. *7*, 53. 10.1038/s41531-021-00187-6.

69. Vissani, M., Isaias, I.U., and Mazzoni, A. (2020). Deep brain stimulation: a review of the open neural engineering challenges. J. Neural Eng. 10.1088/1741-2552/abb581.

70. Randazzo, M.J., Kondylis, E.D., Alhourani, A., Wozny, T.A., Lipski, W.J., Crammond, D.J., and Richardson, R.M. (2016). Three-dimensional localization of cortical electrodes in deep

brain stimulation surgery from intraoperative fluoroscopy. NeuroImage *125*, 515–521. 10.1016/j.neuroimage.2015.10.076.

71. Neudorfer, C., Butenko, K., Oxenford, S., Rajamani, N., Achtzehn, J., Goede, L., Hollunder, B., Ríos, A.S., Hart, L., Tasserie, J., et al. (2023). Lead-DBS v3.0: Mapping deep brain stimulation effects to local anatomy and global networks. NeuroImage *268*, 119862. 10.1016/j.neuroimage.2023.119862.

72. Destrieux, C., Fischl, B., Dale, A., and Halgren, E. (2010). Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. NeuroImage *53*, 1–15. 10.1016/j.neuroimage.2010.06.010.

73. Ewert, S., Plettig, P., Li, N., Chakravarty, M.M., Collins, D.L., Herrington, T.M., Kühn, A.A., and Horn, A. (2018). Toward defining deep brain stimulation targets in MNI space: A subcortical atlas based on multimodal MRI, histology and structural connectivity. NeuroImage *170*, 271–282. 10.1016/j.neuroimage.2017.05.015.

74. Oostenveld, R., Fries, P., Maris, E., and Schoffelen, J.-M. (2011). FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. Computational Intelligence and Neuroscience *2011*, 1–9. 10.1155/2011/156869.

75. Bush, A., Chrabaszcz, A., Peterson, V., Saravanan, V., Dastolfo-Hromack, C., Lipski, W.J., and Richardson, R.M. (2022). Differentiation of speech-induced artifacts from physiological high gamma activity in intracranial recordings. NeuroImage *250*, 118962. 10.1016/j.neuroimage.2022.118962.

76. Vinck, M., van Wingerden, M., Womelsdorf, T., Fries, P., and Pennartz, C.M.A. (2010). The pairwise phase consistency: a bias-free measure of rhythmic neuronal synchronization. Neuroimage *51*, 112–122. 10.1016/j.neuroimage.2010.01.073.

77. Berens, P. (2009). **CircStat**□: A *MATLAB* Toolbox for Circular Statistics. J. Stat. Soft. *31*. 10.18637/jss.v031.i10.

78. Shirhatti, V., Borthakur, A., and Ray, S. (2016). Effect of Reference Scheme on Power and Phase of the Local Field Potential. Neural Computation *28*, 882–913. 10.1162/NECO_a_00827.

79. Eckhorn, R., Bauer, R., Jordan, W., Brosch, M., Kruse, W., Munk, M., and Reitboeck, H.J. (1988). Coherent oscillations: A mechanism of feature linking in the visual cortex?: Multiple electrode and correlation analyses in the cat. Biol. Cybern. *60*, 121–130. 10.1007/BF00202899.

80. Pooresmaeili, A., Poort, J., Thiele, A., and Roelfsema, P.R. (2010). Separable Codes for Attention and Luminance Contrast in the Primary Visual Cortex. J. Neurosci. *30*, 12701–12711. 10.1523/JNEUROSCI.1388-10.2010.

81. Maris, E., and Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. Journal of Neuroscience Methods *164*, 177–190. 10.1016/j.jneumeth.2007.03.024.

82. Tinkhauser, G., Torrecillos, F., Duclos, Y., Tan, H., Pogosyan, A., Fischer, P., Carron, R., Welter, M.-L., Karachi, C., Vandenberghe, W., et al. (2018). Beta burst coupling across the motor circuit in Parkinson's disease. Neurobiology of Disease *117*, 217–225. 10.1016/j.nbd.2018.06.007.

83. Allen, M., Poggiali, D., Whitaker, K., Marshall, T.R., van Langen, J., and Kievit, R.A. (2021). Raincloud plots: a multi-platform tool for robust data visualization. Wellcome Open Res *4*, 63. 10.12688/wellcomeopenres.15191.2.