Official Study Title:

Flexible representation of speech in the supratemporal plane

NCT Number:

NCT05209386

Document:

Statistical Analysis Plan

Document Date:

July 15, 2020

STATISTICAL DESIGN AND POWER ANALYSIS

Statistical Design, Power and Analytic Approach. We will test neurosurgical patients (ages 15-25 years) undergoing evaluation and treatment who require chronic (>7 d) sEEG implantation and recording of the temporal lobe for seizure localization or brain mapping. One set of behavioral tasks, and corresponding electrophysiological measures, will allow us to address Aim 1. Another will address Aim 2. Patient permitting, our goal is to have the same patient complete all tasks. All participants will undergo neuropsychological tests and will have detailed history/demographic information as part of their clinical battery.

Power and Sample Size. Robust behavioral effects will allow examination of perceptual weighs in neural representation. Using behavioral effect sizes from **Figure 2** pilot data to estimate the sample size required for a predicted power of 0.8 (two-tailed alpha at .05) yields a sample of N=25. This leaves open the issue of power for neural measures. Our pilot EEG data with the same task/stimuli revealed robust effects with N=23 (**Figure A**), reassuring that N=25 is a reliable target for EEG. The SNR advantages of sEEG versus EEG, and our sEEG pilot data in **Figure 3**, suggest that this sample size will be more than sufficient for sEEG measures, especially as they will be utilized in a within-patient/within-electrode experimental design.

Analyses. *Specific pre-registered analyses* will assess the hypotheses outlined in the Approach section of the proposal. Specifically, we will test the following hypotheses:

> (**H1**) broadly, the relative perceptual weight of VOT and F0, as measured behaviorally, will be reflected in cortical response, (**H2**) with modulation as a function of baseline perceptual weights, (**H3**) shifts experimentally invoked by a change in listening context by presenting speech in noise, (**H4**) and by introducing an 'accent' that shifts short-term input regularities across VOT and F0. In the latter case, our approach will allow us to test the specific directional hypothesis (**H5**) that F0 perceptual weights in the DBSL paradigm are both exaggerated by Canonical input regularities that cleanly convey a VOTxF0 correlation consistent with English and that F0 perceptual weights are down-weighted upon introduction of a regularity that violates the typical pattern of English (supported by scalp EEG pilot, **Figure A**). Our use of sEEG allows us to evaluate these hypotheses across the supratemporal plane thereby testing the strong, and falsifiable, hypothesis (**H6**) that adaptive plasticity effects are present in HG versus (**H7**) apparent only at higher levels of the cortical hierarchy. Our ability to test these hypotheses is complemented by sEEG electrode placement in cortical regions outside STP (see **Figure 3**) which will support secondary hypotheses and serve as control electrode sites.

The study design is justified our extensive behavioral research demonstrating the feasibility of the project rationale. On the electrophysiological side, our pilot data (**Figures A,3**) demonstrate the feasibility of recording robust sEEG and EEG signals responsive to the acoustic dimensions we manipulate. This will provide clear, informative, interpretable data with which to evaluate the hypotheses listed above.

Evaluating the Hypotheses: Behavioral Analyses. We will evaluate the behavioral impact of the VOT and F0 acoustic dimensions on classification using *mixed-effects logistic regression* (with patient as a random effect, stimulus VOT and F0 as fixed effects and classification responses as the outcome). Following our prior work,[14,25-29] perceptual weights for the dimensions will be computed for each patient as the correlation between dimension values and the proportion of *peach* classifications across all stimuli in the VOTxF0 stimulus grid with absolute values of the correlation coefficients normalized to sum to one as an *index of relative perceptual weight* in quiet (Aim 1) and in noise (Aim 2). To examine the impact of a change in 'accent' in the DBSL paradigm, we will use mixed logit models with responses as a function of patient, block, test stimulus F0 and the interaction between block and F0 (Aim 2). Patient will be modeled as a random effect, with the other factors as fixed effects.

Evaluating the Hypotheses: Neural Analyses. We will take a multi-pronged analysis approach. Our pilot data in **Figure 3** demonstrate that *high-γ* activity (HGA) is modulated by graded acoustic details across the VOT and F0 dimensions across electrodes placed in the STP. Following the analysis pipeline used in our pilot data analyses, we will specifically examine stimulus-time-locked HGA to the *Test* stimuli, which possess consistent, perceptually-ambiguous VOT and differentiated F0. We have specific, directional predictions (detailed above) regarding how HGA to F0-differentiated *Test* stimuli will vary according to perceptual weight in behavior. To briefly recap, we expect that HGA to *Test* stimuli will be better differentiated (1) in the Speech-in-Noise compared to the Clear baseline context (because F0 carries greater perceptual weight); (2) in the Canonical, compared to the Reverse, context; (3) in the Canonical, compared to the Baseline Clear Speech context (because stimulus statistics exaggerate the dimension regularity in behavior).

High-gamma amplitude will be calculated using an approach that utilizes the Hilbert transform. Specifically, the signal will be filtered into 8 subbands, logarithmically spaced between 70-150 Hz. For each subband, we will calculate the amplitude (absolute value) of the analytic signal, which is estimated using the Hilbert transform. Each subband will be normalized to its baseline mean and standard deviation, estimated across trials. The HGA estimate is then the mean across these subbands.

We will use *least-squares linear regression neural encoding models* to investigate relationships between acoustic stimulus dimensions and STP neural responses during the baseline quiet context, in which stimuli sample a 2-d F0 x VOT grid. This approach will allow us to identify the subset of electrodes and temporal windows that encode at least one of these 2 dimensions at baseline; targeted analysis on this subset (described below) will then be used to compare listening contexts.

Encoding model inputs will consist of F0 and VOT, with an output of channel- and time-specific HGA. For a given electrode, individual models will be built using single trial data and a sliding window, allowing us to identify the temporal window relative to stimulus onset that yields significant models. Model quality will be assessed in two ways. First, using models built on all trials, we will calculate the regression F-statistic, which determines if any coefficients are significant. This will be compared to null distributions estimated with permutation tests that shuffle data across trials. Second, goodness-of-fit will be assessed for significant models by performing leave-one-out cross-validation and calculating R^2, the proportion of variance in neural activity explained by the model. Finally, we will assess the relative encoding strength of each acoustic dimension using model coefficient t-statistics. In summary, this approach will allow us to identify the timing and anatomical location of F0 and VOT encoding during baseline quiet listening conditions for which dimensions are sampled orthogonally (as in **Figure 1a**).

To characterize shifts in neural encoding across listening contexts, we will investigate Noise using encoding models and Canonical/Reverse contexts using *cluster-based approaches*.[32,49] Encoding models will be built from the Noise context for all channels and timepoints, using the same approach as Baseline. If a channel/timepoint is significant in either Baseline or Noise, the F0 and VOT coefficient t-statistics will be compared across contexts. We hypothesize that encoding of F0 will strengthen and VOT will weaken in Noise, as measured by changes in t-statistic magnitudes across multiple models. Next, we will compare neural responses for each of the two *Test* stimuli (which were embedded in the F0 x VOT grids, **Figure 1a**) between Baseline vs. Canonical and Canonical vs. Reverse. The clustering approach will look only at Baseline significant channels and time windows and involves randomly assigning listening context labels to single-trial data followed by a t-test at each time step. Across all permutations, a criterion value will be established for each timepoint (>95% of absolute value of t). For each of these permutations, we will next determine whether its value exceeds criterion across timepoints, and for how many timepoints it exceeds criterion (a ‘cluster’). For each cluster, t values will be summed and assigned to all points in the cluster, with the largest summed cluster value stored for each permutation. This will create a null distribution of 1000 cluster values. We then will establish whether the cluster size calculated across real neural data (organized according to listening context) exceeds the 95% permutation-based cluster values such that $p<.001$ indicates an observed cluster is greater than all permutation-based clusters. Using this approach, we will identify context-dependent shifts in HGA responses, which we predict will reflect observed shifts in perceptual weights. Namely, we hypothesize that HGA responses in F0-encoding channels will be enhanced in the Canonical context relative to both Baseline and the Reverse context.

Rigor and Reproducibility. Analyses will be controlled for multiple comparisons, with sex as a co-variate in our analyses. We will use pre-registration and provide access to all the deidentified source data.