

Examining Functional Connectivity Between Broca's Area and Premotor Cortex During Speech Production

Aidan Truel

Supervisors: Jacob Ayers, Richard Hahnloser

Institute for Neuroinformatics, University of Zurich and ETH Zurich

aidangerrick.truel@uzh.ch, jayers@ethz.ch, rich@ini.ethz.ch

Abstract—Spoken communication is a central part of our lives, which can be impaired in many ways following brain damage. Neuroscientists have long been interested in how our cortex supports speech production and comprehension. Broca's area has been historically associated with speech production, but its exact role in speech remains unclear, with recent studies suggesting it may not be directly involved in articulation. Using intracortical recordings from an ALS patient with multielectrode arrays in Broca's area and premotor cortex, we analyzed functional connectivity during word and sentence production. Granger causality tests revealed that Broca's area predicts premotor activity at short time lags, while premotor cortex predicts Broca's area activity more strongly at longer time delays, with higher predictive value for sentences than words. These findings suggest that Broca's area contributes to articulatory planning prior to and during speech rather than direct motor execution.

Index Terms—Broca's Area, Premotor Cortex, Granger Causality, Articulation

I. INTRODUCTION

In 1861, the French anatomist Paul Broca observed that damage to the left inferior frontal gyrus (IFG) was associated with non-fluent aphasia, a condition characterized by halting or effortful speech. This region, later termed Broca's area, became the first brain region linked to a specific function — speech production. The subsequent identification of Wernicke's area led to the influential Wernicke-Geschwind model [3], which solidified speech production and comprehension as textbook examples of cortical specialization.

Despite a century and a half of research, there is no consensus about the precise role of Broca's area in speech production. Early lesion studies suggested it was critical for articulation. More recent models of speech production and comprehension have attempted to unify a large number of results from functional magnetic resonance imaging (fMRI) [4], electroencephalography (EEG), and diffusion tensor imaging (DTI) [2]. While ambitious in scope, such models do little to elucidate the role of Broca's area beyond situating it within a larger network of speech related areas.

Even more recently, intracranial recording and perturbation studies have shed light on the temporal dynamics of neural activity underlying speech and the behavioral consequences of temporary lesions. Intracranial recordings (ECoG) provide

spatial and temporal resolution not available in non-invasive modalities, such as electroencephalography (EEG) and functional magnetic resonance imaging (fMRI). Multiple groups have analyzed ECoG data and found evidence that Broca's area is active prior to, but not during, speech. Computing cross correlations between high gamma activity in Broca's area and motor cortices, Flinker et al. found correlated activity with an average delay of 241 ms [1]. Likewise, Magrassi et al. cross correlated local field potentials (LFP, low frequency ECoG) in Broca's area with speech volume and found an average delay of 245 ms [5]. Furthermore, Flinker et al., using multivariate granger causality techniques with high gamma (70-150 Hz) power, found reciprocal information flow between auditory cortices and Broca's area during word perception, followed by information flow between Broca's area and motor cortices preceding single word production. Based on these results, they hypothesized that Broca's area forms an articulatory code to be implemented by downstream motor areas. This view is also supported by a study in which Broca's area and motor cortex were cooled during speech, thus slowing down neural activity. The authors observed that cooling Broca's area interferes with speech timing, while cooling motor cortex decreases the quality of speech. The authors proposed that Broca's area is responsible for the generation of sequences in speech production.

Advances in neural decoding of speech provide another avenue to assess the role of Broca's area. Several studies have successfully decoded speech from neural activity using fMRI [9], electrocorticography (ECoG) [7], and microelectrode arrays (MEA) [10]. However, these successes may rely on patterns of neural activity in motor-related areas rather than dedicated speech-processing regions. For instance, a recent intracortical brain-computer interface (BCI) study [10] achieved a 24% word error rate in a large-vocabulary speech decoding task with real-time neural data and a language model. The subject was a patient with bulbar-onset amyotrophic lateral sclerosis (ALS), which specifically affects muscles in the face and neck, and activities such as swallowing and speech. For this reason, authors were able to obtain invasive recordings after neurosurgically implanting four MEAs. Two arrays were placed in premotor cortex (superior and inferior BA6v) and

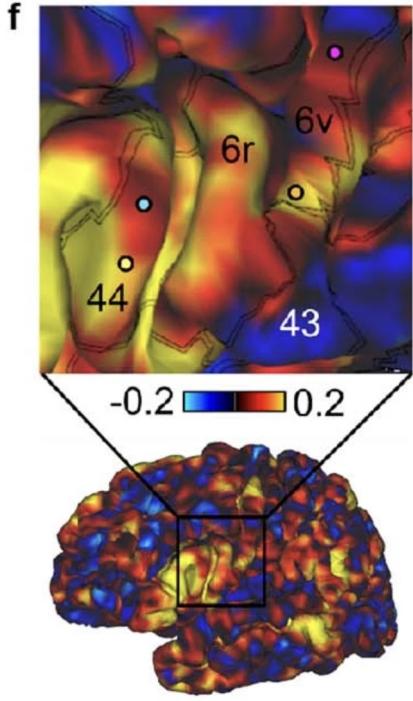


Fig. 1. From [10] Extended data Fig. 2f. MEA placement alongside a language related resting state network from an fMRI scan of the patient

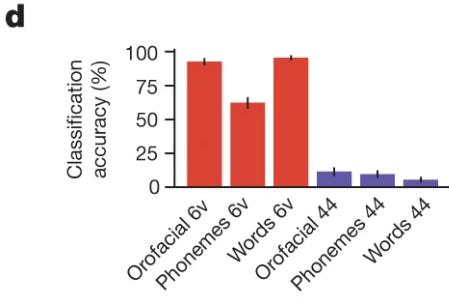


Fig. 2. Accuracy for classification of movements, words and phonemes based on neural activity in BA44 and BA6v from [10] Fig. 1d

two were placed in Broca's area (superior and inferior BA44). Refer to Fig. 1 for MEA placement. Surprisingly, the researchers found no clear tuning to words, phonemes, or muscle movements in BA44, leading them to exclude those electrodes from further decoding analyses (see Fig. 2).

The present study aims to further investigate the interaction between Broca's area (BA44) and premotor cortex (BA6v) during speech production using intracortical data from the ALS patient. The publicly available data lacks the temporal resolution of ECoG, since the raw voltages were squared and averaged over 20ms time bins. However, the MEAs provide much greater spatial resolution and this presents a unique opportunity to investigate the neural circuitry underlying speech. We first replicate the finding from [10] and [1] that Broca's area is not active during word production, using

simple methods that do not rely on classification accuracy. We then test the hypothesis that BA6v is functionally downstream of BA44 and analyze the connectivity patterns between the two regions. Finally, we compare BA44 activity during word and sentence production, as well as periods between trials to test the hypothesis that Broca's area is specifically involved in processing sequences. Functional connectivity is inferred using Granger causality at different time delays.

Our results suggest that BA44 is not uniformly active during speech but shows specific engagement in sentence production. Moreover, we find a higher proportion of Granger causal interactions from superior BA44 electrodes at early time lags, and a higher proportion of interactions from inferior BA6v electrodes at later time lags. Furthermore, there were more interactions at longer latencies for sentences than words.

By leveraging high-resolution intracortical recordings, this study provides new insights into Broca's functional role in speech production. Together, our findings reinforce the hypothesis that Broca's area plays a role in shaping articulatory plans prior to implementation in subsequent motor areas during speech.

II. METHODS

A. Data

Data were collected from four 8x8 MEAs, including two in BA44 and two in BA6v (Fig. 1) implanted in a patient with ALS [10]. In each area, one of the arrays was more dorsal/superior and the other more ventral/inferior. The collection of electrode channels in each array will subsequently be referred to as sup44, inf44, sup6v and inf6v.

During data collection, a word or sentence was presented, which the participant then attempted to speak. The period of speaking was followed by a delay period in which the subject was silent and neural data was continuously recorded. Voltages from each MEA channel were high-pass filtered at 250 Hz to capture neural spiking activity. The mean of squared voltages was computed for each 20ms period and recorded for subsequent analysis. This data, referred to as spikePow, is used for all further processing in this work.

The neural data was recorded during 20 repetitions of each of 51 distinct words. Neural data during sentence production was recorded on 24 separate days, but data was only analyzed from one recording session, in which 300 distinct sentences were spoken.

B. Channel Activity

Histograms of channel activity were compared using the Jensen-Shannon Divergence (JSD), a symmetric, smoothed version of KL-Divergence [6]. A large JSD indicates that two distributions are very different. To compare two probability mass functions (PMFs) P and Q, the JSD is computed as

$$JSD(P||Q) = \frac{1}{2}D_{KL}(P||M) + \frac{1}{2}D_{KL}(Q||M) \quad (1)$$

where

$$M = \frac{1}{2}(P + Q) \quad (2)$$

TABLE I
TOTAL NUMBER OF GRANGER CAUSALITY TESTS PERFORMED: DIRECTIONS (2) X (WORDS X TRIALS PER WORD) X CAUSAL ELECTRODES X CAUSED ELECTRODES X TIME LAGS

Condition	sup44 ↔ sup6v	sup44 ↔ inf6v	inf44 ↔ sup6v	inf44 ↔ inf6v	Row Sum
Word Trial (Speaking)	$2 \times 51 \times 20 \times 64 \times 64 \times 12 = 100,270,080$	100,270,080	100,270,080	100,270,080	401,080,320
Word Trial (Delay)	$2 \times 51 \times 20 \times 64 \times 64 \times 12 = 100,270,080$	100,270,080	100,270,080	100,270,080	401,080,320
Sentence Trial (Speaking)	$2 \times 300 \times 64 \times 64 \times 8 = 19,660,800$	19,660,800	19,660,800	19,660,800	78,643,200
Sentence Whole Session	$2 \times 64 \times 64 \times 8 = 65,536$	65,536	65,536	65,536	262,144
Word Whole Session	$2 \times 64 \times 64 \times 8 = 65,536$	65,536	65,536	65,536	262,144
Column Sum	220,332,032	220,332,032	220,332,032	220,332,032	881,328,128

is a mixture distribution of P and Q, and the KL-Divergence is computed as

$$D_{KL}(P(x)||Q(x)) = \sum_x P(x) \ln \frac{P(x)}{Q(x)} \quad (3)$$

3

C. Granger Causality

To determine whether a time series Y Granger causes a time series X, two models are fit to X using ordinary least squares (OLS). The autoregressive (AR) model uses only past values of X

$$X[t] = e_x[t] + \sum_{d=1}^T a_d X[t-d] \quad (4)$$

where a_d are the model coefficients, e_x is the vector of error terms between the AR model and the time series, for each time point and T is the maximum delay.

Then, a vector autoregressive (VAR) model is fit using past values of Y as well as X

$$X[t] = \epsilon_x[t] + \sum_{d=1}^T a_{xx,d} X[t-d] + a_{xy,d} Y[t-d] \quad (5)$$

The Granger causality is then computed as

$$GC_{Y \rightarrow X} = \ln \left(\frac{\text{Var}(\epsilon_x)}{\text{Var}(e_x)} \right) \quad (6)$$

Finally, the null hypothesis that $GC_{Y \rightarrow X} = 0$ is assessed using an F test and a χ^2 test and the p-value is retained. The p-value of the F test was used for further analyses, as it is a more stringent test.

The causal Granger (GC) interactions tests were carried out between every channel in BA44 and every channel in BA6v using an open source Python package [8]. The tests were carried out during the speaking and delay periods for each of 1020 word trials and 300 sentence trials. In addition, tests were conducted on neural activity from the entire recording session for both words and sentences. For each condition, tests were conducted for Granger causal influences from 44 onto 6v and for 6v onto 44. See Table I for a summary of the total number of tests performed.

For all presented results, a single time lag was chosen, indicated in the figure captions. The array of p-values was binarized, retaining only those below the significance threshold $\alpha = 0.01$, after Bonferroni correction. That is, to correct for

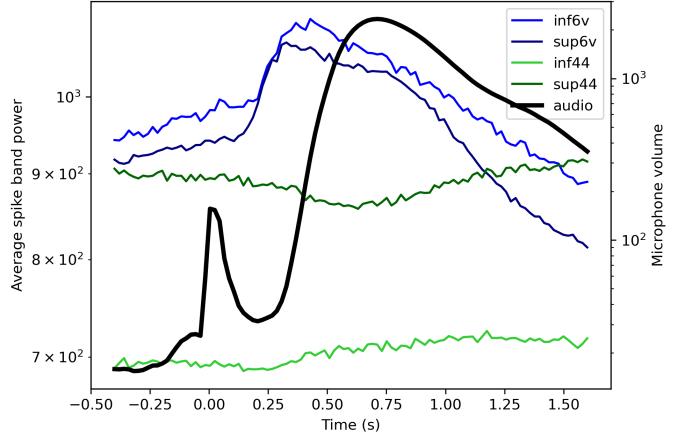


Fig. 3. The time course of neural activity, averaged over trials and channels, was plotted for words on a log scale

multiple comparisons, the p-value was divided by the number of tests performed within each condition ($64 \times 64 \times \text{Number of Trials}$).

III. RESULTS

A. Temporal Dynamics

1) *Word Trials:* There is a slight increase in volume when the word is presented (0s) and the subject begins speaking 250ms after word presentation at the earliest. Activity in both premotor regions follows a similar trajectory, ramping up after cue presentation and peaking at the initiation of speech. Broca's area exhibits larger differences between inferior and superior regions, with higher sup44 activity. Averaged activity in Broca's area does not change as drastically throughout the recording, but on average, sup44 channels decrease activity, while inf44 channels increase activity, during word production.

2) *Sentence Trials:* In contrast to the smooth curves in the word trial averages, Broca's activity was highly variable over time on sentence trials. Activity in inf44 was higher, on average, than sup44 but both were much higher than premotor channels. Activity in inf6v is lower than sup6v before stimulus presentation but higher for the first 6s of speaking, before dropping below sup6v again.

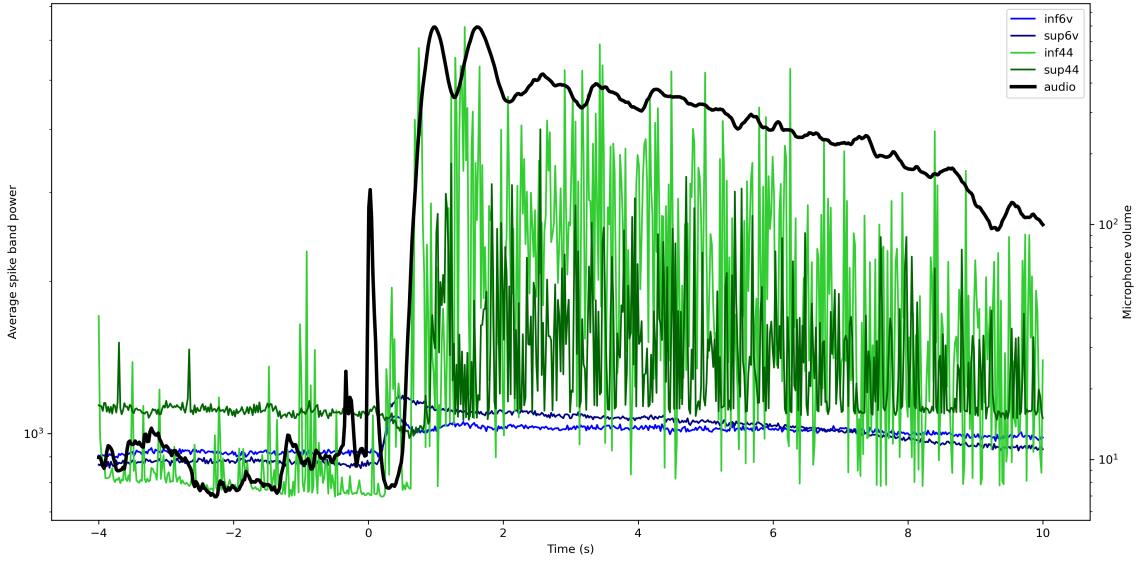


Fig. 4. The time course of neural activity, averaged over trials and channels, was plotted for sentence trials on a log scale

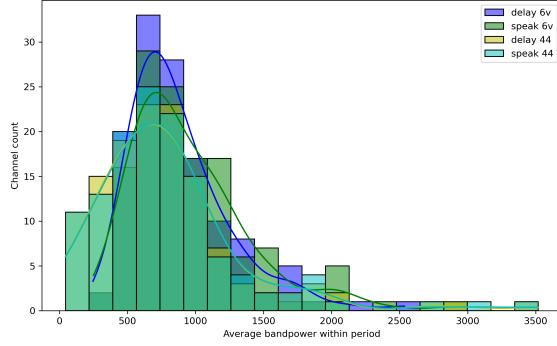


Fig. 5. Average word activity (per channel) during speak vs. delay periods

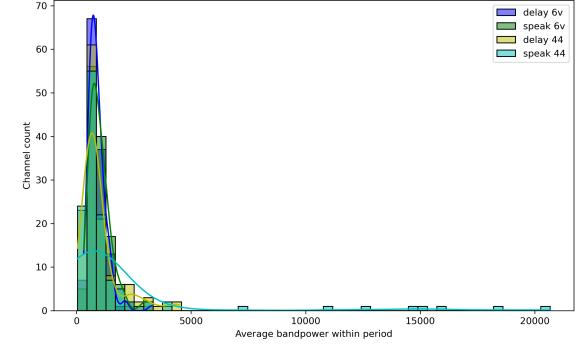


Fig. 6. Average sentence activity (per channel) during speak vs. delay periods

B. Channel Histograms

The distribution of activity over channels in BA44 and BA6v was also plotted, with histograms and fitted kernel density estimates (KDE) of average activity of each channel. For BA44 and BA6v, activity was aggregated in superior and inferior channels and plotted during speaking and delay conditions. During word production, the number of BA44 channels at higher bandpower was consistently lower than BA6v, with BA6v activity during speaking the highest. In contrast, during sentence production, the average channel activity was highest in BA44 while speaking. This appears to be driven by extremely high activity (> 5000) in eight channels.

The distributions were quantitatively compared with the Jensen-Shannon divergence. For word trials (Fig. 7), speak

and delay conditions in 6v were slightly more different than in 44. The highest differences were between 6v and 44, and all were roughly equal. For sentence trials (Fig. 8), the most similar distributions were 6v activity during speak and delay. The greatest differences was between 44 during speaking and 6v (both delay and speak conditions). The difference between area 44 activity during speaking and delay periods is quite similar to the difference between 44 delay and 6v (during speaking and delay).

C. Granger Causality

1) *Single Trial Tests:* After testing for Granger causality on individual word trials, the number of significant interactions between areas was compared. Interaction counts were summed over word, trial, and channel axes, and the number of signifi-

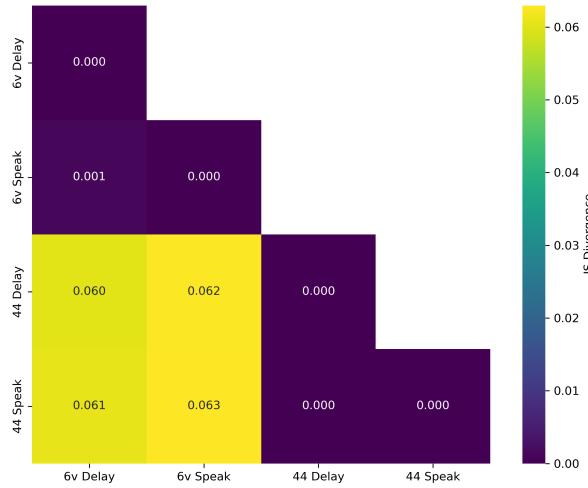


Fig. 7. Jensen Shannon Divergence Heatmap for Words

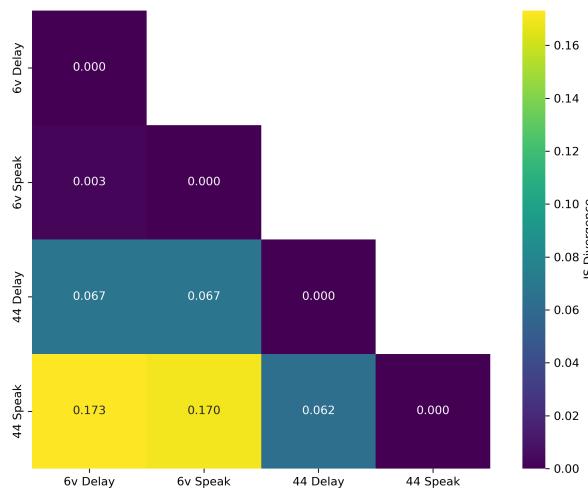


Fig. 8. Jensen Shannon Divergence Heatmap for Sentences

cant interactions is small relative to the maximum number of interactions ($51 * 20 * 64 * 64 = 4,177,920$). Contrary to our expectations, there were roughly equal counts of significant interactions from $44 \rightarrow 6v$ and $6v \rightarrow 44$. Furthermore, the activity number of significant interactions was roughly equal for speaking and delay periods. This could be because Broca's area is not very active during the production of individual words.

2) Whole Session Tests: We ran Granger causality tests between each pair of channels for the entire duration of recording (for both words and sentences). The proportion of significant interactions was roughly equal for words and sentences at short delays, however at longer delays, there were more significant interactions for sentences between every pair of arrays. Furthermore, at short delays, there was a greater number of feedforward interactions ($44 \rightarrow 6v$) originating from sup44, while at long delays, there was a much greater number of feedback interactions ($6v \rightarrow 44$) originating from

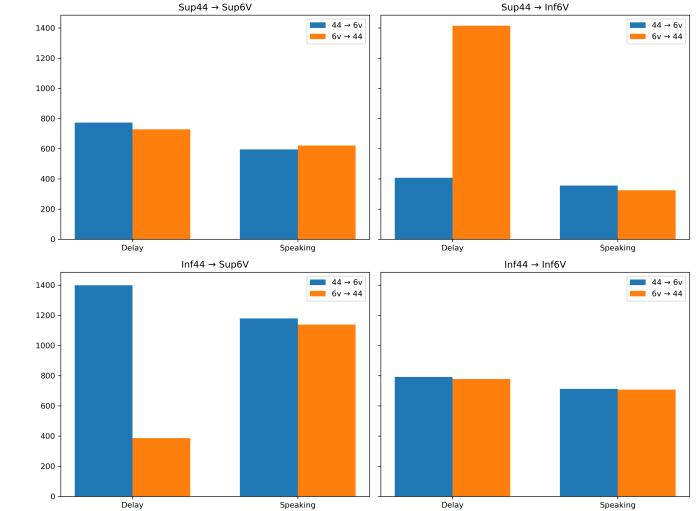


Fig. 9. Number of significant Granger Causal Interactions During Speaking and Delay Periods (60ms)

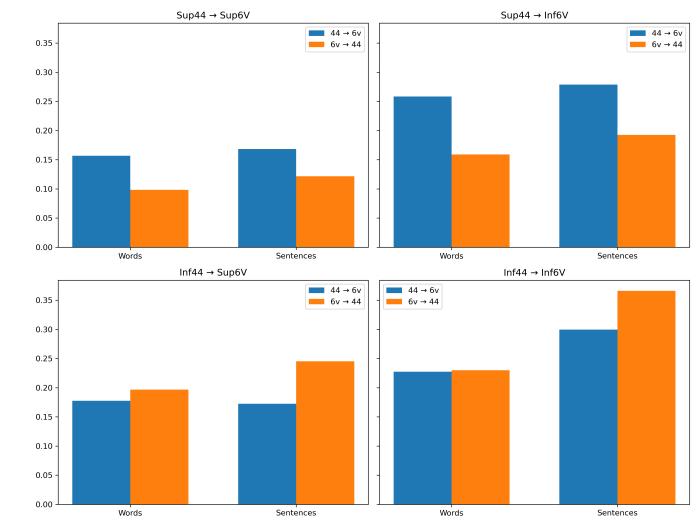


Fig. 10. Proportion of significant inter-region Granger causal interactions at 40ms delay

inf6v.

D. Summary

While premotor activity clearly increases between cued word presentation and speech onset, changes in Broca's are less clear and the overall activity level is lower than that of premotor cortex. In contrast, during sentence production Broca's activity is far greater and much more variable over time than that of premotor cortex. The cause of this variability remains unclear but it suggests that the recorded regions of Broca's area differ substantially in activity between words and sentences. This is supported by histograms of channel activity, in which activity in Broca's area differs little between

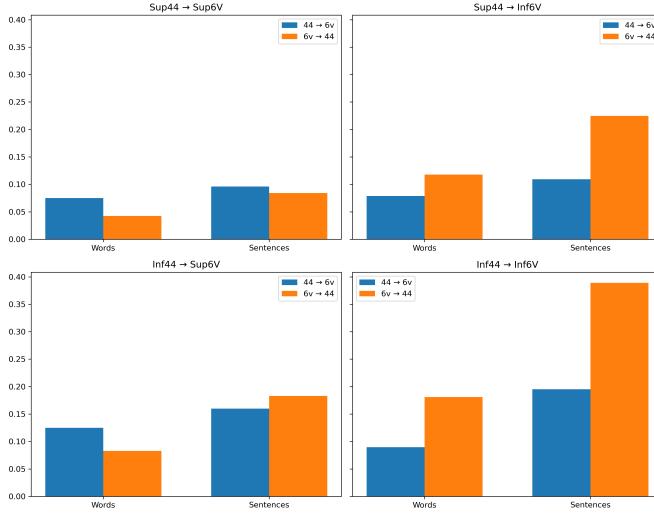


Fig. 11. Proportion of significant inter-region Granger causal interactions at 160ms delay

speaking and delay conditions for words but massively for sentences. Furthermore, Broca's activity on word trials is lower than premotor cortex activity, however activity in Broca's area during sentence production is the highest overall. This is driven by extremely high activity in eight channels.

During word production, the number of channels between Broca's area and premotor cortex with significant Granger causal interactions is roughly equal during speaking and delay periods. Furthermore, the number of feedforward influences is approximately the same as the number of feedback influences. However, there are a much greater number of feedback influences from inf6v to sup44 than feedforward, and much more feedforward than feedback between inf44 to sup6v.

For whole session Granger causality tests, there were more significant interactions at late delays (160ms) for sentences than words. Additionally, there were significantly greater feedforward than feedback GC interactions from sup44 to both BA6v arrays at early time lags. However, at later time lags, there were more GC interactions from inf6v to BA44.

IV. DISCUSSION

Neural recordings with increasing spatial and temporal resolution have enabled fine-grained analyses of Broca's area and its role in speech. The microelectrode array (MEA) data analyzed in the present work offer even higher spatial resolution than ECoG, enabling a more localized assessment of Broca's activity and its interaction with premotor cortex.

Unlike earlier studies, we did not observe robust activation in Broca's area prior to single-word production. This discrepancy likely stems from differences in the recording modality. Flinker et al. [1] and Magrassi et al. [5] used ECoG signals in the sub 150 Hz range, while our analysis focused on frequencies above 250 Hz. Moreover, ECoG electrodes pick up postsynaptic potentials from a larger population of neurons,

likely capturing network dynamics that were not detected with the specific MEAs we analyzed.

The distinction between single-word and sentence production is intriguing. Hickok and Poeppel have previously proposed an auditory dorsal stream, which includes Broca's area and interfaces with the motor system. They further hypothesized that there are multiple levels of organization, including a level for speech segments and a level for sequences of segments. This could explain the staggering difference between activity during single-word production and sentence production. This pattern may indicate that lower-frequency signals (<150 Hz) coordinate speech segments, while higher-frequency signals reflect the coordination of sequences.

Alternatively, the arrays may be preferentially positioned in subregions of Broca's area more tuned to speech sequences than individual segments. Consistent with prior literature, we did not find evidence of Broca's activation overlapping with the articulation of single words. However, the strong activation and functional connectivity observed during sentence production suggest that Broca's area contributes to the online coordination of articulation in longer sequences. This extends previous work that has largely focused on isolated words, and supports the view that Broca's area is specifically involved in sequence processing.

Moreover, Granger causality analysis revealed temporally structured interactions with early feedforward influence from superior BA44 to BA6v, followed by later feedback from inferior BA6v to BA44. While the functional interpretation of this dynamic remains uncertain, it may reflect a recurrent loop involved in planning and adjusting motor sequencing during speech.

Overall, the elevated activity and directional connectivity observed during sentence production support the hypothesis that Broca's area plays a central role in orchestrating speech sequences rather than executing individual words. Future work using broader spatial coverage and targeted behavioral manipulations will be necessary to refine these functional interpretations.

REFERENCES

- [1] A. Flinker, A. Korzeniewska, A. Y. Shestyuk, *et al.*, "Redefining the role of Broca's area in speech," *Proceedings of the National Academy of Sciences*, vol. 112, no. 9, pp. 2871–2875, 2015.
- [2] A. D. Friederici, "The brain basis of language processing: from structure to function," *Physiological Reviews*, vol. 91, no. 4, pp. 1357–1392, 2011.
- [3] N. Geschwind, "The Organization of Language and the Brain: Language disorders after brain damage help in elucidating the neural basis of verbal behavior," *Science*, vol. 170, no. 3961, pp. 940–944, 1970.
- [4] G. Hickok and D. Poeppel, "The cortical organization of speech processing," *Nature Reviews Neuroscience*, vol. 8, no. 5, pp. 393–402, 2007.

- [5] L. Magrassi, G. Aromataris, A. Cabrini, V. Annovazzi-Lodi, and A. Moro, “Sound representation in higher language areas during language generation,” *Proceedings of the National Academy of Sciences*, vol. 112, no. 6, pp. 1868–1873, 2015.
- [6] M. L. Menéndez, J. A. Pardo, L. Pardo, and M. D. C. Pardo, “The Jensen-Shannon divergence,” *Journal of the Franklin Institute*, vol. 334, no. 2, pp. 307–318, 1997.
- [7] S. L. Metzger, K. T. Littlejohn, A. B. Silva, *et al.*, “A high-performance neuroprosthesis for speech decoding and avatar control,” *Nature*, vol. 620, no. 7976, pp. 1037–1046, 2023.
- [8] S. Seabold, J. Perktold, *et al.*, *statsmodels: Statistical modeling and econometrics in Python*, Version 0.14.4, 2010. [Online]. Available: <https://www.statsmodels.org/>.
- [9] J. Tang, A. LeBel, S. Jain, and A. G. Huth, “Semantic reconstruction of continuous language from non-invasive brain recordings,” *Nature Neuroscience*, vol. 26, no. 5, pp. 858–866, 2023.
- [10] F. R. Willett, E. M. Kunz, C. Fan, *et al.*, “A high-performance speech neuroprosthesis,” *Nature*, vol. 620, no. 7976, pp. 1031–1036, 2023.