IberSPEECH 2024
11-13 November 2024, Aveiro, Portugal

Comparative Analysis of Power Dynamics from Invasive EEG Recordings During Overt and Covert Speech Production in Epilepsy Patients

Asma Hasan Sbaih[1], Marc Ouellet[2], José L. Pérez-Córdoba[1], Sneha Raman[3], Ana B. Chica[2], Owais Mujtaba Khanday[1], Alberto Galdón[4], Gonzalo Olivares[4] and Jose A. Gonzalez-Lopez[1]*

[1]Dept. Signal Theory, Telematics and Communications, University of Granada, Spain
[2]Dept. of Experimental Psychology, University of Granada, Spain
[3]Dept. of Biosciences and Bioengineering, Indian Institute of Technology, Mumbai, India
[4]Hospital Universitario Virgen de las Nieves, Granada, Spain
*Corresponding authors: {asmasbaih, mouellet, jlpc, joseangl}@ugr.es

Abstract

In this study, invasive EEG recordings of four epilepsy patients were used to investigate the neural dynamics of speech production. Patients were asked to say aloud or imagining to say aloud Vowel-Consonant-Vowel (VCV) syllables or the name of different pictures. We analyzed the mean power across different frequency bands and across the different conditions. In order to facilitate the recognition of the different patterns, we also employed UMAP (Uniform Manifold Approximation and Projection) visualization. Our findings revealed distinct patterns of power distribution and modulation in each frequency band and for the different tasks, suggesting differential engagement of neural circuits.

Keywords: Epilepsy, Intracranial electrocorticography, imagined speech, Invasive recordings, frequency bands, brain activity, mean power.

1. Introduction

Invasive EEG recordings in epilepsy patients undergoing pre-surgical evaluation provide a unique opportunity to capture high-temporal and -spatial resolutions neural activity [1, 2]. In this study, we analyze stereoelectroencephalography (sEEG) data from four epilepsy patients carrying out different speech production tasks: saying aloud (overt speech) or imagining to say aloud, but without moving the articulators (covert speech), Vowel-Consonant-Vowel (VCV) syllables or the names of pictures from six semantic categories.

The motivation for this study is to gain a deeper understanding of how overt and covert speech production differ in the epileptic brain. Identifying the neural correlates of covert speech holds potential for advancing BCIs that enable communication for individuals unable to speak. Previous research has largely focused on overt speech, leaving a critical gap in understanding covert speech, particularly in epilepsy patients where abnormal brain activity poses additional challenges.

Decoding covert speech from neural activity could be the basis for future brain-computer interface (BCI) devices that help individuals with speech impairments, but with intact cognition, to recover speech. For example, intracortical microelectrode arrays were implanted to a patient suffering from amyotrophic lateral sclerosis, which allowed him to speak again via a synthesizer [3]. The implant that provided the richest information was located over the ventral premotor cortex (area 6v), an area associated with the activation of the muscles involved in articulation.

Unfortunately, one of the disadvantages to do research with epileptic patients is that, for obvious reasons, electrodes locations are chosen according to clinical reasons only [1]. In other words, researchers cannot ask to have the electrodes being located on a specific area of interest and the location of the electrodes differ from one participant to the other. The aim of this study was to assess how much speech-related information can be recovered from neural signals recorded from epileptic patients who have deep brain electrodes implanted for clinical purposes.

In addition, we considered the variations in different EEG frequency bands because it has been demonstrated that variations in specific frequency bands (ranging from the delta to the gamma bands) can be related to particular cognitive and speech processes [4, 5, 6]. Of a particular interest for the purpose of this study, it has been observed that while overt speech is better decoded in the high gamma frequency band, decoding covert speech is better achieved with the activity in lower frequency bands, such as low-beta or theta bands [7]. What is more, the authors of this study also highlighted the potential role of the delta band (1-4 Hz) to design new BCIs. It is for these reasons that we decided to compare the mean power activity across various EEG frequency bands: delta (1-4 Hz); theta (4-8 Hz); alpha (8-12 Hz); beta (13-30 Hz); gamma (30-70 Hz); and high gamma (70-120 Hz).

We expected to uncover specific neural patterns, focusing on power spectral densities on different frequency bands, associated with each of the tasks.

2. Methods

2.1. Participants

Participants were individuals with drug-resistant epilepsy who have undergone implantation of invasive intracranial electrodes to precisely localize seizure onset zones and to map functional areas of the brain for potential surgical intervention. These individuals experience recurrent epileptic seizures, significantly impacting their daily functioning. All participants were adults, had normal or corrected-to-normal vision and no prior experience with the task [8]. All participants were able to comprehend instructions, read, and communicate effectively, enabling their participation in our study. Participants received a monetary compensation of 100€ for their participation. Signed informed consent was collected prior to their inclusion in the study. Participants were informed about their right to withdraw

10.21437/IberSPEECH.2024-8

from the experiment at any time. Our Regional Biomedicine Ethic Research Committee approved the study.
For this work, we selected a sample of four patients (S05, S06, S10, S11) to be included in the analyses (mean age 43.4 ± 9.5 years; 3 females, 1 male).

2.2. Tasks description

Participants were first asked to read aloud or to imagine reading aloud the same list of 50 vowel-consonant-vowel (VCV) pseudowords in each task. The pseudowords were generated by combining the five Castillan Spanish vowels (/a, e, i, o, u/) with ten unvoiced consonants (/p, m, f, t, n, r, s, l, k, x/), especially chosen for their different places of articulation according to the International Phonetic Alphabet (IPA) [9].
Each trial started with the presentation of a fixation cross during 1500 to 2000 ms (Fix), followed by the presentation of the pseudoword during 1500 ms (See) and an inter-stimulus-interval during 1500 to 2000 ms (ISI) preceded the prompt that instructed the participant to read aloud or imagine speaking the pseudoword during 1500 ms (Speech).
In the Picture Naming task, participants named or imagined naming 30 pictures from six semantic categories (Fruits, Body parts, Animals, Musical instruments, Clothes, and Inanimate objects). Each trial started with the presentation of a fixation cross during 1500 to 2000 ms (Fix), followed by the presentation of the picture during 700 ms (See) and ended with the prompt to instruct the participant to say aloud or imagine saying the name of the picture during 2000 ms (Speech).
In the VCV and the Picture Naming tasks, speech production changed between covert and overt every 10 trials. Depending on the participant's fatigue, each task could be repeated a varying number of times. Table 1 and Table 2 show the total number of trials in the VCV and Picture Naming tasks, respectively, per participant.

Table 1: *Number of trials in the VCV task*

Patient_id	Vocabular_size	Repetitions	Total_trials Overt speech	Total trials Covert speech
S05	50	8	400	400
S06	50	6	300	300
S10	50	4	200	200
S11	50	7	350	350

Table 2: *Number of trials in the Picture Naming task*

Patient_id	Number of categories	Images per category	Repetitions	Total trials Overt speech	Total trials Covert speech
S05	6	5	12	360	360
S06	6	5	6	180	180
S10	6	5	4	120	120
S11	6	5	6	180	180

2.3. Recording setup and data processing

Both audio and EEG were recorded while participants performed the experimental task above. Invasive EEG were recorded from the participants using a 256-channel Natus Quantum amplifier (Natus Neuro, Middleton WI) at a 1024Hz sampling rate. Simultaneously, a Blue Yeti microphone (Blue Microphones, California) sampled the voice data at 44.1 kHz. EEG and audio were synchronized using Lab Streaming Layer (LSL)[10].

2.4. Signal processing

The neural activity in each of the following EEG frequency bands was extracted: delta (1-4 Hz), theta (4-8 Hz), alpha (8-12 Hz), beta (13-30 Hz), gamma (30-70 Hz), and high gamma (70-120 Hz). Average power-spectral density (PSD) was computed for each band using Welch's method [11] for the whole duration of each segment of each trial under examination, that is: Fix, See, ISI (for the VCV task only), or Speech (see Section 2.2). The Welch method involves segmenting the EEG signal into overlapping windows, applying a window function to each segment, computing the periodogram for each segment, and then averaging these periodograms to produce a PSD estimate. This method provides a reliable and stable estimate of the power spectrum with reduced variance compared to a single segment fast Fourier transform (FFT) [12]. Thus, for each trial and channel, a single value representing the average power for that segment was obtained. After examination, no significant activity was found on the lower delta band, so it was discarded from the analyses.

3. Analysis & Results

3.1. Statistical Analysis of Mean Power Distribution

To ensure accuracy, outliers in the High Gamma band were filtered out. This involved identifying and removing power values that deviated significantly from the mean, which could be due to artifacts or noise.
Figure 1 shows that the Theta, Beta, and Gamma bands exhibited similar and compact distributions across conditions in the VCV task. In contrast, the Alpha and High Gamma bands showed the most significant differences across the four conditions, with wider and more prominent distributions in the 'Speech' condition, indicating higher power and variability. The Alpha and High Gamma bands' variability in the 'Speech' condition suggests differential neural engagement.

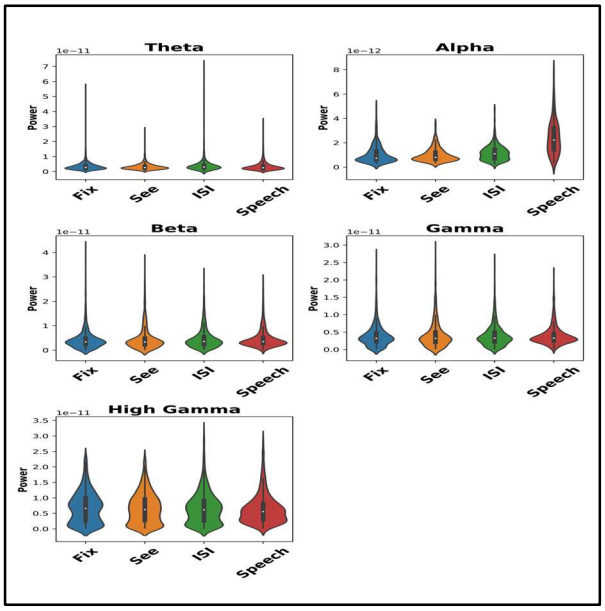

Figure 1*: Mean Power Distribution per frequency band in the four segments of the trials in the VCV task*

Figure 2 illustrates the power distributions of EEG frequency bands for all the patients in the VCV task in the speech segment according to the type of speech: Overt or Covert. Significant differences between the Overt and Covert speech were most notable in the Gamma and High Gamma bands, where the Overt condition showed higher mean power and greater variability. The Alpha band also exhibited higher power and a distinct peak in the Overt speech condition, indicating less variability compared to Covert speech. The Theta and Beta bands showed slightly higher mean power in the Overt speech condition. Overall, the Overt speech condition generally showed higher mean power across all frequency bands, indicating increased changes in neural activity when speaking aloud.

Table 3 shows the mean power distribution of EEG frequency bands for all the participants in the Picture Naming task. The most notable differences are in the Alpha and High Gamma bands in the 'Speech' condition, where power was more prominent and variable. The Theta and Beta bands also showed a great variability, especially in the 'Speech' and 'Fix' conditions. The Gamma band showed similar distributions across conditions with minimal variability. These findings suggest changes in neural activity during speech production.

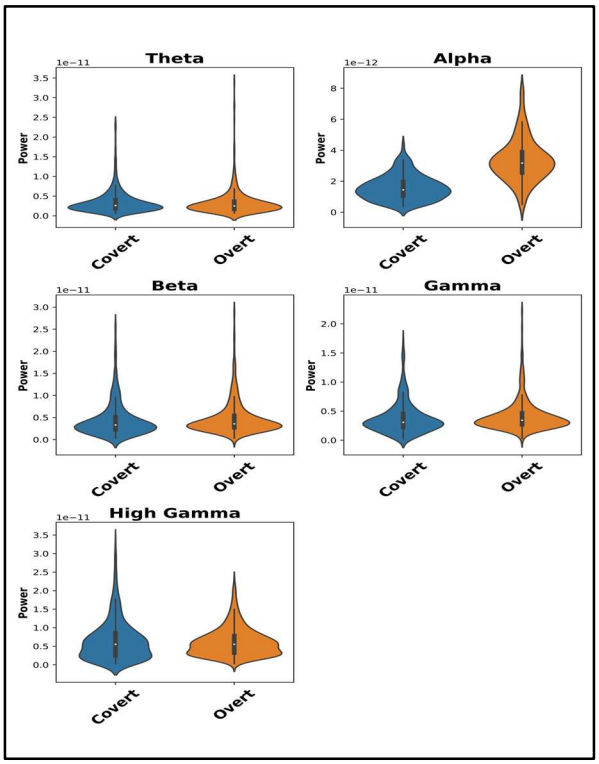

Figure 2: *Mean power distribution per frequency band in the 'Speech' condition between Overt and Covert speech in the VCV task*

Table 3*: Mean power distribution per frequency band in the three segments of the trials in the Picture naming task, where the units of W/Hz (watts per hertz), which indicates the amount of power within a specific frequency bandwidth.*

Frequency Band	Fix (PSD (W/Hz))	Read (PSD (W/Hz))	Speech (PSD (W/Hz))
Theta	$5.38e^{-13}$ to $6.66e^{-11}$	$4.58e^{-13}$ to $2.44e^{-11}$	$5.66e^{-13}$ to $1.79e^{-11}$
Alpha	$3.19e^{-13}$ to $0.67e^{-11}$	$3.27e^{-13}$ to $0.67e^{-11}$	$4.1e^{-13}$ to $0.46e^{-11}$
Beta	$2.17e^{-13}$ to $6.38e^{-11}$	$2.25e^{-13}$ to $4.89e^{-11}$	$2.34e^{-13}$ to $2.13e^{-11}$
Gamma	$1.63e^{-13}$ to $6.24e^{-11}$	$1.75e^{-13}$ to $3.79e^{-11}$	$2.14e^{-13}$ to $1.6e^{-11}$
High-Gamma	$1.38e^{-13}$ to $2.56e^{-11}$	$1.68e^{-13}$ to $2.93e^{-11}$	$2.02e^{-13}$ to $5.36e^{-11}$

Table 4 shows the EEG frequency bands power distributions in the Picture Naming task in both Covert and Overt speech conditions. The Overt condition generally showed higher mean power across all bands, reflecting greater variation in neural activity when speaking aloud. The greater differences were observed in the Gamma and High Gamma bands, where Overt speech exhibited higher peak and mean power, along with greater variability. The Alpha band also showed a distinct peak and higher power in the Overt speech condition, indicating more consistent neural activity when words are spoken aloud. Regarding the Theta and Beta bands, a slightly higher peak was observed in the Overt speech. These findings mirror the results of the VCV task, highlighting increased EEG power when speaking aloud compared to when imagining speech.

Table 4*:Mean power distribution per frequency band in the 'Speech' condition between Overt and Covert speech in the Picture naming task*

Frequency Band	Covert speech Range (PSD Mean Power (W/Hz))	Overt speech Range (PSD Mean Power (W/Hz))
Theta	6.25e−13 to 2.35e−11	6.70e−13 to 3.40e−11
Alpha	3.58e−13 to 4.32e−12	4.98e−13 to 7.85e−12
Beta	2.65e−13 to 2.62e−11	3.58e-13 to 2.91e-11
Gamma	2.10e-13 to 1.74e-11	2.99e-13 to 2.22e-11
High-Gamma	2.21e-13 to 3.35e-11	3.00e-13 to 2.30e-11

3.2. UMAP (Uniform Manifold Approximation and Projection)

Next, we used UMAP, a dimensionality reduction technique useful for visualizing high-dimensional data [13] by clustering EEG data corresponding to the same category of stimuli in a low-dimensional 2D space. Here, UMAP was applied to high gamma features from participant S11 in the VCV and the Picture Naming tasks. As a reference, we also performed UMAP on MFCC features obtained from the audio waveforms. Figure 3 shows vowels discrimination (/a, e, i, o, u/) in the VCV task using MFCC audio data (Figure 3a), EEG features corresponding to the Overt (Figure 3b) or Covert (Figure 3c) speech, or both types of speech together (Figure 3d). While formant frequencies corresponding to the MFCC audio data allowed a vowels clustering that was clearly delineated, EEG data for the Overt speech showed a less defined clustering. Worse yet, a completely diffused classification of the vowels was observed with the EEG data corresponding to the Covert or both types of speech together. These results demonstrate the challenge of classifying vowels using sEEG data recorded from

electrodes placed in regions not typically associated with language processing. It also highlights that vowels classification in Covert speech does not seem to benefit from the high-gamma band information.

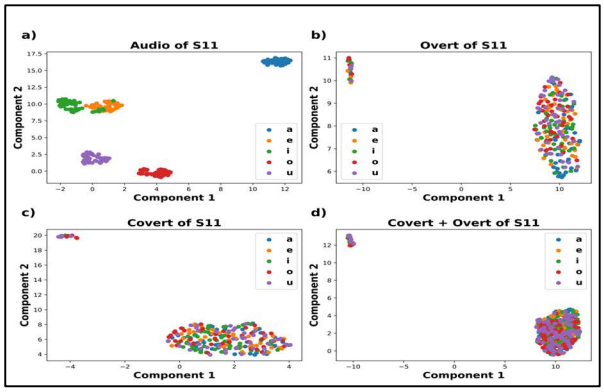

Figure 3: *UMAP visualization of the EEG data of participant S11in the Vowels classification*

Figure 4 shows semantic categories discrimination in the Picture naming task using MFCC audio data (Figure 4a), and EEG features corresponding to the Overt (Figure 4b) or Covert (Figure 4c) speech. The clustering corresponding to the MFCC audio data was identifiable, but poorly delineated. Unfortunately, the classification of the different semantic categories was completely diffused with the Overt or Covert EEG data. Overall, the results suggest that sEEG data from this particular epilepsy patient on the high gamma band cannot be used to classify words into semantic categories.

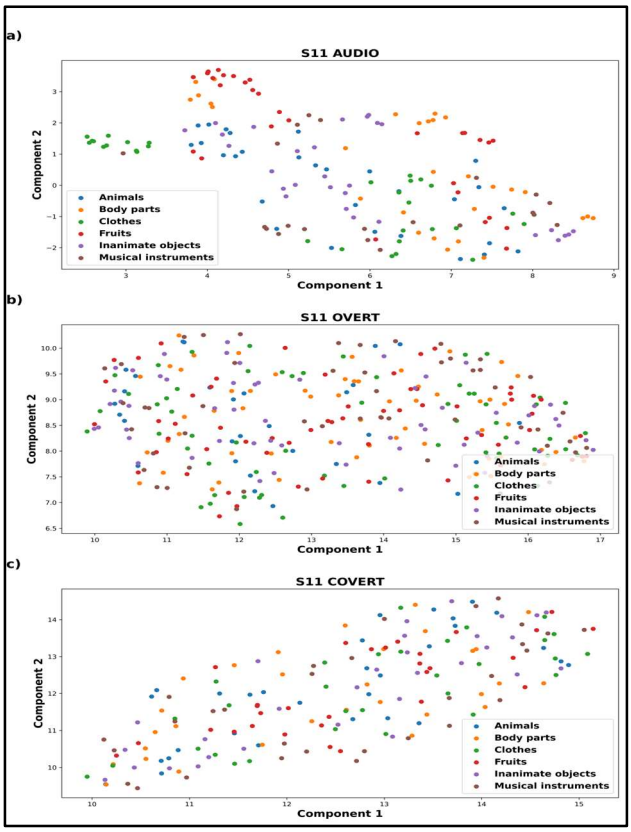

Figure 4: *Visualization Data of S11 Picture Naming classification*

5. Conclusion

EEG power analysis across different speech production tasks revealed consistent neural activity patterns. The theta band showed moderate variability but stable activity across the tasks. The alpha and beta bands exhibited a broader distribution during overt speech tasks compared to imagined speech, indicating distinct neural processes in speech production. The high gamma band showed heightened activity and greater variability during overt speech tasks, highlighting its potential as a cognitive biomarker in speech production. In addition, UMAP visualizations pointed to a possible classification of vowels with high-gamma band data when speaking aloud. Unfortunately, the same classification procedure when imagining speech seems to be impossible. It also seems impossible to classify the semantic categories of words with this type of data.

In conclusion, our study analyzed EEG mean power distribution and utilized UMAP visualization to investigate the neural mechanisms involved in speech production and perception tasks. These findings underscore EEG's potential to differentiate speech conditions and modalities, contributing to our understanding of speech processing's neural correlates and informing advancements in neuroscience, cognitive science, and speech technology. In future work, we plan to deploy state--of-the-art deep learning algorithms to decode speech from the sEEG recordings.

6. Acknowledgements

This work was supported by grant PID2022-141378OB-C22 funded by MCIN/AEI/10.13039/501100011033 and by ERDF/EU.

7. References

[1] J. A. González-López, A. Gómez-Alanis, J. M. Martín Doñas, J. L. Pérez-Córdoba, and A. M. Gomez, "Silent Speech Interfaces for Speech Restoration: A Review," *IEEE Access*, vol. 8, pp. 177995-178021, Sept. 2020, [doi: 10.1109/ACCESS.2020.3026579]

[2] J. P. Lachaux, D. Rudrauf, and P. Kahane, "Intracranial EEG and human brain mapping," *Journal of Physiology-Paris*, vol. 97, no. 4, pp. 613–628, Jul. 2003, [doi: 10.1016/j.jphysparis.2004.01.018]

[3] F.R. Willett, E.M. Kunz, C. Fan, et al., "A high-performance speech neuroprosthesis," *Nature*, vol. 620, pp. 1031–1036, Jan. 2023, [doi: 10.1038/s41586-023-06377-x]

[4] W. Klimesch, "EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis," *Brain Research Reviews*, vol. 29, no. 2, pp. 169–195, Apr. 1999, [doi: 10.1016/S0165-0173(98)00056-3]

[5] E. Başar, C. Başar-Eroglu, S. Karakaş, and M. Schürmann, "Gamma, alpha, delta, and theta oscillations govern cognitive processes," *International journal of psychophysiology*, vol. 39, no. 2–3, pp. 241–248, 2001 [doi: 10.1016/S0167-8760(00)00145-8]

[6] B. Voytek and R. T. Knight, "Dynamic Network Communication as a Unifying Neural Basis for Cognition, Development, Aging, and Disease," *Biological Psychiatry*, vol. 77, no. 12, pp. 1089–1097, Jun. 2015, doi: 10.1016/j.biopsych.2015.04.016.

[7] T. Proix, J. Delgado Saa, A. Christen, et al., "Imagined speech can be decoded from low- and cross-frequency intracranial EEG features," Nature Communications, vol. 13, p. 48, Jan. 2022, [doi: 10.1038/s41467-021-27725-3]

[8] I. Hernáez Rioja *et al.*, "ReSSInt project: voice restoration using Silent Speech Interfaces," in *IberSPEECH 2022*, ISCA, Nov. 2022, pp. 226–230, [doi: 10.21437/IberSPEECH.2022-46]

[9] International Phonetic Association, "The international phonetic alphabet (revised to 2020)," 2020. [Online]. Available: https://www.internationalphoneticassociation.org/IPAcharts/IPA_chart_orig/pdfs/IPA_Kiel_2020_full.pdf

[10] C. Kothe *et al.*, "The Lab Streaming Layer for Synchronized Multimodal Recording," Feb. 14, 2024, *bioRxiv*. doi: 10.1101/2024.02.13.580071.

[11] A. S. Siva, S. G. Rameshkumar, and K. Dhayalini, "Supraharmonic Analysis by Welch's-Power Spectral Density Estimation," in 2024 Third International Conference on Power, Control and Computing Technologies (ICPC2T), Jan. 2024, pp. 357–362. doi: 10.1109/ICPC2T60072.2024.10474764.

[12] P. K. Rahi, R. Mehra, and M. Scholar, "Analysis of Power Spectrum Estimation Using Welch Method for Various Window Techniques," 2014.

[13] L. McInnes, J. Healy, and J. Melville, "UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction," Sep. 17, 2020, *arXiv*: arXiv:1802.03426. doi: 10.48550/arXiv.1802.03426.