Semantic Representations of Speech Production in Intracranial EEG

Citation for published version (APA):

Herff, C., Verwoert, M., Amigó-Vega, J., & Ottenhoff, M. C. (2023). Semantic Representations of Speech Production in Intracranial EEG. In *2023 IEEE International Conference on Systems, Man, and Cybernetics: Improving the Quality of Life, SMC 2023 - Proceedings: Improving the Quality of Life, SMC 2023 - Proceedings* (pp. 4764-4769). IEEE. https://doi.org/10.1109/SMC53992.2023.10394550

Document status and date:
Published: 01/10/2023

DOI:
10.1109/SMC53992.2023.10394550

Document Version:
Publisher's PDF, also known as Version of record

Document license:
Taverne

Semantic Representations of Speech Production in Intracranial EEG

Christian Herff
School for Mental Health and Neuroscience
Maastricht University
Maastricht, the Netherlands
c.herff@maastrichtuniversity.nl

Maxime Verwoert
School for Mental Health and Neuroscience
Maastricht University
Maastricht, the Netherlands
m.verwoert@maastrichtuniversity.nl

Joaquín Amigó-Vega
Computer Science Department
Gran Sasso Science Institute
L'Aquila, Italy
joaquin.amigo@gssi.it

Maarten C. Ottenhoff
School for Mental Health and Neuroscience
Maastricht University
Maastricht, the Netherlands
m.ottenhoff@maastrichtuniversity.nl

***Abstract*—Speech neurprostheses have the potential to provide severely paralyzed patients with a means of communication. To enable the best possible decoding of speech processes from neural data, it is important to chose a representation of speech that is both meaningful to the decoding process and represented well in the neural recordings.**

Previously, acoustic, articulatory, and textual representations of speech have been decoded from neural recordings. Semantic representations of speech could add additional information about the content of the produced speech. In this study, we show that semantic embeddings for individual words, as extracted by a *word2vec*-model, can be used to reconstruct neural activity during speech production across wide-spread cortical and subcortical areas. We elucidate the temporal dynamics of reconstruction quality and show that a slight right hemisphere preference exists. These findings could be used to add semantic information into speech neuroprostheses in the future.

***Index Terms*—stereotactic EEG, Brain-Computer Interfaces, NLP, BCI, semantics, word2vec**

I. Introduction

Speech neuroprostheses [1]–[4] are envisioned to provide a means of communication to patients who lost the ability to speak, due to a stroke or neurodegenerative diseases. Such a speech neuroprosthesis would measure neural activity and decode meaningful representations of speech that can then be reproduced in textual or acoustic form. Consequently, it is crucial to identify meaningful representations of speech that can be decoded from neural activity.

Textual representation of speech that comprises of phonemes that are combined to build words have been decoded from electrocorticography (ECoG) [5]–[9], stereotactic EEG (sEEG) [10] and Utah arrays [11]. Subsequently, these approaches have been implemented in anarthric patients using both ECoG [12] and microelectrode arrays [13].

Instead of a textual representation, other approaches have targeted an articulatory representation of speech, in which the configuration of the articulatory tract, including lips, larynx, tongue and jaw, is decoded. Reliable representations of articulator movements and configurations have been decoded [14] and described [15] in ECoG recordings. Similarly, lip movements, recorded by video, have been decoded from ECoG recordings [16].

A third representation of speech that has been thoroughly studied is acoustics. In this approach, a spectral representation of the audio waveform is usually decoded. The spectrogram is then re-synthesized to an audio waveform by a *Vocoder* that recreates the missing phase information. Several studies have reconstructed spectrograms of speech from ECoG [17]–[20], sEEG [21] and microarrays [11]. A simplified version of this approach has even been used to synthesize imagined speech processes in real-time [22].

These three representations of speech all come with advantages and disadvantages for the use in speech neuroprostheses. A textual representation carries a lot of the information content and the decoding process can be greatly aided by the addition of language models and dictionaries. The use of this additional knowledge sources, however, restricts the expressive power of the user to certain words or common phrases. Furthermore, a textual representation does not carry all the other information of speech, such as intonation, prosody and accentuation [23]. The full expressive power of speech can be realized by decoding an articulatory or acoustic representation, but these high-dimensional representations are more error-prone and cannot benefit from additional knowledge sources as easily. All these previously investigated representations miss a semantic representation of the content of the spoken words. Semantic representation can add to the advantages of the other repre-

C.H. acknowledges funding by the Dutch Research Council (NWO) through the research project 'Decoding Speech In SEEG (DESIS)' with project number VI.Veni.194.021. MV is supported by the INTENSE consortium, which has received funding from the Dutch Research Council (NWO) with Grant Number 17619.

sentation by supplying information about the meaning of the spoken words. Ideally, this could mean that words might not be perfectly decoded, but semantically similar words can be extracted. Potentially, this could be used for expressive speech BCIs or in combination with one of the other approaches.

Semantic representations of perceived natural sounds have been investigated by linking different intermediate-to-semantic level features, extracted using different embeddings models from Natural Language Processing (NLP) to sound-to-event, to recorded fMRI activity [24]. In [25], the authors used a *wav2vec*-model to provide embeddings of audio to explain the cortical fMRI responses to perceived speech. In an ECoG study, Goldstein et al. [26] used the embeddings of the large language model GPT2 to predict high-gamma time series of electrodes to perceived audio books. In another study employing GPT2, Cai et al. [27] modelled cortical activity in sEEG electrodes in natural conversations. By employing natural conversations, instead of simple listening conditions, the author demonstrated that neural activity can not only be predicted during listening, but also during speech production. Furthermore, the authors provide insights into the transition periods between speech perception and production.

In this study, we investigate if semantic representations of word production can be found in stereotactic EEG even when only individual words are produced. In single word production, words are not embedded into meaningful sentences and thus lack a lot of the semantic grounding. Stereotactic EEG is ideally suited for this investigation as it supplies a sparse coverage across cortical and sub-cortical areas [28], which is necessary for the expected scattered representation of semantic encoding [29].

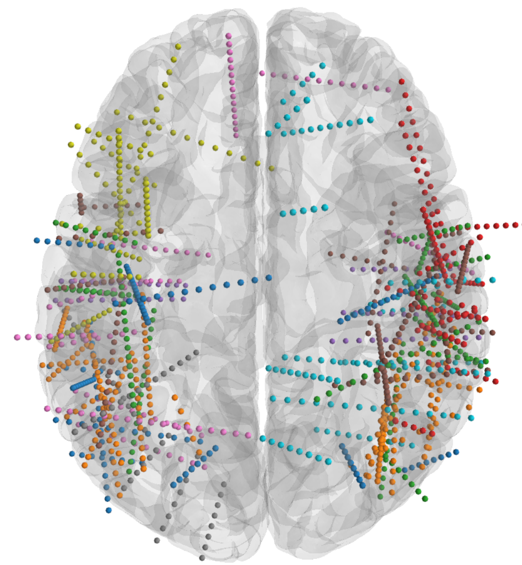

Fig. 1. Electrode locations for all 10 participants warped to a standard MNI brain. Each color indicates locations for one participant. Colors match colors in the results Fig 3**a**.

II. Material and Methods

A. Participants

Ten patients (5 female, 5 male) with intractable epilepsy between 16 and 50 years of age participated in our experiment. Patients underwent surgery in the clinical treatment for their epilepsy. Participation in the experiment occured on a voluntary basis and participants provided written informed consent. The experiment received ethical approval from both Institutional Review Boards at Maastricht Unviversity and Epilepsy Center Kempenhaeghe.

B. Experiment Design

Participants read out individual words shown on a computer screen. Each word was displayed for 2 seconds followed by a relaxation interval of 1 second. Words were taken from the IFA dutch corpus [30] enriched with the numbers one through ten. A total of 100 randomly chosen words were recorded, resulting in a total dataset length of 300 seconds per participant.

C. Data Recording

Each participant was implanted with between 5 and 19 platinum-iridium stereotactic EEG electrode shafts (Microdeep intracerebral electrodes; Dixi Medical, Beçanson, France). Electrode locations are purely determined by clinical necessity and are not influenced by the experiment at all. Each shaft (diameter of 0.8 mm) has a varying number of electrode contacts with a length of 2mm each, resulting in a total of between 54 and 127 electrode contacts per participant. Combined, we analyzed data from 1103 contacts (543 in the left and 557 in the right hemisphere). Electrode data was recorded using Micromed SD LTM amplifiers (Micromed S.p.A., Treviso, Italy) with a sampling rate of 1024 or 2048 Hz. Simultaneously, we recorded audio data from the build-in microphone of the recording notebook. Experimental timing, audio data and neural data were synchronized using Lab-StreamingLayer [31]. Audio data was not analyzed for this study. We previously made all data used in this study available [32].

D. Electrode Localization

Electrodes were localized by co-registering a pre-surgical MR (T1-weighted) and post-surgical CT. The MR image is then parcellated into cortical areas using the Destrieux atlas [33] in *freesurfer* [34]. Electrode locations are marked in the CT using *img_pipe* [35]. This yields anatomical labels for each electrode contact. Cortical meshes and electrode locations are then warped to a standard MNI-Brain for joint visualization (see Fig. 1).

E. Signal Processing

We focused our analysis on the broadband gamma (70-170 Hz) range, as it is known to exhibit highly localized information for a variety of cognitive processes, including music perception and imagination [36], speech [37] and language [38] tasks and is thought to reflect ensemble spiking [39]. To extract broadband gamma, we bandpass filtered the signals

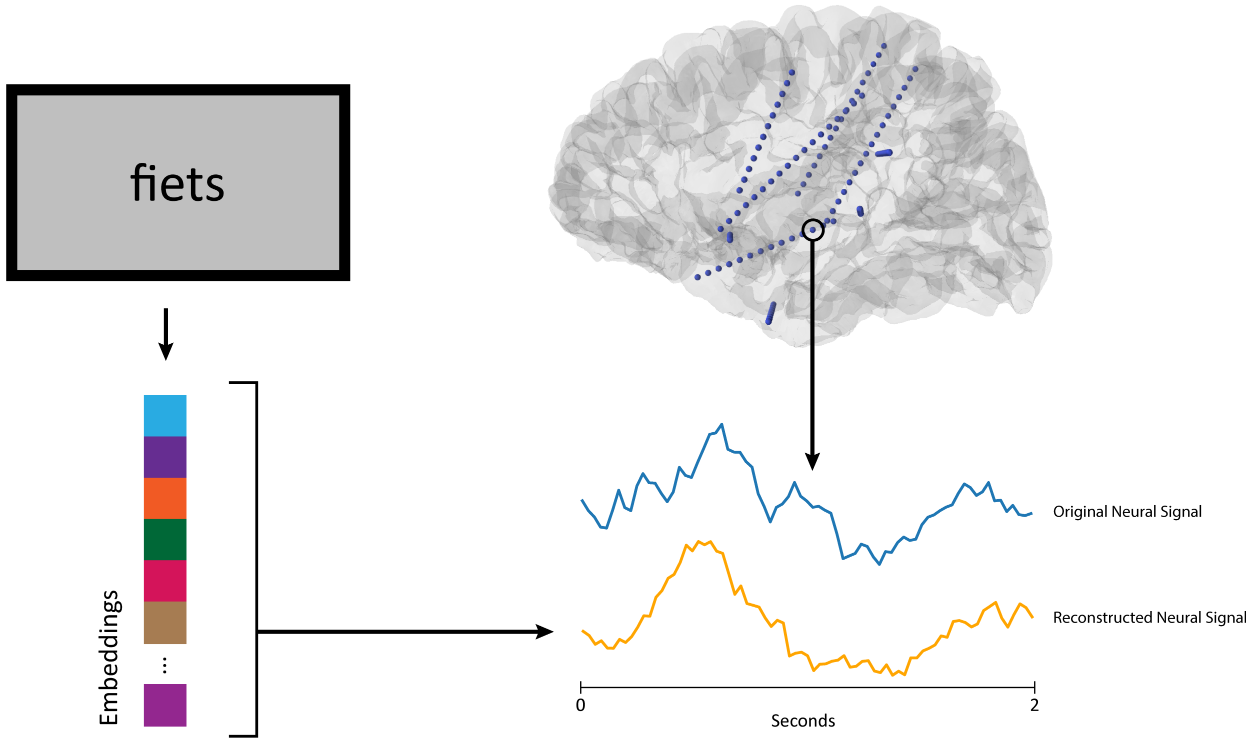

Fig. 2. **Study overview:** Words are presented to a participant, who is then asked to say the word aloud. We record intracranial EEG from stereotactic electrodes during this word production. Word embeddings are generated for the presented words. We train a prediction model to reconstruct neural activity during speech production based on these word embeddings. For a new word, the neural activity can then be predicted from the word embedding (orange time series) and compared to the original neural activity (blue time series).

between 70 and 170 Hz using a zero-phase IIR bandpass filter (filter order 4). Subsequently, we attenuated the first and second harmonic of the 50 Hz line noise by applying two IIR bandstop filters (filter order 4). Neural features were then extracted by averaging the absolute of the Hilbert transform in windows of 200 ms with a window shift of 25 ms.

For each spoken word, we extracted the entire speech production period (2 seconds) and assigned it to the corresponding cued word. This resulted in 75 time points for each electrode contact in each trial.

F. Semantic Embeddings

Semantic representations of the Dutch words were generated by a *word2vec*-model [40]–[42] trained on a dump of the Dutch wikipedia with 392 million words. The *word2vec* embedding provides a numerical representation of Dutch words that captures the semantic and syntactic qualities of words and ensures that semantically similar words have smaller cosine-similarities than non-similar words. The trained model is publicly available [43] and provides a 160 dimensional embedding vector for a given Dutch word. Word embeddings were available for 85 out of the 100 words produced by each participant. We discarded words without a word embedding available.

G. Encoding Model

We trained an ordinary least squares linear regression to predict the time series of a single electrode contact from the embedding of the cued word (see Fig. 2). This means that a 160 dimensional embedding vector is multiplied with a matrix β with shape $[160, 75]$ to predict the broadband gamma activity of one electrode contact during speech production.

H. Evaluation

We train and evaluate the linear regression predictor for each electrode contact individually in a 10-fold cross validation. We then calculate the Pearson correlation across all trials for each individual time point relative to word presentation. This means that we are looking at the differences in average response to word presentation in each time point as opposed to looking at the average response during speech production.

I. Randomized Baseline

We evaluated chance level correlations by randomly permuting the embedding vectors and training the encoding model again exactly as described previously. This procedure was repeated 1,000 times for each channel and the 95% largest correlation across all time points and channels was used as the significance threshold ($\alpha = 0.05$). As we take the largest

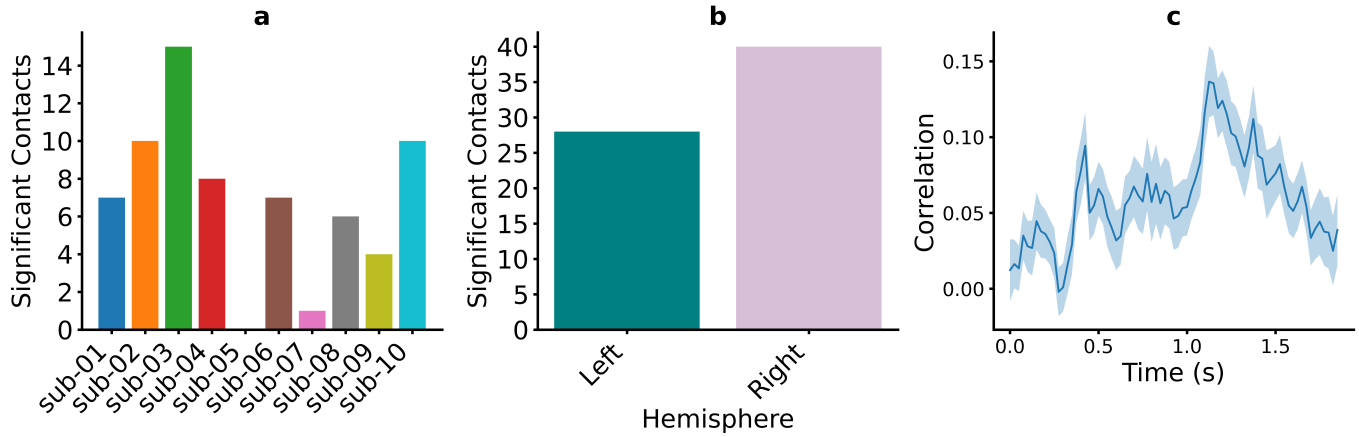

Fig. 3. **Results: a** Number of significant channels for each participant. **b** Number of significant contacts per hemisphere for all participants combined. **c** Average Correlations of significant channels over time. Shaded areas indicate standard error of the mean.

correlation across all channels and time points, we correct for multiple comparisons using the max-t correction [44].

III. Results

A. Contacts encoding semantics in almost all participants

The encoding model could predict neural activity during speech production for at least one contact in all but one participant (Fig. 3**a**). This shows that despite the very different electrode locations (Fig 1) and sparse and distributed sampling of stereotactic EEG, semantic information can be found. Previous studies in fMRI have also found semantic representations, albeit in speech perception, tiling the entire cerebral cortex [29]. In participant 3, the activity of 12% of all electrode contacts could be reconstructed with significant correlations.

B. More contacts with semantic encoding in the right hemisphere

Our participants had very balanced coverage of the left and right hemisphere (543 in the left versus 557 in the right hemisphere), but 42% more contacts were significant in the right hemisphere (28 left, 40 right, see Fig. 3**b**). It is interesting to see this right hemisphere dominance, despite the fact that speech production is lateralized to the left in most people [45]. This indicates that the semantic processing is, at least in part, independent of the actual speech production process.

C. Correlations peak late during speech production

To identify when reconstructed neural activity is most similar to actual activity during speech production, we look at the average Pearson correlation across all significant electrode contacts (Fig. 3**c**). Average correlation has a first peak 500 ms after the word cue and a second, larger peak, one second after the word is shown on the screen. This indicates that neural responses to semantics are present even after word production, as participants are usually finished with speaking at that time.

D. Semantic encoding can be found across many cortical areas

To identify where channels with significant encoding are located, we visualize them in a joint MNI model (Fig 4). Significant channels can be found in all lobes on both hemispheres. There is a clear accumulation of significant channels around the central sulcus and in superior temporal areas, which are known to be involved in speech production [46], [47]. Interestingly, the activity of a number of contacts in occipital and parietal areas can also be reconstructed with significant correlations. This, once again, points to a wide-spread representation of semantics as also found in fMRI [29].

IV. Discussion

Semantic representations of individual words can be used to reconstruct the neural activity during speech production measured by stereotactic EEG. For the semantic representations, we used word embeddings generated by a *word2vec*-model. Interestingly, the neural activity can still be reconstructed despite the fact that the words are not grounded into a wider semantic context (i.e. in full sentences). This indicates how semantic processing occurs in a wide-spread cortical network even when semantic processing is not required by the task.

We find electrodes with significant reconstruction in all but one participant and observe a wide-spread representation across many cortical and sub-cortical areas. This is in accordance with prior studies into semantic encoding [29]. Even occipital and parietal contacts could be reconstructed with significant correlations. This indicates that semantic processing might already start with the reading of the visually presented words.

We observe a slight preference for the right hemisphere, while speech production is traditionally thought to be left hemisphere dominant. This indicates that semantic processing recruits additional neural circuitry to mere speech production.

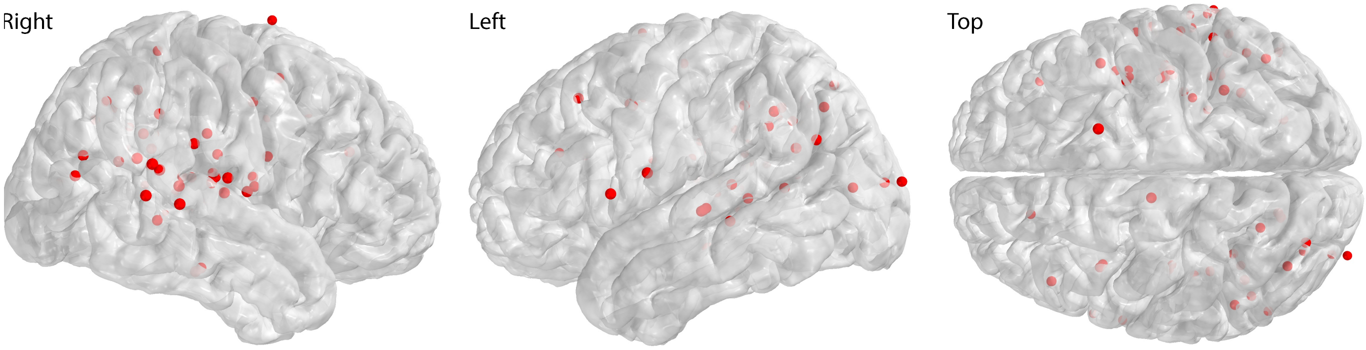

Fig. 4. Locations of channels with significant encoding of semantic information in right, left and top views. Non-significant channels are omitted in this visualization. Channels with significant encoding can be found across a variety of cortical and sub-cortical areas and not only in regions traditionally associated with speech production.

From a temporal perspective, we observe that semantic representation seem to be best encoded shortly after the word cue and about one second after word production, which indicates that semantic processing takes place even after the actual word production.

In the future, it would be of interest to compare these findings to the encoding of acoustic, textual and articulatory features to see which regions overlap and which are distinctively encoding certain aspects of the speech production process. Potentially, a clear hierarchy in the encoding of these processes can be identified, as has been previously done in the processing of articulated, whispered and imagined speech [48].

The fact that semantic representations can indeed be found in intracranial recordings points to a potential role in speech neuroprostheses. Semantic embeddings could be predicted, instead of used for an encoding model, in the future to inform speech neuroprostheses about an additional dimension of intended speech production.

V. Conclusion

We have shown that word embeddings of individual words can be used to predict the neural activity during the speech production process of these words. These findings have the potential to add additional information to the decoding process in speech neuroprosthesis.

References

[1] F. Bocquelet, T. Hueber, L. Girin, S. Chabardès, and B. Yvert, "Key considerations in designing a speech brain-computer interface," *Journal of Physiology-Paris*, vol. 110, no. 4, pp. 392–401, 2016.

[2] Q. Rabbani, G. Milsap, and N. E. Crone, "The potential for a speech brain–computer interface using chronic electrocorticography," *Neurotherapeutics*, vol. 16, no. 1, pp. 144–165, 2019.

[3] C. Herff and T. Schultz, "Automatic speech recognition from neural signals: a focused review," *Frontiers in neuroscience*, vol. 10, 2016.

[4] S. Chakrabarti, H. M. Sandberg, J. S. Brumberg, and D. J. Krusienski, "Progress in speech decoding from the electrocorticogram," *Biomedical Engineering Letters*, vol. 5, no. 1, pp. 10–21, 2015.

[5] C. Herff, D. Heger, A. de Pesters, D. Telaar, P. Brunner, G. Schalk, and T. Schultz, "Brain-to-text: decoding spoken phrases from phone representations in the brain," *Frontiers in neuroscience*, vol. 9, 2015.

[6] E. M. Mugler, J. L. Patton, R. D. Flint, Z. A. Wright, S. U. Schuele, J. Rosenow, J. J. Shih, D. J. Krusienski, and M. W. Slutzky, "Direct classification of all american english phonemes using signals from functional speech motor cortex," *Journal of neural engineering*, vol. 11, no. 3, p. 035015, 2014.

[7] D. A. Moses, N. Mesgarani, M. K. Leonard, and E. F. Chang, "Neural speech recognition: continuous phoneme decoding using spatiotemporal representations of human cortical activity," *Journal of neural engineering*, vol. 13, no. 5, p. 056004, 2016.

[8] J. G. Makin, D. A. Moses, and E. F. Chang, "Machine translation of cortical activity to text with an encoder–decoder framework," Nature Publishing Group, Tech. Rep., 2020.

[9] P. Sun, G. K. Anumanchipalli, and E. F. Chang, "Brain2char: a deep architecture for decoding text from brain recordings," *Journal of neural engineering*, vol. 17, no. 6, p. 066015, 2020.

[10] A. Petrosyan, A. Voskoboinikov, D. Sukhinin, A. Makarova, A. Skalnaya, N. Arkhipova, M. Sinkin, and A. Ossadtchi, "Speech decoding from a small set of spatially segregated minimally invasive intracranial eeg electrodes with a compact and interpretable neural network," *Journal of Neural Engineering*, vol. 19, no. 6, p. 066016, 2022.

[11] G. H. Wilson, S. D. Stavisky, F. R. Willett, D. T. Avansino, J. N. Kelemen, L. R. Hochberg, J. M. Henderson, S. Druckmann, and K. V. Shenoy, "Decoding spoken english phonemes from intracortical electrode arrays in dorsal precentral gyrus," *bioRxiv*, 2020.

[12] D. A. Moses, S. L. Metzger, J. R. Liu, G. K. Anumanchipalli, J. G. Makin, P. F. Sun, J. Chartier, M. E. Dougherty, P. M. Liu, G. M. Abrams *et al.*, "Neuroprosthesis for decoding speech in a paralyzed person with anarthria," *New England Journal of Medicine*, vol. 385, no. 3, pp. 217–227, 2021.

[13] F. R. Willett, E. Kunz, C. Fan, D. Avansino, G. Wilson, E. Y. Choi, F. Kamdar, L. R. H. Hochberg, S. Druckmann, K. Shenoy *et al.*, "A high-performance speech neuroprosthesis," *bioRxiv*, pp. 2023–01, 2023.

[14] E. M. Mugler, M. C. Tate, K. Livescu, J. W. Templer, M. A. Goldrick, and M. W. Slutzky, "Differential representation of articulatory gestures and phonemes in precentral and inferior frontal gyri," *Journal of Neuroscience*, vol. 38, no. 46, pp. 9803–9813, 2018.

[15] J. Chartier, G. K. Anumanchipalli, K. Johnson, and E. F. Chang, "Encoding of articulatory kinematic trajectories in human speech sensorimotor cortex," *Neuron*, vol. 98, no. 5, pp. 1042–1054, 2018.

[16] S. Lesaja, C. Herff, G. D. Johnson, J. J. Shih, T. Schultz, and D. J. Krusienski, "Decoding lip movements during continuous speech using electrocorticography," in *2019 9th International IEEE/EMBS Conference on Neural Engineering (NER)*. IEEE, 2019, pp. 522–525.

[17] G. K. Anumanchipalli, J. Chartier, and E. F. Chang, "Speech synthesis from neural decoding of spoken sentences," *Nature*, vol. 568, no. 7753,

pp. 493–498, 2019. [Online]. Available: https://doi.org/10.1038/s41586-019-1119-1
[18] J. Berezutskaya, Z. V. Freudenburg, M. J. Vansteensel, E. J. Aarnoutse, N. F. Ramsey, and M. A. van Gerven, "Direct speech reconstruction from sensorimotor brain activity with optimized deep learning models," *bioRxiv*, pp. 2022–08, 2022.
[19] C. Herff, L. Diener, M. Angrick, E. Mugler, M. C. Tate, M. A. Goldrick, D. J. Krusienski, M. W. Slutzky, and T. Schultz, "Generating natural, intelligible speech from brain activity in motor, premotor, and inferior frontal cortices," *Frontiers in Neuroscience*, vol. 13, p. 1267, 2019.
[20] M. Angrick, C. Herff, E. Mugler, M. C. Tate, M. W. Slutzky, D. J. Krusienski, and T. Schultz, "Speech synthesis from ecog using densely connected 3d convolutional neural networks," *Journal of neural engineering*, 2019.
[21] J. Kohler, M. C. Ottenhoff, S. Goulis, M. Angrick, A. J. Colon, L. Wagner, S. Tousseyn, P. L. Kubben, and C. Herff, "Synthesizing speech from intracranial depth electrodes using an encoder-decoder framework," *Neurons, Behavior, Data analysis, and Theory*, 2022.
[22] M. Angrick, M. C. Ottenhoff, L. Diener, D. Ivucic, G. Ivucic, S. Goulis, J. Saal, A. J. Colon, L. Wagner, D. J. Krusienski *et al.*, "Real-time synthesis of imagined speech processes from minimally invasive recordings of neural activity," *Communications biology*, vol. 4, no. 1, pp. 1–10, 2021.
[23] C. Herff, G. Johnson, L. Diener, J. Shih, D. Krusienski, and T. Schultz, "Towards direct speech synthesis from ECoG: A pilot study," in *Engineering in Medicine and Biology Society (EMBC), 2016 IEEE 38th Annual International Conference of the*. IEEE, 2016, pp. 1540–1543.
[24] B. L. Giordano, M. Esposito, G. Valente, and E. Formisano, "Intermediate acoustic-to-semantic representations link behavioral and neural responses to natural sounds," *Nature Neuroscience*, pp. 1–9, 2023.
[25] A. R. Vaidya, S. Jain, and A. Huth, "Self-supervised models of audio effectively explain human cortical responses to speech," in *International Conference on Machine Learning*. PMLR, 2022, pp. 21 927–21 944.
[26] A. Goldstein, Z. Zada, E. Buchnik, M. Schain, A. Price, B. Aubrey, S. A. Nastase, A. Feder, D. Emanuel, A. Cohen *et al.*, "Shared computational principles for language processing in humans and deep language models," *Nature neuroscience*, vol. 25, no. 3, pp. 369–380, 2022.
[27] J. Cai, A. E. Hadjinicolaou, A. C. Paulk, Z. M. Williams, and S. S. Cash, "Natural language processing models reveal neural dynamics of human conversation," *bioRxiv*, pp. 2023–03, 2023.
[28] C. Herff, D. J. Krusienski, and P. Kubben, "The potential of stereotactic-eeg for brain-computer interfaces: Current progress and future directions," *Frontiers in Neuroscience*, vol. 14, p. 123, 2020.
[29] A. G. Huth, W. A. De Heer, T. L. Griffiths, F. E. Theunissen, and J. L. Gallant, "Natural speech reveals the semantic maps that tile human cerebral cortex," *Nature*, vol. 532, no. 7600, pp. 453–458, 2016.
[30] R. J. Van Son, D. Binnenpoorte, H. v. d. Heuvel, and L. Pols, "The ifa corpus: a phonemically segmented dutch" open source" speech database," in *7th European Conference on Speech Communication and Technology*. Aalborg, Denmark:[Sn], 2001.
[31] C. Kothe, "Lab streaming layer (lsl)," *https://github.com/sccn/labstreaminglayer*, vol. 26, p. 2015, 2014.
[32] M. Verwoert, M. C. Ottenhoff, S. Goulis, A. J. Colon, L. Wagner, S. Tousseyn, J. P. van Dijk, P. L. Kubben, and C. Herff, "Dataset of speech production in intracranial electroencephalography," *Scientific data*, vol. 9, no. 1, p. 434, 2022.
[33] C. Destrieux, B. Fischl, A. Dale, and E. Halgren, "Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature," *NeuroImage*, vol. 53, no. 1, pp. 1–15, 2010. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811910008542
[34] B. Fischl, "Freesurfer," *Neuroimage*, vol. 62, no. 2, pp. 774–781, 2012.
[35] L. S. Hamilton, D. L. Chang, M. B. Lee, and E. F. Chang, "Semi-automated anatomical labeling and inter-subject warping of high-density intracranial recording electrodes in electrocorticography," *Frontiers in Neuroinformatics*, vol. 11, p. 62, 2017.
[36] S. A. Herff, C. Herff, A. J. Milne, G. D. Johnson, J. J. Shih, and D. J. Krusienski, "Prefrontal high gamma in ecog tags periodicity of musical rhythms in perception and imagination," *Eneuro*, vol. 7, no. 4, 2020.
[37] N. Crone, L. Hao, J. Hart, D. Boatman, R. Lesser, R. Irizarry, and B. Gordon, "Electrocorticographic gamma activity during word production in spoken and sign language," *Neurology*, vol. 57, no. 11, pp. 2045–2053, 2001.
[38] V. L. Towle, H.-A. Yoon, M. Castelle, J. C. Edgar, N. M. Biassou, D. M. Frim, J.-P. Spire, and M. H. Kohrman, "Ecog gamma activity during a language task: differentiating expressive and receptive speech areas," *Brain*, vol. 131, no. 8, pp. 2013–2027, 2008.
[39] S. Ray, N. E. Crone, E. Niebur, P. J. Franaszczuk, and S. S. Hsiao, "Neural correlates of high-gamma oscillations (60–200 hz) in macaque local field potentials and their potential implications in electrocorticography," *Journal of Neuroscience*, vol. 28, no. 45, pp. 11 526–11 536, 2008.
[40] R. Řehřek and P. Sojka, "Software framework for topic modelling with large corpora," 2010.
[41] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," *arXiv preprint arXiv:1301.3781*, 2013.
[42] R. Řehřek, P. Sojka *et al.*, "Gensim—statistical semantics in python," *Retrieved from genism. org*, 2011.
[43] S. Tulkens, C. Emmery, and W. Daelemans, "Evaluating unsupervised dutch word embeddings as a linguistic resource," in *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016)*, N. C. C. Chair), K. Choukri, T. Declerck, M. Grobelnik, B. Maegaard, J. Mariani, A. Moreno, J. Odijk, and S. Piperidis, Eds. Paris, France: European Language Resources Association (ELRA), may 2016.
[44] P. H. Westfall and S. S. Young, *Resampling-based multiple testing: Examples and methods for p-value adjustment*. John Wiley & Sons, 1993, vol. 279.
[45] M. Annett, "Hand preference and the laterality of cerebral speech," *Cortex*, vol. 11, no. 4, pp. 305–328, 1975.
[46] J. A. Tourville and F. H. Guenther, "The diva model: A neural theory of speech acquisition and production," *Language and cognitive processes*, vol. 26, no. 7, pp. 952–981, 2011.
[47] G. Hickok, "Computational neuroanatomy of speech production," *Nature Reviews Neuroscience*, vol. 13, no. 2, p. 135, 2012.
[48] P. Z. Soroush, C. Herff, S. K. Ries, J. J. Shih, T. Schultz, and D. J. Krusienski, "The nested hierarchy of overt, mouthed, and imagined speech activity evident in intracranial recordings," *NeuroImage*, vol. 269, p. 119913, 2023.