Journal of Neuroscience Methods 311 (2019) 253–258

Contents lists available at ScienceDirect

Journal of Neuroscience Methods

journal homepage: www.elsevier.com/locate/jneumeth

Semantic-hierarchical model improves classification of spoken-word evoked electrocorticography

Youngmin Na[a,1], Inyong Choi[b,1], Dong Pyo Jang[c], Joong Koo Kang[d], Jihwan Woo[a,*]

[a] Department of Biomedical Engineering, University of Ulsan, Ulsan, Republic of Korea
[b] Department of Communication Sciences and Disorders, University of Iowa, Iowa City, IA, USA
[c] Department of Biomedical Engineering, Hanyang University, Seoul, Republic of Korea
[d] LTG Neuro Medical Center, Seoul, Republic of Korea

ARTICLE INFO

Keywords:
Electrocorticography
Brain computer interface
Decoding words
Semantic hierarchical structure

ABSTRACT

Classification of spoken word-evoked potentials is useful for both neuroscientific and clinical applications including brain-computer interfaces (BCIs). By evaluating whether adopting a biology-based structure improves a classifier's accuracy, we can investigate the importance of such structure in human brain circuitry, and advance BCI performance. In this study, we propose a semantic-hierarchical structure for classifying spoken word-evoked cortical responses. The proposed structure decodes the semantic grouping of the words first (e.g., a body part vs. a number) and then decodes which exact word was heard. The proposed classifier structure exhibited a consistent ~10% improvement of classification accuracy when compared with a non-hierarchical structure. Our result provides a tool for investigating the neural representation of semantic hierarchy and the acoustic properties of spoken words in human brains. Our results suggest an improved algorithm for BCIs operated by decoding heard, and possibly imagined, words.

1. Introduction

When we listen to speech sounds, our brains process the information from the sound over multi-level hierarchies (Hickok and Poeppel, 2007). Past studies have revealed anatomical structures corresponding to the neural representation of spoken words using various techniques including functional magnetic resonance imaging (Hickok and Poeppel, 2007), electroencephalography (EEG) (Khalighinejad et al., 2017; Vanthornhout et al., 2018), and electrcorticography (ECoG) (Chang et al., 2011; Khoshkhoo et al., 2018). Recent studies have demonstrated that it is possible to use cortical neural responses to "decode" what speech sound (e.g., which word of a closed set) was heard (Crosse et al., 2016; Di Liberto et al., 2018; Overath et al., 2015; Pasley et al., 2012).

Spoken word decoding is a useful tool for investigating the neuroscience of speech processes. By finding what structures of the brain decoder and which features of cortical responses affect the accuracy of the decoding, it is possible to reveal information about the crucial circuitry and mechanisms underlying brain processing of speech sounds. Because spoken-word decoders can also be used in brain-computer interfaces (BCIs), improving their accuracy could provide clinical benefits, for example in decoding the intent of paralyzed patients.

Recently, Huth et al. (2016) concluded that a human brain semantically processes spoken words with a broadly distributed but well-conserved network of semantic classification areas in the human cortex. de Heer et al., (2017) demonstrated that the superior temporal sulcus area reveals the semantic class of a spoken word, whereas medial primary auditory cortex represents low-level speech characteristics such as phonetic features.

We hypothesized that a spoken-word decoder that adopts a semantic-hierarchical structure may yield better decoding performance than a non-hierarchically structured classifier. In this study, we propose a novel two-step approach for decoding ECoG signals based on semantic hierarchy of speech. By testing whether and to what extent the decoding accuracy is improved by adopting this sematic hierarchical structure, we can provide empirical evidence of the importance of sematic processing in the human brain, and its potential application in BCIs.

* Corresponding author.
E-mail addresses: nayoungmin39@gmail.com (Y. Na), inyong-choi@uiowa.edu (I. Choi), dongpjang@gmail.com (D.P. Jang), kangszkr@gmail.com (J.K. Kang), jhwoo@ulsan.ac.kr (J. Woo).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.jneumeth.2018.10.034
Received 22 March 2018; Received in revised form 22 October 2018; Accepted 26 October 2018
Available online 30 October 2018

Table 1
Demographic information.

Subject	Age	Sex	Language dominancy	Grid location	Seizure duration (years)	Diagnosis	Analysis area
S1	42	M	R	R-frontal	8	Right frontal tumor	Broca
S2	19	M	L	R-parietal	6	Right parietal	Auditory
S3	35	M	L	L-temporal	18	Left temporal	Wernicke
S4	26	M	L	L-frontal	12	Left frontal	Broca
S5	38	M	L	L-temporal, L-posterior-superior-temporal, L-posterior-inferior-temporal	26	Left frontal, Left temporal	Broca, Wernicke
S6	33	M	L	R-temporal	24	Right frontal, Right temporal	Auditory

2. Materials and methods

2.1. Subjects and data collection

All study procedures were reviewed and approved by the Asan Medical Center Institutional Review Board, and all work was performed in accordance with the Code of Ethics of the World Medical Association. Six subjects diagnosed with epilepsy participated in this study (Table 1). These subjects underwent clinically-indicated ECoG monitoring using implanted subdural electrodes to localize a seizure focus. The electrode grids were placed over either the left hemisphere (for subjects S3, S4, and S5) or right hemisphere (subjects S1, S2, and S6). Fig. 1 shows the cortical maps representing the electrode locations for each subject. The electrode positions were identified in Talairach coordinates using Curry Software (Compumedics Neuroscan Ltd., Australia) based on CT and MRI images. These positions were then marked on the MNI (Montreal Neurological Institute, Canada) average brain using the Matlab Surfstat toolbox (The Mathworks, Inc., USA) (Worsley et al., 2009). Each grid placed in a subject included 20–126 electrodes, the optimal positions of which were determined using clinical criteria. This led to 530 available electrodes from 6 subjects. Due to limitations on the complexity of data analysis, we selected only ECoG signals from the primary auditory cortex, Broca's area, or Wernicke's area, as identified through cortical stimulation mapping by a neurosurgeon. The clinically-mapped electrodes are shown in green in Fig. 1 and listed in Table 1.

As all subjects were native Korean speakers, eight monosyllabic Korean words /귀gui/ (meaning ear), /눈nun/ (eye), /코ko/ (nose), /목mog/ (neck), /삼sam/ (three), /오o/ (five), /구gu/ (nine), and /십sib/ (ten) were used. These words were all naturally spoken words by a professionally-trained voice actor. The words were categorized into two semantic groups: the first four words are parts of the face, while the latter four words are numbers. Word durations ranged from 275 to 498 ms. The durations of number words were 386, 498, 483, and 275 ms (mean 410 ± 103 ms) and those of face words were 405, 419, 338, and 370 ms (mean 383 ± 36 ms), respectively. To evaluate the differences between the low-level features of number and face words, a 3 × 3-window spectrogram and the mean harmonic-to-noise ratio of words were compared. No significant acoustical differences ($p > 0.05$) between number words and face words were observed (De Lucia et al., 2010).

Each word was presented using a loudspeaker 1 m away from the subject at 60 ± 10 dBA. The eight words were presented in a random order, with 80–110 repetitions of each word. The inter-stimulus interval was 2 s as shown in Fig. 2A. To sustain the subject's attention during the task, subjects were asked to look at a fixation cross on a monitor and repeat the heard word 3 times. The ECoG signals were measured from 20 to 126 subdural electrodes, with electrode diameter and inter-electrode distance of 4.0 and 10.0 mm, respectively, using the EEG-1200system and JE-125AK amplifier (Nihon-Kohden, Japan) at a sampling rate of 1000 Hz. The reference for the recordings was set to the scalp electrode at location Pz. The ECoG signals were epoched from 0.2 s before to 1.0 s after the sound-stimulus onset (Fig. 2A), and then were manually inspected to remove artifact epochs based on the threshold of ± 300 μV. No epochs were rejected in this study.

2.2. Decoding with a semantic-hierarchical model

The event-related spectral perturbation (ERSP) was used to measure stimulus-related cortical activity as opposed to spontaneous activity (Makeig, 1993). Our ERSP computation represented the significance of permutation statistics compared to a baseline-normalized power

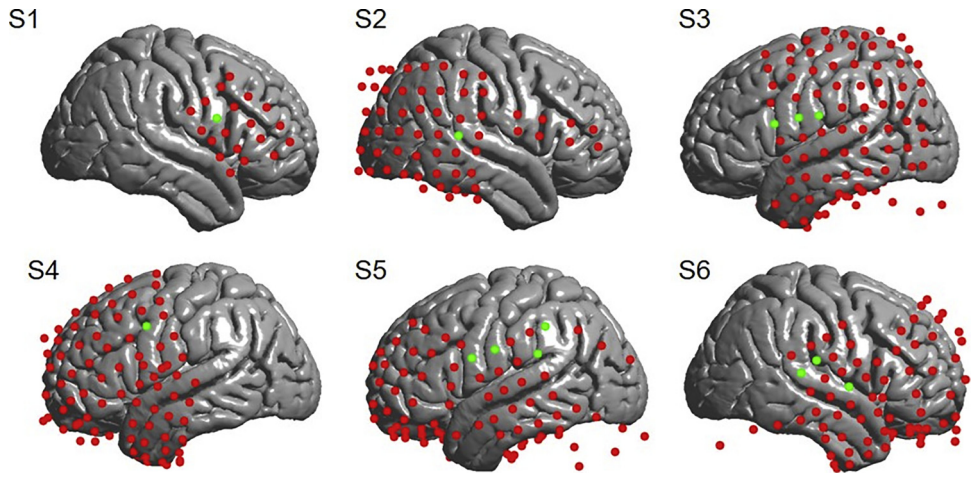

Fig. 1. The red circles indicate electrode locations for the six subjects. At least one language area was covered by the thirteen electrodes shown in green, as determined intraoperatively.

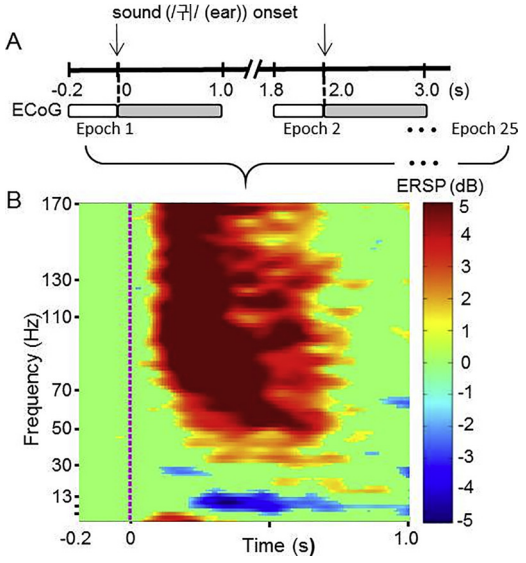

Fig. 2. A: The experimental paradigm consists of 640–880 trials. Sound onsets are separated by 2 s. The subject sees a fixation cross on a monitor until finishing the paradigm. The data set is combined 20–25 trials for each word. **B:** An example of feature extraction from S3 data when the patient hears the word (/귀/). In the frequency domain, the features were separated into seven frequency bands. In the time domain, the features were dependent on the size of time windows and overlap windows. For example, if the time and overlap windows are 200 ms and 100 ms, respectively, the time domain has 9 boundaries within 0–1000 ms. The color axis represents the significance of permutation statistics compared to a baseline-normalized power spectrum in a -200 to 0 ms interval.

spectrum in a -200 to 0 ms interval, in which there was no word stimuli. An ERSP was constructed from twenty-five epochs. Thus, from 80 to 110 epochs of each heard word, we obtained four unique ERSPs.

Fig. 2B shows an example ERSP in response to the stimulus /귀gui/ (ear). The decoding features, used to classify ECoG signals in response to word stimuli, were calculated by averaging ERSPs in each frequency band of delta (2–4 Hz), theta (4–8 Hz), alpha (8–13 Hz), beta (13–30 Hz), low-gamma (30–50 Hz), high-gamma I (70–110 Hz), and high-gamma II (130–170 Hz), within various time windows (20, 30, 40, 50, 100, and 200 ms, all with 50% overlap). The number of features for each 1000 ms ERSP ranged from 693 (for 20 ms time window and 7 frequency bandwidths) to 63 (for 200 ms time window and 7 frequency bandwidths). To select effective features, first, candidate lists of features were calculated separately for the semantic-hierarchical model and the non-hierarchical model using the minimal-redundancy-maximal-relevance criterion (mRMR) (Peng et al., 2005). Then, the features for each of the two models were optimally selected by computing decoding performance as shown in Table 2.

As shown in Fig. 3, ECoG signals were decoded using a multilayer-linear support vector machine (SVM) (Kecman, 2001) with linear kernel function and sequential minimal optimization method. Three of four data sets were used to train the SVM and the remaining set was used to test a decoding strategy. The key feature of this study is an additional classification step using semantic-hierarchical structure. The semantic-hierarchical model has a sub-model which classifies the ECoG signal to one of two semantic groups (face group or number group) prior to decoding the ECoG signal to one of 4 words (two-layer classification); the non-hierarchical model directly decodes ECoG signals to one of 8 words without any sub-model (i.e., single-layer classification). Four-fold cross-validation was employed to calculate the overall decoding accuracy for both the semantic-hierarchical and non-hierarchical models. The performance of each decoding method was evaluated using correct decoding rate (CDR). The CDR was calculated by counting the number of correctly decoded test sets over all 4-fold cross-validations for all 8 words (i.e., 32 total tests). Fig. 4 shows an example of CDR calculation for subject S3 at the recording site 3. The CDRs of the semantic-hierarchical model (CDR_{sh}) and non-hierarchical model (CDR_{nh}) were directly computed for each word and were averaged across 32 combinations (i.e., 8 words x 4 folds of each word).

To investigate the possibility overfitting could lead to false positive results as well as true negative results in this high-dimensional decoding, 'wrong-labeled' ECoG signals and 'shuffled-clustered' ECoG signals were also tested using a hierarchical model. The 'wrong-labeled' case disrupted the word-ECoG relationship by randomly labeling the words associated with the ECoG signals. The 'shuffled-clusteed' case disrupted the semantic labeling by misclustering one or two 'face' words as 'number' words, and vice versa, while still maintaining an accurate word-to-ECoG signal relationship. Although the 8 words could be shuffled with 52 different cases (i.e., 16 for the single-word shuffle and 36 for 2-word shuffles), in this study, we provided 10 shuffled cases (5 for the single-word shuffle and 5 for 2 word shuffles) as seen in Fig. 7.

3. Results

3.1. Spatial distribution of decoding accuracy

The CDRs of 13 ECoG recording sites, pooled across all subjects from those sites which were labeled as being associated with language intraoperatively, are shown on the MNI average brain model in Fig. 5. In the case of the semantic-hierarchical model, the highest CDR of 81% was observed at electrode 3, which had been clinically defined as covering Wernicke's area during electrode placement. This is significantly greater than the highest CDR using the non-hierarchical model, which was 50%.

Fig. 6A compares the CDRs across all subjects for the semantic-hierarchical model, non-hierarchical model, and semantic-hierarchical model using randomly wrong-labeled ECoG signals. The decoding performance of the semantic-hierarchical model was significantly higher than that of the non-hierarchical model ($p < 0.001$, signed rank sum test), and both of these models performed significantly better than the wrong-labeled dataset average CDR, which was close to the random chance level of 12.5%. This shows that the semantic-hierarchical model is not flawed by an overfitting of machine learning to this specific data set. Table 3 details the CDRs for the semantic-hierarchical and non-hierarchical models. The CDRs of the semantic-hierarchical model are higher than the non-hierarchical model at all the recording sites.

The upper panel of Fig. 6B compares the CDRs between the semantic-hierarchical model and non-hierarchical model for each subject. Paired *t*-test was performed for subjects S3, S5, and S6 for whom the numbers of recording sites were greater than 2 (i.e., 3 sites for S3 and S6, 4 sites for S5). The *p*-value was 0.034, 0.011, and 0.002 for subjects S3, S5, and S6, respectively. This statistical test was not performed for subjects S1, S2, and S4, for whom only one recording site could be obtained. The effect sizes for all individual subjects, as computed as $(CDR_{sh}\text{-}CDR_{nh})/CDR_{nh}$, are represented at the bottom panel of Fig. 6B. The mean CDR benefit of the semantic-hierarchical model was 61%, ranged from 39% (subject S6) to 73% (subject S2).

To investigate the specificity of this hierarchical clustering effect, the CDRs were also computed using a shuffled-hierarchical clustering approach, shown in Fig. 7. In the semantic-hierarchical model, the 8 words were clustered into two groups; semantic clustering accuracy was based on comparing the result of the classifier to the correct group assignment. In the cases of shuffled hierarchical clusters, altered group assignments were created by shuffling one word (lower row) or two words (upper row) with those in the other group. The semantic clustering accuracy for the shuffled hierarchical approach was then computed using this altered group assignment. The CDRs at 13 ECoG recording sites are plotted in Fig. 7. These results reveal that the CDR

Table 2
Time window and number of features used at the hierarchical model and non-hierarchical model.

Recording site num.	Semantic hierarchical model						Non-hierarchical model	
	Semantic decoding		Face decoding		Number decoding			
	Time window (ms)	Number of features	Time window (ms)	Number of features	Time window (ms)	Number of features	Time window (ms)	Number of features
1	200	53	200	9	50	40	100	113
2	20	69	40	68	30	45	200	3
3	50	95	100	73	200	53	200	37
4	50	68	100	39	100	33	30	68
5	200	28	30	136	100	59	100	86
6	50	68	200	34	40	17	200	53
7	30	45	40	154	100	73	50	27
8	40	34	200	18	50	40	200	15
9	50	13	200	40	200	15	40	17
10	100	6	40	102	40	85	200	28
11	200	15	40	34	40	68	50	28
12	200	28	200	44	50	13	100	6
13	200	15	100	106	100	13	200	15

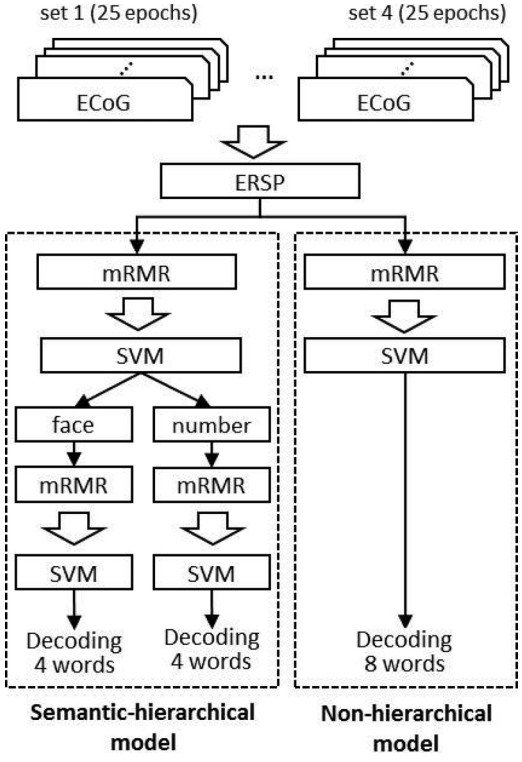

Fig. 3. Schematic of ECoG data processing of the semantic-hierarchical model and non-hierarchical model.

using the semantic-hierarchical cluster is significantly higher than those using shuffled-hierarchical clusters ($p < 0.001$, paired t-test).

4. Discussion

In the present study, a semantic information-based hierarchical model improved the classification accuracy of spoken word-evoked cortical responses measured using ECoG. The CDR of the semantic-hierarchical model was significantly higher than that of either the non-hierarchical model or the shuffled (not semantically clustered) hierarchical model.

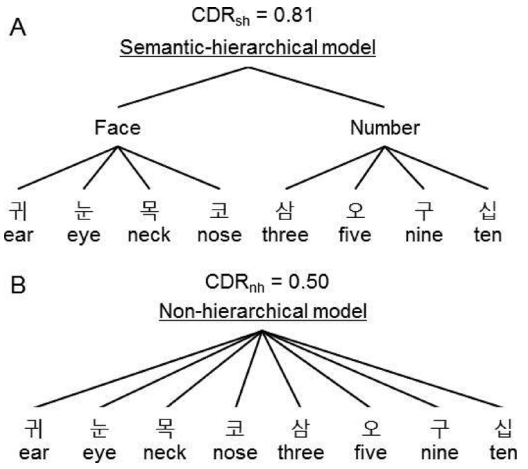

Fig. 4. Schematic for computing a correct decoding rate (CDR) for **A:** the semantic-hierarchical model and **B:** non-hierarchical model.

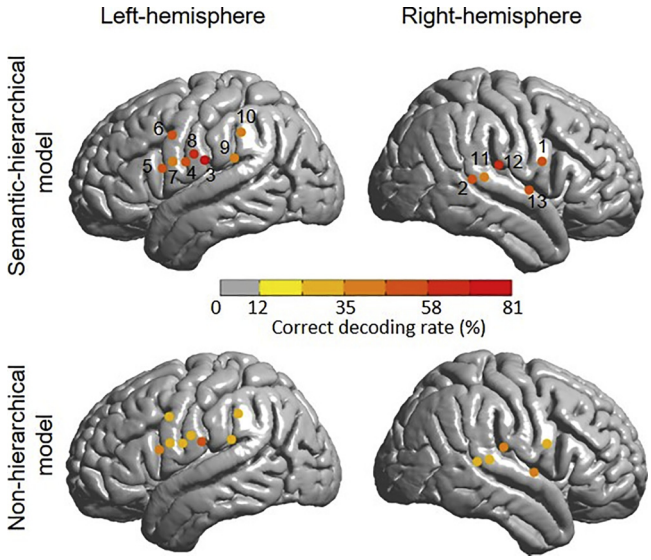

Fig. 5. Correct decoding rates at each locus of our ECoG coverage. These ECoG sites are pooled from all subjects. The number in the brain model denotes the recording site as shown in Table 3, and the color axis denotes CDR from 0% to 81%.

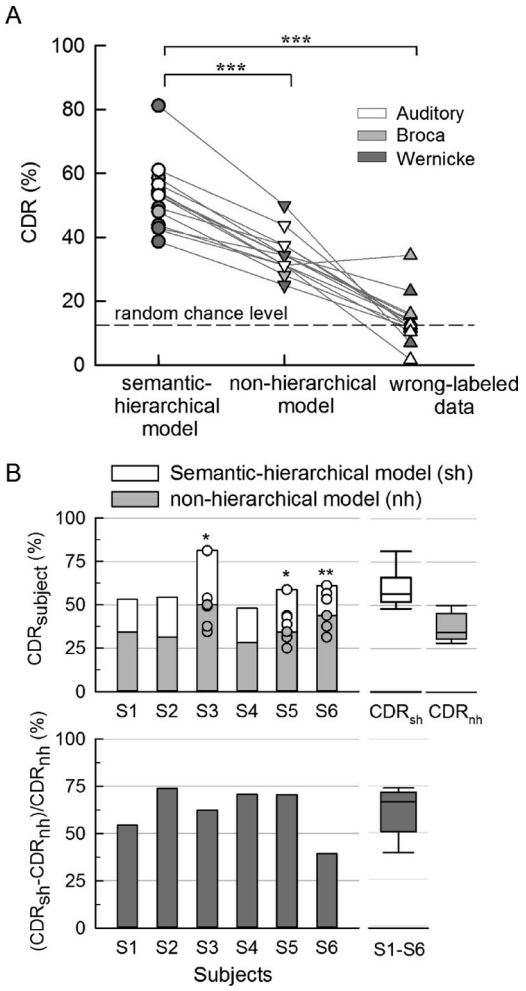

Fig. 6. A: Comparison of correct decoding rates (CDRs) of a semantic-hierarchical model with those of a non-hierarchical model and a semantic-hierarchical model using a wrong-labeled dataset. **B:** The highest CDRs ($CDR_{subject}$) of a semantic-hierarchical model (white bar) and non-hierarchical model (grey bar) for each subject (upper panel) and the percent difference between CDR_{sh} and CDR_{nh} (lower panel). The circle plots in the bars indicate the individual CDRs. ($^{*}p < 0.05$, $^{**}p < 0.01$, and $^{***}p < 0.001$).

4.1. Semantic processing in the human brain

Several studies provide empirical bases for interpreting why the semantic-hierarchical approach improves classification accuracy of cortical evoked responses. Huth et al. (2016) and others have demonstrated that language is semantically represented in the human brain. Such semantical processing relies on both anatomical and functional connectivity (Binder and Desai, 2011; Hickok and Poeppel, 2007; Moseley et al., 2012; Moseley and Pulvermüller, 2014; Yang et al., 2017). It follows that first classifying brain signals into semantic clusters might have a benefit when decoding heard words. Because the left inferior frontal cortex is directly involved in word-semantic processing and the superior temporal gyrus is the classic language semantic processor (Binder and Desai, 2011; Poldrack et al., 1999; Pulvermüller, 2013), for the present study we selected 13 electrodes, covering those areas, from which to decode ECoG signals, as described in Table 3. Wernicke's area is known to extract auditory features for recognition (Hickok and Poeppel, 2007; Rauschecker and Scott, 2009; Recanzone and Cohen, 2010; Robson et al., 2014; Steinschneider, 2011). This might explain why ECoG signals from electrode 3 (clinically mapped as Wernicke's area) could be decoded well.

4.2. Speech perception and speech imagery

Generally, BCI systems aim to decode or classify brain signals in response to speech imagery (covert speech). Speech imagery tends to produce similar neural representations to those of actual speech perception, as both imagery and perception rely on the same brain functions (Brigham and Kumar, 2010). For proof-of-concept of the implementation of the semantic-hierarchical model, and due to the technical difficulty of testing speech imagery, this study used ECoG signals from a speech perception task. This approach yielded accurate and effective decoding.

4.3. Study limitations and future directions

Although ECoG provides high temporal and spatial resolution, the spatial coverage of electrodes is limited. In this study, the electrodes, chosen from 6 subjects, non-symmetrically covered the right hemisphere (5 electrodes) and left hemisphere (8 electrodes). There were no aligned or spatially-paired electrodes between left and right hemispheres. Therefore, one limitation of this study is that we cannot directly compare decoding performance across hemispheres, or according to lateralization of brain function.

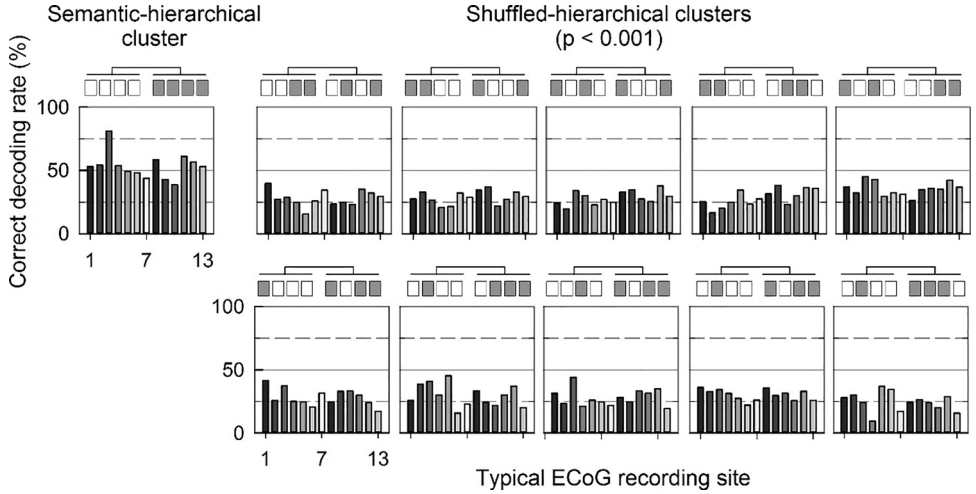

Fig. 7. Effect of hierarchical clustering based on semantic information. The clustering conditions are represented in each panel. The white and grey color of rectangles represents the face (ear, eye, neck, and nose) and number (three, five, nine, and ten) semantic groups, respectively.

Table 3
Anatomical and clinical language mapping with correct decoding rate of both semantic-hierarchical and non-hierarchical models at each site.

Recording site num.	Anatomical name	Clinical language mapping	Subject	Correct decoding rate (%)	
				Semantic-hierarchical model	Non-hierarchical model
1	Inferior Frontal Gyrus	Broca	S1	53.1	34.4
2	Superior Temporal Gyrus	Auditory	S2	54.4	31.3
3	Postcentral Gyrus	Wernicke	S3	81.2	50.0
4	Precentral Gyrus	Wernicke	S3	53.7	34.4
5	Inferior Frontal Gyrus	Wernicke	S3	49.2	37.5
6	Precentral Gyrus	Broca	S4	48.0	28.1
7	Precentral Gyrus	Broca	S5	43.8	31.3
8	Precentral Gyrus	Broca	S5	58.6	34.4
9	Superior Temporal Gyrus	Wernicke	S5	42.9	34.4
10	Inferior Parietal Lobule	Wernicke	S5	38.7	25.0
11	Superior Temporal Gyrus	Auditory	S6	61.0	43.8
12	Postcentral Gyrus	Auditory	S6	56.5	37.5
13	Superior Temporal Gyrus	Auditory	S6	53.1	31.3

Several areas of improvement for BCI systems can be suggested based on our findings. Because of the success of the semantic-hierarchical model on ECoG signals generated from a speech perception task, this model should, in the future, be tested with brain signals evoked by speech imagery. We were unable to use ECoG signals from multiple electrodes for any individual brain, as the subdural electrode arrays were implanted for clinical purposes in epilepsy patients, and therefore the locations of the electrodes were not identical. Thus, we processed ECoG channels individually in this study. The use of more channels of data can increase the accuracy of speech perception decoding (Kellis et al., 2010; Pasley et al., 2012). Thus, decoding multi-channel data, for example by including both Wernicke's and Broca's areas from a single subject, combined with the semantic-hierarchical model, may increase BCI performance. In addition, increasing the numbers of informative features from machine-learning characteristics (Setiono and Liu, 1997) and extending semantic classes dependent on different brain areas (Huth et al., 2016) may also help develop a more effective and efficient BCI system.

Conflicts of interests

The authors declare that there is no conflict of interest in relation to this work.

Acknowledgments

This work was supported by the 2016 Research Fund of the University of Ulsan, Korea.

References

Binder, J.R., Desai, R.H., 2011. The neurobiology of semantic memory. Trends Cogn. Sci. (Regul. Ed.) 15, 527–536. https://doi.org/10.1016/j.tics.2011.10.001.

Brigham, K., Kumar, B.V.K.V., 2010. Subject identification from Electroencephalogram (EEG) signals during imagined speech. IEEE 4th Int. Conf. Biometrics Theory, Appl. Syst. BTAS 2010. https://doi.org/10.1109/BTAS.2010.5634515.

Chang, E.F., Edwards, E., Nagarajan, S.S., Fogelson, N., Dalal, S.S., Canolty, R.T., Kirsch, H.E., Barbaro, N.M., Knight, R.T., 2011. Cortical spatio-temporal dynamics underlying phonological target detection in humans. J. Cogn. Neurosci. 23, 1437–1446. https://doi.org/10.1162/jocn.2010.21466.

Crosse, M.J., Di Liberto, G.M., Bednar, A., Lalor, E.C., 2016. The multivariate temporal response function (mTRF) toolbox: a MATLAB toolbox for relating neural signals to continuous stimuli. Front. Hum. Neurosci. 10, 1–14. https://doi.org/10.3389/fnhum.2016.00604.

de Heer, W.A., Huth, A.G., Griffiths, T.L., Gallant, J.L., Theunissen, F.E., 2017. The hierarchical cortical organization of human speech processing. J. Neurosci. 37, 6539–6557. https://doi.org/10.1523/JNEUROSCI.3267-16.2017.

De Lucia, M., Clarke, S., Murray, M.M., 2010. A temporal hierarchy for conspecific vocalization discrimination in humans. J. Neurosci. 30, 11210–11221. https://doi.org/10.1523/JNEUROSCI.2239-10.2010.

Di Liberto, G.M., Lalor, E.C., Millman, R.E., 2018. Causal cortical dynamics of a predictive enhancement of speech intelligibility. Neuroimage 166, 247–258. https://doi.org/10.1016/j.neuroimage.2017.10.066.

Hickok, G., Poeppel, D., 2007. The cortical organization of speech processing. Nat. Rev. Neurosci. 8, 393–402 https://doi.org/nrn2113 [pii]\r10.1038/nrn2113.

Huth, A.G., de Heer, W.A., Griffiths, T.L., Theunissen, F.E., Gallant, J.L., 2016. Natural speech reveals the semantic maps that tile human cerebral cortex. Nature 532, 453–458. https://doi.org/10.1038/nature17637.

Kecman, V., 2001. Learning and Soft Computing. The MIT Presshttps://doi.org/10.1007/s13398-014-0173-7.2.

Kellis, S., Miller, K., Thomson, K., Brown, R., House, P., Greger, B., 2010. Decoding spoken words using local field potentials recorded from the cortical surface. J. Neural Eng. 7https://doi.org/10.1088/1741-2560/7/5/056007. 056007.

Khalighinejad, B., Cruzatto da Silva, G., Mesgarani, N., 2017. Dynamic encoding of acoustic features in neural responses to continuous speech. J. Neurosci. 37, 2176–2185. https://doi.org/10.1523/JNEUROSCI.2383-16.2017.

Khoshkhoo, S., Leonard, M.K., Mesgarani, N., Chang, E.F., 2018. Neural correlates of sine-wave speech intelligibility in human frontal and temporal cortex. Brain Lang. https://doi.org/10.1016/j.bandl.2018.01.007. In press.

Makeig, S., 1993. Auditory event-related dynamics of the EEG spectrum and effects of exposure to tones. Electroencephalogr. Clin. Neurophysiol. 86, 283–293. https://doi.org/10.1016/0013-4694(93)90110-H.

Moseley, R., Carota, F., Hauk, O., Mohr, B., Pulvermüller, F., 2012. A role for the motor system in binding abstract emotional meaning. Cereb. Cortex 22, 1634–1647. https://doi.org/10.1093/cercor/bhr238.

Moseley, R.L., Pulvermüller, F., 2014. Nouns, verbs, objects, actions, and abstractions: local fMRI activity indexes semantics, not lexical categories. Brain Lang. 132, 28–42. https://doi.org/10.1016/j.bandl.2014.03.001.

Overath, T., McDermott, J.H., Zarate, J.M., Poeppel, D., 2015. The cortical analysis of speech-specific temporal structure revealed by responses to sound quilts. Nat. Neurosci. 18, 903–911. https://doi.org/10.1038/nn.4021.

Pasley, B.N., David, S.V., Mesgarani, N., Flinker, A., Shamma, S.A., Crone, N.E., Knight, R.T., Chang, E.F., 2012. Reconstructing speech from human auditory cortex. PLoS Biol. 10. https://doi.org/10.1371/journal.pbio.1001251.

Peng, H., Long, F., Ding, C., 2005. Feature selection based on mutual information: criteria of Max-Dependency, Max-Relevance, and Min-Redundancy. IEEE Trans. Pattern Anal. Mach. Intell. 27, 1226–1238. https://doi.org/10.1109/TPAMI.2005.159.

Poldrack, R.A., Wagner, A.D., Prull, M.W., Desmond, J.E., Glover, G.H., Gabrieli, J.D.E.E., 1999. Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex 1. Neuroimage 35, 15–35. https://doi.org/10.1006/nimg.1999.0441.

Pulvermüller, F., 2013. How neurons make meaning: brain mechanisms for embodied and abstract-symbolic semantics. Trends Cogn. Sci. (Regul. Ed.) 17, 458–470. https://doi.org/10.1016/j.tics.2013.06.004.

Rauschecker, J.P., Scott, S.K., 2009. Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. Nat. Neurosci. 12, 718–724. https://doi.org/10.1038/nn.2331.

Recanzone, G.H., Cohen, Y.E., 2010. Serial and parallel processing in the primate auditory cortex revisited. Behav. Brain Res. 206, 1–7. https://doi.org/10.1016/j.bbr.2009.08.015.

Robson, H., Zahn, R., Keidel, J.L., Binney, R.J., Sage, K., Lambon Ralph, M.A., 2014. The anterior temporal lobes support residual comprehension in Wernicke's aphasia. Brain 137, 931–943. https://doi.org/10.1093/brain/awt373.

Setiono, R., Liu, H., 1997. Neural-network feature selector. IEEE Trans. Neural Networks 8, 654–662. https://doi.org/10.1109/72.572104.

Steinschneider, M., 2011. Unlocking the role of the superior temporal gyrus for speech sound categorization. J. Neurophysiol. 105, 2631–2633. https://doi.org/10.1152/jn.00238.2011.

Vanthornhout, J., Decruy, L., Wouters, J., Simon, J., Francart, T., 2018. Speech intelligibility predicted from neural entrainment of the speech envelope. bioRxiv 19, 181–191. https://doi.org/10.1101/246660.

Worsley, K., Taylor, J., Carbonell, F., Chung, M., Duerden, E., Bernhardt, B., Lyttelton, O., Boucher, M., Evans, A., 2009. SurfStat: a Matlab toolbox for the statistical analysis of univariate and multivariate surface and volumetric data using linear mixed effects models and random field theory. Neuroimage 47, S102. https://doi.org/10.1016/S1053-8119(09)70882-1.

Yang, Y., Dickey, M.W., Fiez, J., Murphy, B., Mitchell, T., Collinger, J., Tyler-Kabara, E., Boninger, M., Wang, W., 2017. Sensorimotor experience and verb-category mapping in human sensory, motor and parietal neurons. Cortex 92, 304–319. https://doi.org/10.1016/j.cortex.2017.04.021.