

Published in final edited form as:

IEEE Trans Neural Syst Rehabil Eng. 2018 May ; 26(5): 1084–1092. doi:10.1109/TNSRE.2018.2821058.

## Spatial-Temporal Dynamics of the Sensorimotor Cortex: Sustained and Transient Activity

E. Salari, Z.V. Freudenburg [Member, IEEE], M.J. Vansteensel, and N.F. Ramsey

Department of Neurology and Neurosurgery, Brain Center Rudolf Magnus, University Medical Center Utrecht, 3584CX, Utrecht, The Netherlands

### Abstract

How the sensorimotor cortex is organized with respect to controlling different features of movement is unclear. One unresolved question concerns the relation between the duration of an action and the duration of the associated neuronal activity change in the sensorimotor cortex. Using subdural electrocorticography electrodes, we investigated in five subjects, whether high frequency band (HFB; 75-135 Hz) power changes have a transient or sustained relation to speech duration, during pronunciation of the Dutch /i/ vowel with different durations. We showed that the neuronal activity patterns recorded from the sensorimotor cortex can be directly related to action duration in some locations, whereas in other locations, during the same action, neuronal activity is transient, with a peak in HFB activity at movement onset and/or offset. This data sheds light on the neural underpinnings of motor actions and we discuss the possible mechanisms underlying these different response types.

### Index Terms

Sensorimotor cortex; Movement; Duration; ECoG; Speech

## I Introduction

VOLUNTARY body movements are controlled by the sensorimotor areas of the brain [1, pp. 756–770], [2]. Studies on the relation between sensorimotor cortex activity and body movements have revealed, among others, that the sensorimotor cortex is somatotopically organized. This somatotopic organization seems quite detailed, since individual fingers [3]–[5] and even separate muscles [6], [7] can be distinguished from one another within this brain area. Besides a topographical organization, there is evidence from both monkey and human studies that the sensorimotor cortex has a role in controlling different *features* of a movement. It has been shown for instance that the activity of single sensorimotor neurons can be related to the position of a body part [8], [9], but also to movement direction [10], [11] or velocity [8], [9], [11], change in force [12] or to motor planning [13], [14]. How the sensorimotor cortex is organized with respect to controlling these different movement features remains to be determined. Importantly, the relationship between neural activity and movement features may not be straightforward, since there is evidence that repeating the

same movement shortly after one another is not accompanied by the same magnitude of neuronal signal change for every repetition, despite equal behavioural output [15]. This indicates that neuronal activity is not always linearly related to movement and can depend on past actions. Clearly, the role of the sensorimotor cortex in controlling body movements is complex and the underlying mechanisms are not yet completely understood. Investigating the detailed relationship between the spatial-temporal neural patterns within this area and overt body movements may contribute to our understanding of the functioning of this area, which is also of importance for neural engineering purposes such as the development of brain-computer interface (BCI) systems.

One unresolved question involves the relation between the duration of neuronal activity and the duration of a motor action. In other words, is neuronal activity continuously or transiently related to motor output? Evidence for the existence of both transient and sustained responses mainly comes from single cell primate studies [13], [14], [16]–[18]. The topic received relatively little attention in studies with human subjects, but evidence for both transient and sustained responses have also been found in humans [19]–[21] and some studies have shown an effect of duration on the neural response profiles [19], [21]–[23]. It remains to be determined, however, how exactly the duration of movement is encoded in the brain with respect to these two types of responses, and whether or not extended action duration is associated with a corresponding temporal extension of neuronal activity. Furthermore, a detailed spatial mapping of both sustained and transient responses in humans is still missing for complex movements such as those involved in speech.

In the current study, we investigated the relation between sensorimotor cortex activity duration and action duration. Since people can quite easily vary the duration of the pronunciation of vowels, and thereby the duration of a motor action, we focused on the motor cortex areas involved in articulator movements and speech pronunciation. Articulator movements are known to be controlled by the ventral parts of the sensorimotor cortex [24]–[27] with the larynx and tongue being represented more ventrally and the jaw and lips more dorsally [24]. Indeed, it has been demonstrated that articulator movements [28], as well as speech units [29], [30] can be distinguished ('classified') reliably from this area.

We recorded neural signals with electrocorticography (ECoG), a technique that benefits from a unique combination of high temporal resolution, comparable to electroencephalography (EEG), and high cortical sampling specificity [5]. Using this technique, it has been demonstrated that movements of, for instance, the tongue, lips, hand or foot, are accompanied by an increase in high-frequency-band (HFB; >50 Hz) power in the sensorimotor cortex [20], [24], [31]. This HFB power increase is thought to be associated with underlying neuronal firing [32]–[34]. Therefore, we used HFB power changes associated with sustained pronunciation of single vowels to investigate whether neural activity has a transient or sustained relation to speech duration.

## II Method

### A Subjects

Five subjects (age 14-41y; median 21y, 3 females), who were implanted with subdural ECoG electrodes for the treatment of epilepsy in the University Medical Center Utrecht participated in this study. Three subjects (A, B & E) had coverage with standard clinical grids (exposed electrode diameter of 2.3 mm with a 10 mm inter-electrode distance) over the ventral sensorimotor cortex (vSMC; left hemisphere in 2 subjects, right in 1 subject). Two subjects (C & D) gave permission to place, besides the clinical grids, an extra high-density (HD) electrode grid (exposed electrode diameter 1 mm for subject B and 1.17 mm for subject C, inter-electrode distance 4 mm for both subjects) over the left mouth sensorimotor cortex for research purposes. For these subjects, only the HD electrodes were used for the current analysis.

This study was approved by the ethics committee of the University Medical Center Utrecht and is in accordance with the Declaration of Helsinki (2013). All subjects gave written informed consent.

### B Task

Subjects were asked to perform a vowel durations task in which they pronounced the Dutch /i/ vowel for 1, 2 or 3 seconds. This phoneme was chosen as it is easy to pronounce for variable durations and because it engages multiple articulators, including the tongue [35, p. 5], [36], which is well represented in sensorimotor cortex [24], [37]. The task was presented on a computer screen, which was placed at a distance of about 1 m. A trial started with a 500 ms cue indicating the pronunciation duration (1, 2 or 3 seconds) to prepare subjects. After 1000 ms, a visual cue (presented by the letters 'ie', corresponding, in Dutch, to the /i/ sound from the international phonetic alphabet) instructed the subjects to start the pronunciation and hold it for as long as the visual cue was visible, which was followed by an inter-trial interval (fixation cross presented) of 2000 ms. Trials were randomized and each type was repeated 15 times per run (3 subjects performed two runs).

### C Data Acquisition

Signals were recorded at a sampling rate of 512Hz for subjects A, B, C (run 2), and E, at 2048Hz for subject C (run 1; all Micromed, Treviso, Italy) and at 2000Hz for subject D (Blackrock Microsystems LLC, Salt Lake City, USA). Three subjects performed the task twice (A, C & D), and for these subjects the data of the two runs were concatenated (for subject C the data of the first run was down-sampled to 512Hz for concatenation). Vowel pronunciation was recorded using microphones installed in the patient's room. All data was processed and analyzed using Matlab software (The Mathworks, Inc., Natick, MA, USA), unless specified otherwise.

### D Data pre-processing

First, deviations in the power-density distributions, line noise values and raw voltage distributions (some of which are described in [38]) were used to identify electrodes with noisy or flat signals, which were subsequently removed from further analysis. Electrodes

that were classified as ictal by a neurologist were also removed from further analysis. For the remaining electrodes, line noise and harmonics thereof were removed using a 3rd order butterworth filter ('butter' function and 'filtfilt' function from Matlab) and a common average re-reference was performed. Finally, signals were visually inspected and any trials with excessive noise (which sometimes occurs as a result of cable movements) were removed from the analysis.

Subsequently, electrode positions were visualized on the 3D rendering of the pre-surgical MRI scan using an in-house developed procedure [39], [40]. Electrodes that were located over the sensorimotor cortex (pre- and postcentral gyrus as indicated by the freesurfer MRI segmentation [41]) were identified by visual inspection. All further analyses were performed using only these electrodes (25, 18, 38, 114 and 8 electrodes for subjects A-E respectively), excluding the identified noisy and ictal electrodes.

For every retained sensorimotor cortex electrode, the high frequency band (75-135 Hz) power was calculated per sample point by applying a Gabor wavelet function [42] for all frequencies in the HFB range in bins of 1 Hz, with a full width half maximum (fwhm) of 4 wavelets per frequency, followed by a log transformation ( $10^{\ast}\log_{10}$ ) and averaging (over frequencies between 75 and 135Hz) of the resulting values. HFB power values (for all time points) were subsequently normalized by subtracting the mean signal value of the whole time series and dividing the result by the standard deviation of the whole time series. Finally, this signal was smoothed with a moving average window of 500 ms (window centered around each sample). We chose the 500 ms window because we found this to be the optimal window for (speech) movement classification in a previous study. In that study (currently under review with a journal) the smoothing window was determined using an elaborate optimization algorithm for several parameters, using classification of spatiotemporal HFB patterns as optimization outcome measure. The current frequency band (75-135Hz) was chosen such that the lower bound matches that of previous experiments [20]. The upper bound was determined by hardware filters. The audio signal was aligned with the brain signal and the voice onsets and offsets were automatically determined by a vowel detection algorithm [43] and subsequently checked and corrected using Praat annotation software [44]. Deviations in voice onset and offset from the cued timing were corrected in the brain signal by interpolation or down sampling, so that the brain signals of different trials could be adequately aligned. This step was necessary to allow for comparison of different (complete) trials, including both voice onset and offset. Only by making sure that each trial has the same duration, can complete trials be averaged and visualized as a whole, including a reliable, undiluted estimation of the neural responses associated with both voice onset and offset and the period in between. Visual inspection of the original and corrected time-series did not reveal any major differences other than the expected slight timing differences, indicating that the interpolation and down-sampling did not have a major influence on the signal values.

## E Statistical procedures

A regression analysis with three predictors was performed on the whole (corrected) time series to describe the electrode responses to the task (Figure 1) and to determine which

electrodes responded significantly to the task. The predictors consisted of a transient response around voice onset, a sustained response between voice on- and offset, and a transient response around voice offset. The first and third predictor were created by convolving a Gaussian function to an impulse function with impulses at voice onset and offset respectively (corrected such that the Gaussian peak would be at voice onset or offset). The width of the Gaussian (fwhm) was estimated for every subject separately (mean = 0.65 seconds, range = 0.51 - 0.76 seconds) using data from a different task. Similar to the vowel durations task, subjects were required to repeatedly pronounce the same vowel (/i/), but in this case very briefly for three times at a 1 Hz repetition rate per trial. For this task, we determined the significant electrodes by doing a simple  $r^2$ -analysis, contrasting speech versus silence periods. We aligned all trials (13 per run, subjects A, C & D performed the task twice) to the first pronunciation and calculated the fwhm of the average peak (over trials and significant electrodes) in HFB power related to the first pronunciation. Note that, although potentially the ‘shape’ of a (transient) response can be different for different electrodes and therefore there doesn’t necessarily have to be one standard response, inspection of the data showed that this mean response described the data well and could be used for finding significant electrodes. The second predictor was created in a similar way as the transient predictors but with a block function between the impulses of the first and third predictor. Since the task had trials of 1, 2 or 3 second duration, we could evaluate whether sustained responses showed a relation with action duration.

The onset of movement-related activity can be different for different locations in the brain. This effect of shifting activity onset has been reported before in both monkey and human studies [15], [20], [24], [45] and has for speech been attributed to the involvement of different articulators [24]. We repeated the regression analysis procedure while shifting all predictors in steps of 0.1 second, from 0.5 seconds before voice onset/offset to 0.5 seconds after, to capture activity related to the movement as accurate as possible. For every shift, a model was created using the following formula,

$$model = \sum_{i=1}^4 P_i \times \beta_i \quad (1)$$

where P represents a predictor and  $\beta$  the corresponding regression beta values. In this formula, the summation is over 4 instead of 3 predictors due to the addition of the intercept.

The model with the highest variance explained was used for further analysis. The variance explained was calculated by the following formula,

$$(SS_{mean} - SS_{model})/SS_{mean} \quad (2)$$

where  $SS_{mean}$  is the sum of squared differences between the signal and the mean of that signal, and  $SS_{model}$  the sum of squared differences between the signal and the model.

The chosen model was tested for significance of explaining the data by using the analysis of variance statistic, with  $\alpha=0.05$  (false discovery rate (FDR) corrected). Note that for such

a statistical test the degrees of freedom in the denominator is  $n-k$ , where  $k$  is the number of predictors (including the intercept), and  $n$  the number of observations, assuming that all sample points are independent measurements. However, this would lead to an overestimation of the significance since the frequency conversion creates dependence among consecutive sample points over some time span. In addition, the HFB signal is a proxy for the underlying neural events, which by itself have an inherent temporal width. Therefore,  $n$  (the number of observations) was not set to one observation per sample point but to one observation per second of signal time ( $n =$  the total signal duration in seconds). This one-second value was chosen to fully capture a neural impulse response peak (max fwhm was 0.76 seconds).

For electrodes for which the model explained the data significantly, the beta values of the three predictors, resulting from the regression analysis, were tested for significance by converting them to  $t$ -values using the following formula:

$$T_p = \frac{\beta_p}{se(\beta_p)} \quad (3)$$

where  $\beta_p$  is the beta value for a predictor, and  $se(\beta_p)$  was calculated as follows,

$$se(\beta_p) = \sqrt{C_{jj}} \quad (4)$$

where  $C_{jj}$  is the diagonal value of the variance covariance matrix for the corresponding predictor.

The calculated  $t$ -value was subsequently converted to a one-sided  $p$ -value. The significance level was set to  $\alpha = 0.05$ , FDR corrected. Based on this, electrodes were assigned to one of six classes or deemed non-responsive. Electrodes that showed a significant beta value for the first predictor but not the other two predictors were classified as transient at voice onset (class 1). Electrodes that showed a significant beta value for the first and last predictor but not the second predictor were classified as transient at voice onset and voice offset (class 2) and if only the third predictor was significant, it was classified as transient at voice offset (class 3). Electrodes that had a significant positive beta value for the second predictor were classified as either only sustained (class 4), sustained with a peak at voice onset (class 5) or sustained with a peak at voice onset and voice offset (class 6), depending on whether the first and or third predictor were also significant. A sustained response with a peak only at voice offset was not found in our data and is therefore not mentioned further. If no predictors were significant, an electrode was classified as non-responsive.

For visualization purposes, the signal was epoched (see Figure 2) and trials of the same condition (1, 2 or 3 seconds) were averaged. This led to an average HFB-trace per electrode and per condition. These were subsequently averaged over electrodes and over subjects for each condition and class separately (see Figure 3).

Finally, we performed a correlation analysis between the model shift timing that explained the data best, serving as an indication of the HFB response onset timing relative to a movement (as indicated by voice onset or voice offset) and the anatomical dorsal-ventral localization. This was done to see if ventral areas showed relatively later responses than more dorsal areas as suggested by Bouchard and colleagues [24].

### III Results

Subjects performed the task well and no trials were removed based on behavioral performance. The average reaction time was 0.46 seconds (SD=0.23) and duration of pronunciation deviated from the intended speech duration, by -0.25 (SD=0.33) seconds on average, meaning that subjects usually pronounced the vowel a little shorter than instructed.

Each electrode that had a significant response (i.e. for which the model explained the data significantly) was classified as belonging to one of six classes as mentioned above (see method section E and Figure 3). Of all electrodes, on average (over subjects and weighted by number of sensorimotor cortex electrodes per subject) 29.1% (59/203, SD=7.4%) were classified as having a transient response without an additional sustained response, and 17.2% (35/203, SD=8.5%) as showing a sustained response with or without a transient peak. These sustained responses were related to the duration of the action, showing an increasing brain response duration with increasing vowel duration (Figure 3). The remaining electrodes (53.7%, 112/203, SD=15.1%) did not show a significant response to this analysis. Of the electrodes that were classified as transient, 59.3% (35/59, SD=13.9%) was classified as transient at voice onset only, 30.5% (18/59, SD=13.0%) as transient at voice onset and offset, and 10.2% (6/59, SD=5.3%) as transient at voice offset only. Interestingly, the transient electrodes that showed a peak at voice offset seem to have reduced activity between voice onset and offset compared to baseline, whereas the electrodes that only showed a peak at voice onset did not (see Figure 3a,c,e). Of all electrodes that were classified as sustained, 54.3% (19/35, SD=18.7%) also showed a peak at voice onset, 34.3% (12/35, SD=19.9%) a peak at voice onset and offset, and another 11.4% (4/35, SD=14.5%) was sustained without a clear peak of activity at voice onset or offset.

When looking at the anatomical localization of electrodes with different response profiles (Figure 4), no clear anatomical organization could be observed between transient and sustained electrodes, except for the subjects with the HD electrodes (subject C & D). For subject C, a cluster of transient responses seemed to be more anteriorly located, and a cluster of sustained responses more posteriorly. Subject D showed two clusters of sustained responses, one located ventral-posteriorly and the other more dorsal-anteriorly. For these subjects, we also plotted the timing of model shift for each sensorimotor electrode, serving as a marker for neural activity onset (see Figure 5 for the model shift timing). Notably, most electrodes started their activity before voice onset and showed a peak in activity mostly on or just after voice onset, see also Figure 3 for this. Furthermore, the more ventral electrodes seemed activate later than the dorsal electrodes for subject C, with a negative relation between the ventral-dorsal localization and response onset ( $R=-0.34$ ,  $p=0.08$ ), see Figure 5. For subject D, the same result was visible ( $R=-0.39$ ,  $p=0.01$ ) although there also seemed to be an anterior-posterior division with later responses mostly located posteriorly.

## IV Discussion

### A General discussion

In this study, we showed that the neuronal activity patterns recorded from the sensorimotor cortex can be directly related to action duration in some locations, whereas in other locations, during the same action, neuronal activity is transient, with a peak in HFB activity at voice onset and/or offset. To our knowledge, this is the first time that both transient and sustained neuronal dynamics of sensorimotor activity have been simultaneously and systematically mapped with respect to speech-actions of different durations in humans.

When looking at the results obtained in the three subjects with standard clinical grids over the sensorimotor mouth area, populations with sustained and transient responses seemed to be spatially scattered, without a clear anatomical organization or separation. Interestingly, however, data from the patients with a high-density grid allowed us to take a more detailed look at regional differences between response profiles and revealed that in the ventral sensorimotor mouth area, electrodes showing sustained responses occurred more posteriorly. Also, for one subject that had HD coverage more dorsally, a cluster of sustained responses was seen more anteriorly. These anatomical differences have not been shown before and might suggest that populations of neurons with transient and sustained response profiles are anatomically separated. Moreover, the high-density grid data revealed that activity that was associated with the onset of pronunciation usually started before voice onset, with a peak of activity close to voice onset for the more dorsal electrodes and usually somewhat later for the ventral electrodes. This result corresponds to the findings from Bouchard and colleagues [24] who found that articulators that are represented more dorsally (i.e. lips and jaw), show earlier responses, close to voice onset or before that, and features of the tongue position (located ventrally) are represented later, i.e. after voice onset. They also showed that the responses for the larynx (representation both ventrally and dorsally) were locked to voice onset, which could explain some of the early ventral responses in our data. Taken together, our data suggest that there is an underlying distinction between sustained and transient responses and also between early and late responses, with respect to their anatomical location, but more high-density ECoG data will be needed to confirm these findings.

Another phenomenon we observed was the occurrence, in both precentral and postcentral areas, of a (second) transient peak close to voice offset in some electrodes. In addition, for several electrodes that were classified as having only a peak at voice onset, it seemed that there was a small second peak (although not statistically significant). It could be that this voice offset peak was sometimes present but masked due to noise. Interestingly, Ball and colleagues [19] described similar offset-related increases in HFB power for arm movements and Hermes and colleagues [15] found the same result for finger movements, mainly within the postcentral gyrus (which is related to somatosensory functions; [37]) and attributed this to the notion that some cells in this area can fire with two directions of movements [18]. Also cells in the precentral gyrus have been found to be active with two directions; during flexion and/or extension of joints [17]. It is likely that the voice offset related peaks we observed are associated with articulators moving back to their rest position. Furthermore, we found that the electrodes that showed this offset related peak displayed a reduction in

activation compared to baseline between voice onset and offset. Previous functional magnetic resonance imaging (fMRI) and transcranial magnetic stimulation (TMS) studies have linked motor related deactivation responses (for hand movements) in the ipsilateral cortex to inhibition of the opposite hand and claim this to be necessary to reduce interfering movements [46]. Possibly, inhibition of neurons responsible for movements which are offset related and oppose the articulator positions during the pronunciation cause a deactivation in the HFB power.

## B Correspondence to previous research

We showed that different patterns of activity can be recorded with electrodes placed on the surface of the brain and our results are in correspondence to earlier primate studies. For instance, we found that both transient-only and transient-followed-by-sustained responses show a peak of activity at movement onset, followed by a decrease in activity (which was larger in transient responses than in sustained responses). This is in correspondence to earlier findings of single cell recordings in primates that first show a burst of activity during a dynamic movement phase followed by a decrease to a steady level of elevated activity during a tonic phase. In contrast, other cells are highly active only during the tonic phase after a movement [16], which corresponds to the sustained only responses that we found.

Previous work [20], [21] suggested that both duration related and duration unrelated responses are involved in hand and speech movements in humans, which is in correspondence to our data. We extend those findings, in that we experimentally show that the sustained responses are related to the duration of an action and that transient responses are related to movement onset or offset. Furthermore, we mapped these two types of responses on the sensorimotor cortex and show their anatomical positioning. However, in contrast to the findings of Crone and colleagues [20], who indicated that transient responses are more associated with higher-gamma frequencies (75-100 Hz) and sustained responses with low-gamma frequencies (40-50 Hz), we observed sustained responses in the higher frequencies. Although they suggest that lower frequencies are more involved in, for instance, motor output or sustained attention and higher frequencies in motor planning or initiation of movements, our results suggest that high frequency power can be associated with motor output as well. Task differences could, however, explain this inconsistency (see below for discussion).

## C Function of sustained and transient responses

One possible explanation for the sustained responses we observed is that they reflect the movement of body parts that are continuously moving or that apply continuous force during speech (in contrast to transient responses for non-continuously moving body parts). For instance, pulmonic egressive airstreams (air pushed from the lungs) enable speech [36, pp. 57–58] and during continuous speech, lung volume decreases steadily, which is realized mostly by internal and external intercostal muscle activity [47]. This might explain why in the study by Crone and colleagues [20], sustained responses were not found for higher-gamma frequencies, since subjects did not constantly move. They were asked to make a movement as fast as possible (e.g. tongue protrusion) and keep it in that position for a couple of seconds. Alternatively, sustained responses may reflect continuous somatosensory

feedback associated with sustained sound production. The data from our subjects with HD-grids may support this possibility in that clusters of sustained responses were observed in more posterior positions (towards the somatosensory cortex) in the ventral parts of the sensorimotor cortex. Indeed, the postcentral gyrus is mainly involved in controlling somatosensory processing, whereas the precentral gyrus is more involved in motor execution [37]. However, there is also overlap in function between pre- and postcentral gyrus activity with respect to both motor and sensory responses [17], [18], [37]. Therefore, some deviations from this classical pre- and postcentral division for sustained (and transient) responses can be expected. In both subjects with HD-coverage, sustained responses in the more ventral parts of the sensorimotor cortex were predominantly observed posterior to the central sulcus, suggesting that for ventral areas there might be an anterior-posterior distinction for sustained and transient responses (although more data is needed to investigate this). However, in one of these subjects, in the more dorsal areas, sustained responses were also located anterior to the central sulcus. Although this contradicts the anterior-posterior division from the ventral parts, we only observed this for one subject. It is therefore unclear if this finding was subject specific or a general phenomenon. A third possibility for the role of sustained responses could be the encoding of the position of a body part. Bouchard and colleagues [48] have suggested that the HFB-power is correlated with the position of the lips, which could suggest that the sustained responses reflect articulator position. However, since their results are correlational, no causal conclusion can be derived from this. Support for the position hypothesis, however, is also provided by single cell recordings in both primates and humans, that show position specific activity for sensorimotor cortex cells [8], [9] especially during a static phase after a movement [16]–[18].

Transient responses, on the other hand, can be expected if neuronal activity is related to initiation of the movement, and could therefore be associated with articulators or body parts that move only during the initiation or ending of sound production (in contrast to sustained responses for continuously moving body parts). For instance, the tongue has been shown to start moving just before voice onset, stay in a steady position throughout single syllable pronunciation and then move back just after voice offset [48]. Another explanation would be that transient responses reflect movement planning. Indeed, non-human primate vocalization research [45], [49] as well as research in songbirds [50] has described responses often preceding or surrounding voice onset. These profiles were most often seen in premotor areas, an area generally associated with motor planning [1, pp. 770–771], [51]–[53]. In our study, we found transient responses in the precentral and postcentral parts of the sensorimotor cortex, with activity usually starting before, but peak activity just on or after, voice onset (and/or voice offset). In correspondence to our results, activity of some neurons in the precentral gyrus have been found to depend on movement instructions during a preparatory (non-moving) phase before monkeys get a cue to move (push versus pull an object with their arm). Errors in movements can be predicted from these neurons before the actual movement [14]. This indicates that indeed, besides in the premotor areas, motor planning may be represented within the precentral gyrus as well.

## D Implications for brain computer interfaces

Our findings may help in the development of brain-computer interfaces (BCIs) for subjects that suffer from complete or nearly complete paralysis. Although it has been shown that speech units can be classified from sensorimotor cortex activity [29], [30], [54], [55], accuracy scores usually do not meet the standards for home use BCI-controlled language communication applications. This might potentially improve when more is known about the relationship between small speech units, and variations in pronunciation of these, and the brain signal. Our current study contributes to our understanding of the sensorimotor cortex and the relation between neuronal activity and behavioral output, which could lead to better classification of syllables, phonemes and eventually words and sentences.

## E Limitations

There are several limitations to this study that need to be addressed. First, only a limited number of subjects participated in this study and only two subjects had high-density electrode coverage. As described before, only subject C & D (with HD-grids) showed separate clusters of transient and sustained responses. It could be that, due to the sparse sampling with clinical grids, we missed active cortical sites in the other three patients. Alternatively, one could argue that in two of the three subjects with the clinical grids, statistical power was low, given that they performed only one run (and therefore half of the trials). However, this does not explain why the one subject with clinical grids who performed two runs, also did not show a functional organization. Also, since the electrodes in the clinical grids are larger, they record from a larger population of neurons than the HD-grid electrodes, which may lead to spatial blurring, mixing of response profiles and a smaller signal to noise ratio (for a simulation of this see [56]). It may be speculated that, because of the larger number of electrodes per square cm and the fact that these electrodes are also smaller, high-density recordings are especially suitable to accurately study the detailed organization of the sensorimotor cortex.

Second, we corrected the data for incorrect response timing (by interpolation and down sampling) to be able to compare the responses of different trials and different subjects. We don't believe this procedure has induced the differences between sustained and transient responses we observed, since during individual trials, both sustained and transient responses were found (in different locations) while the correction was the same for all electrodes.

Third, since the subjects heard their own voice during the task, we cannot rule out that the auditory stimulation contributed to the cortical responses. However, since sustained movement related responses in the sensorimotor cortex during single cell recordings in primates have been found in the absence of auditory stimulation we argue that the current results cannot be solely attributed to auditory stimulation.

## V Conclusion

We demonstrate here that some focal regions of the sensorimotor mouth area show sustained responses associated with action duration, whereas other sites show transient responses coupled to movement onset and/or offset. Sustained responses may be associated with

continuous movement or force, somatosensory feedback or articulator position, whereas transient responses could reflect the (initiation of) short-duration movements or motor planning. We believe our findings warrant further research into the nature of cortical responses during elementary articulations, which may improve our understanding of the cortical representation of complex motor actions such as speech, and thereby improve decodability for brain-computer interfaces.

## Acknowledgment

The authors thank the staff of the clinical neurophysiology department, the neurosurgeons and the subjects for their contribution.

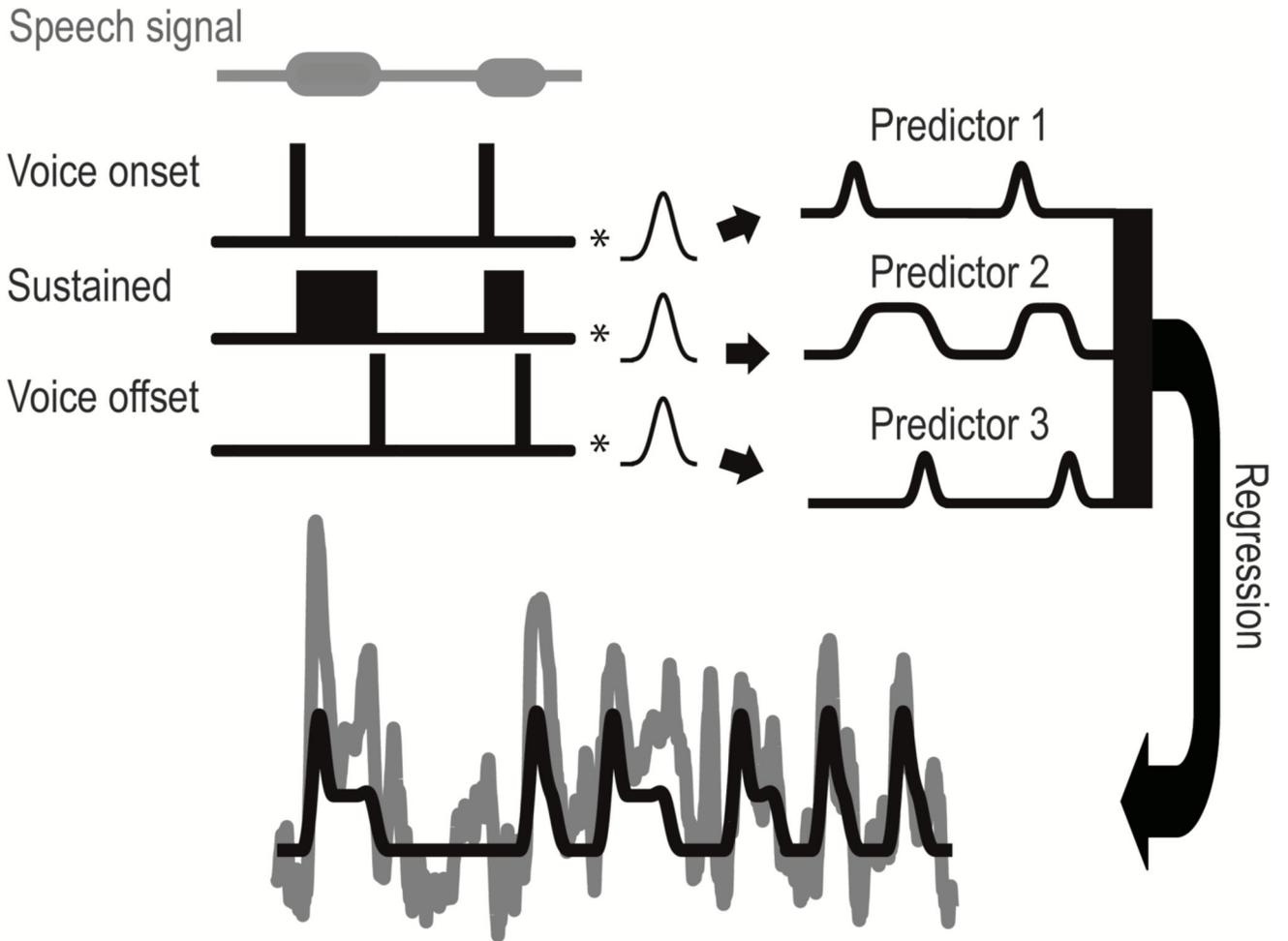
This work was supported by the European Union (ERC-Advanced 'iConnect' grant 320708).

## References

- [1]. Kandel E, Schwartz J, Jessell T. Principles of Neural Science. McGraw-Hill Medical; 2000.
- [2]. Taylor CSR, Gross CG. Twitches Versus Movements: A Story of Motor Cortex. *The Neuroscientist*. 2003 Oct; 9(5):332–342. [PubMed: 14580118]
- [3]. Dechent P, Frahm J. Functional somatotopy of finger representations in human primary motor cortex. *Hum Brain Mapp*. 2003 Apr; 18(4):272–283. [PubMed: 12632465]
- [4]. Miller KJ, Zanos S, Fetz EE, den Nijs M, Ojemann JG. Decoupling the Cortical Power Spectrum Reveals Real-Time Representation of Individual Finger Movements in Humans. *J Neurosci*. 2009 Mar; 29(10):3132–3137. [PubMed: 19279250]
- [5]. Siero JC, Hermes D, Hoogduin H, Luijten PR, Ramsey NF, Petridou N. BOLD matches neuronal activity at the mm scale: A combined 7 T fMRI and ECoG study in human sensorimotor cortex. *NeuroImage*. 2014 Nov.101:177–184. [PubMed: 25026157]
- [6]. Hadoush H, Sunagawa T, Nakanishi K, Endo K, Ochi M. Motor somatotopy of extensor indicis proprius and extensor pollicis longus. *Neuroreport*. 2011 Aug; 22(11):559–564. [PubMed: 21673606]
- [7]. Espadaler J, Rogi M, Deletis V, Leon A, Quijada C, Conesa G. Representation of cricothyroid muscles at the primary motor cortex (M1) in healthy subjects, mapped by navigated transcranial magnetic stimulation (nTMS). *Clin Neurophysiol*. 2012 Nov; 123(11):2205–2211. [PubMed: 22621909]
- [8]. Truccolo W, Friehs GM, Donoghue JP, Hochberg LR. Primary Motor Cortex Tuning to Intended Movement Kinematics in Humans with Tetraplegia. *J Neurosci*. 2008 Jan; 28(5):1163–1178. [PubMed: 18234894]
- [9]. Wang W, Chan SS, Heldman DA, Moran DW. Motor Cortical Representation of Position and Velocity During Reaching. *J Neurophysiol*. 2007 Jun; 97(6):4258–4270. [PubMed: 17392416]
- [10]. Georgopoulos, Kalaska JF, Caminiti R, Massey JT. On the relations between the direction of two-dimensional arm movements and cell discharge in primate motor cortex. *J Neurosci*. 1982 Nov; 2(11):1527–1537. [PubMed: 7143039]
- [11]. Moran DW, Schwartz AB. Motor Cortical Representation of Speed and Direction During Reaching. *J Neurophysiol*. 1999 Nov; 82(5):2676–2692. [PubMed: 10561437]
- [12]. Georgopoulos, Ashe J, Smyrnis N, Taira M. The Motor Cortex and the Coding of Force. *Science*. 1992; 256(5064):1692–1695. [PubMed: 1609282]
- [13]. Donoghue JP, Sanes JN, Hatsopoulos NG, Gaál G. Neural Discharge and Local Field Potential Oscillations in Primate Motor Cortex During Voluntary Movements. *J Neurophysiol*. 1998 Jan; 79(1):159–173. [PubMed: 9425187]
- [14]. Tanji J, Evarts EV. Anticipatory activity of motor cortex neurons in relation to direction of an intended movement. *J Neurophysiol*. 1976 Sep; 39(5):1062–1068. [PubMed: 824409]

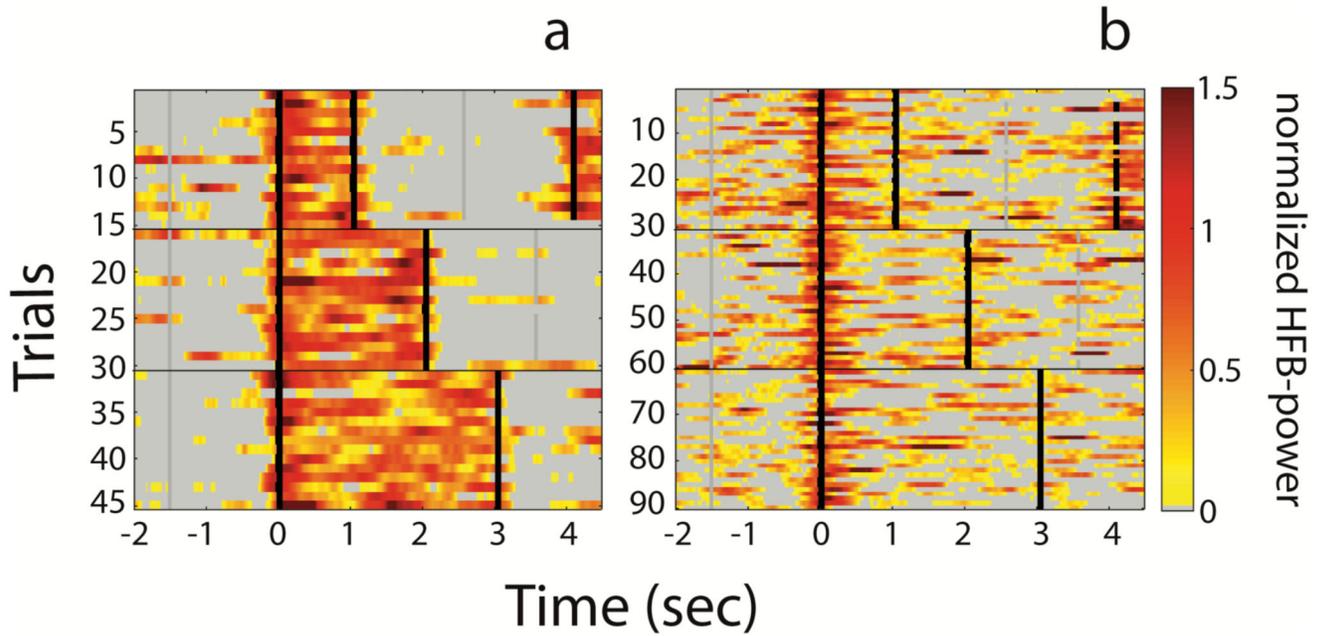
- [15]. Hermes D, Siero JCW, Aarnoutse EJ, Leijten FSS, Petridou N, Ramsey NF. Dissociation between Neuronal Activity in Sensorimotor Cortex and Hand Movement Revealed as a Function of Movement Rate. *J Neurosci*. 2012 Jul; 32(28):9736–9744. [PubMed: 22787059]
- [16]. Cheney PD, Fetz EE. Functional classes of primate corticomotoneuronal cells and their relation to active force. *J Neurophysiol*. 1980 Oct; 44(4):773–791. [PubMed: 6253605]
- [17]. Fetz EE, Finocchio DV, Baker MA, Soso MJ. Sensory and motor responses of precentral cortex cells during comparable passive and active joint movements. *J Neurophysiol*. 1980 Apr; 43(4): 1070–1089. [PubMed: 6766994]
- [18]. Soso MJ, Fetz EE. Responses of identified cells in postcentral cortex of awake monkeys during comparable active and passive joint movements. *J Neurophysiol*. 1980 Apr; 43(4):1090–1110. [PubMed: 6766995]
- [19]. Ball T, et al. Movement related activity in the high gamma range of the human EEG. *NeuroImage*. 2008 Jun; 41(2):302–310. [PubMed: 18424182]
- [20]. Crone NE, Miglioretti DL, Gordon B, Lesser RP. Functional mapping of human sensorimotor cortex with electrocorticographic spectral analysis. II. Event-related synchronization in the gamma band. *Brain*. 1998 Dec; 121(12):2301–2315. [PubMed: 9874481]
- [21]. Conant DF, Bouchard KE, Leonard MK, Chang EF. Human sensorimotor cortex control of directly-measured vocal tract movements during vowel production. *J Neurosci*. 2018 Feb.;2382–17.
- [22]. Wang PT, , et al. Electrocorticogram encoding of upper extremity movement duration. 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society; 2014. 1243–1246.
- [23]. Wang PT, et al. Characterization of electrocorticogram high-gamma signal in response to varying upper extremity movement velocity. *Brain Struct Funct*. 2017 Nov; 222(8):3705–3748. [PubMed: 28523425]
- [24]. Bouchard KE, Mesgarani N, Johnson K, Chang EF. Functional organization of human sensorimotor cortex for speech articulation. *Nature*. 2013 Mar; 495(7441):327–332. [PubMed: 23426266]
- [25]. Crone NE, et al. Electrographic gamma activity during word production in spoken and sign language. *Neurology*. 2001 Dec; 57(11):2045–2053. [PubMed: 11739824]
- [26]. Pei X, Leuthardt EC, Gaona CM, Brunner P, Wolpaw JR, Schalk G. Spatiotemporal dynamics of electrocorticographic high gamma activity during overt and covert word repetition. *NeuroImage*. 2011 Feb; 54(4):2960–2972. [PubMed: 21029784]
- [27]. Towle VL, et al. ECoG gamma activity during a language task: differentiating expressive and receptive speech areas. *Brain*. 2008 Aug; 131(8):2013–2027. [PubMed: 18669510]
- [28]. Bleichner MG, Jansma JM, Salari E, Freudenburg ZV, Raemaekers M, Ramsey NF. Classification of mouth movements using 7 T fMRI. *J Neural Eng*. 2015; 12(6):66026.
- [29]. Kellis S, Miller K, Thomson K, Brown R, House P, Greger B. Decoding spoken words using local field potentials recorded from the cortical surface. *J Neural Eng*. 2010 Oct.7(5):56007.
- [30]. Mugler EM, et al. Direct classification of all American English phonemes using signals from functional speech motor cortex. *J Neural Eng*. 2014 Jun.11(3):35015.
- [31]. Miller KJ, et al. Spectral Changes in Cortical Surface Potentials during Motor Movement. *J Neurosci*. 2007 Feb; 27(9):2424–2432. [PubMed: 17329441]
- [32]. Manning JR, Jacobs J, Fried I, Kahana MJ. Broadband shifts in LFP power spectra are correlated with single-neuron spiking in humans. *J Neurosci Off J Soc Neurosci*. 2009 Oct.29(43):13613.
- [33]. Miller KJ, Sorensen LB, Ojemann JG, den Nijs M. Power-Law Scaling in the Brain Surface Electric Potential. *PLOS Comput Biol*. 2009 Dec.5(12):e1000609. [PubMed: 20019800]
- [34]. Ray S, Maunsell JHR. Different Origins of Gamma Rhythm and High-Gamma Activity in Macaque Visual Cortex. *PLOS Biol*. 2011 Apr.9(4):e1000610. [PubMed: 21532743]
- [35]. Booij G. *The Phonology of Dutch*. Clarendon Press; 1999.
- [36]. Rietveld A, van Heuven V. *Algemene Fonetiek*. 2nd ed. Bussum: Coutinho; 2001.
- [37]. Penfield W, Boldrey E. Somatic motor and sensory representation in the cerebral cortex of man as studied by electrical stimulation. *Brain J Neurol*. 1937; 60:389–443.

- [38]. Liu Y, Coon WG, de Pestors A, Brunner P, Schalk G. The effects of spatial filtering and artifacts on electrocorticographic signals. *J Neural Eng.* 2015; 12(5):56008.
- [39]. Branco M, Gaglianese A, Hermes D, Saad ZS, Petridou N, Ramsey NF. Pipeline for ECoG electrode localization on brain surface: towards a one click approach. *ResearchGate.* 2016
- [40]. Hermes D, Miller KJ, Noordmans HJ, Vansteensel MJ, Ramsey NF. Automated electrocorticographic electrode localization on individually rendered brain surfaces. *J Neurosci Methods.* 2010 Jan; 185(2):293–298. [PubMed: 19836416]
- [41]. Fischl B. FreeSurfer. *NeuroImage.* 2012 Aug.62(2):774. [PubMed: 22248573]
- [42]. Bruns A. Fourier-, Hilbert- and wavelet-based signal analysis: are they really different approaches? *J Neurosci Methods.* 2004 Aug; 137(2):321–332. [PubMed: 15262077]
- [43]. Hermes DJ. Vowel-onset detection. *J Acoust Soc Am.* 1990 Feb; 87(2):866–873. [PubMed: 2307780]
- [44]. Boersma P. Praat, a system for doing phonetics by computer. *Glott Int.* 2002; 5
- [45]. Coudé G, et al. Neurons Controlling Voluntary Vocalization in the Macaque Ventral Premotor Cortex. *PLOS ONE.* 2011 Nov.6(11):e26822. [PubMed: 22073201]
- [46]. Allison JD, Meador KJ, Loring DW, Figueroa RE, Wright JC. Functional MRI cerebral activation and deactivation during finger movement. *Neurology.* 2000 Jan; 54(1):135–135. [PubMed: 10636139]
- [47]. Draper MH, Ladefoged P, Whitteridge D. Respiratory Muscles in Speech. *J Speech Lang Hear Res.* 1959 Mar; 2(1):16–27.
- [48]. Bouchard KE, et al. High-Resolution, Non-Invasive Imaging of Upper Vocal Tract Articulators Compatible with Human Brain Recordings. *PLOS ONE.* 2016; 11(3):e0151327. mrt. [PubMed: 27019106]
- [49]. Hage SR, Nieder A. Single neurons in monkey prefrontal cortex encode volitional initiation of vocalizations. *Nat Commun.* 2013 Sep.4:2409. [PubMed: 24008252]
- [50]. Tang C, Chehayeb D, Srivastava K, Nemenman I, Sober SJ. Millisecond-Scale Motor Encoding in a Cortical Vocal Area. *PLOS Biol.* 2014 Dec.12(12):e1002018. [PubMed: 25490022]
- [51]. Culham J. Cortical Areas Engaged in Movement: Neuroimaging Methods A2International Encyclopedia of the Social & Behavioral Sciences. Second Edition. Wright James D, editorOxford: Elsevier; 2015. 21–29.
- [52]. Rizzolatti G, Luppino G. Premotor Cortex A2International Encyclopedia of the Social & Behavioral Sciences. Second Edition. Wright James D, editorOxford: Elsevier; 2015. 846–851.
- [53]. Vargas-Irwin CE, Franquemont L, Black MJ, Donoghue JP. Linking Objects to Actions: Encoding of Target Object and Grasping Strategy in Primate Ventral Premotor Cortex. *J Neurosci.* 2015 Jul; 35(30):10888–10897. [PubMed: 26224870]
- [54]. Brumberg JS, Wright EJ, Andreasen DS, Guenther FH, Kennedy PR. Classification of Intended Phoneme Production from Chronic Intracortical Microelectrode Recordings in Speech-Motor Cortex. *Front Neurosci.* 2011 May.5
- [55]. Herff C, et al. Brain-to-text: decoding spoken phrases from phone representations in the brain. *Neural Technol.* 2015; 9:217.
- [56]. Camuñas-Mesa LA, Quiroga RQ. A Detailed and Fast Model of Extracellular Recordings. *Neural Comput.* 2013 May; 25(5):1191–1212. [PubMed: 23470125]



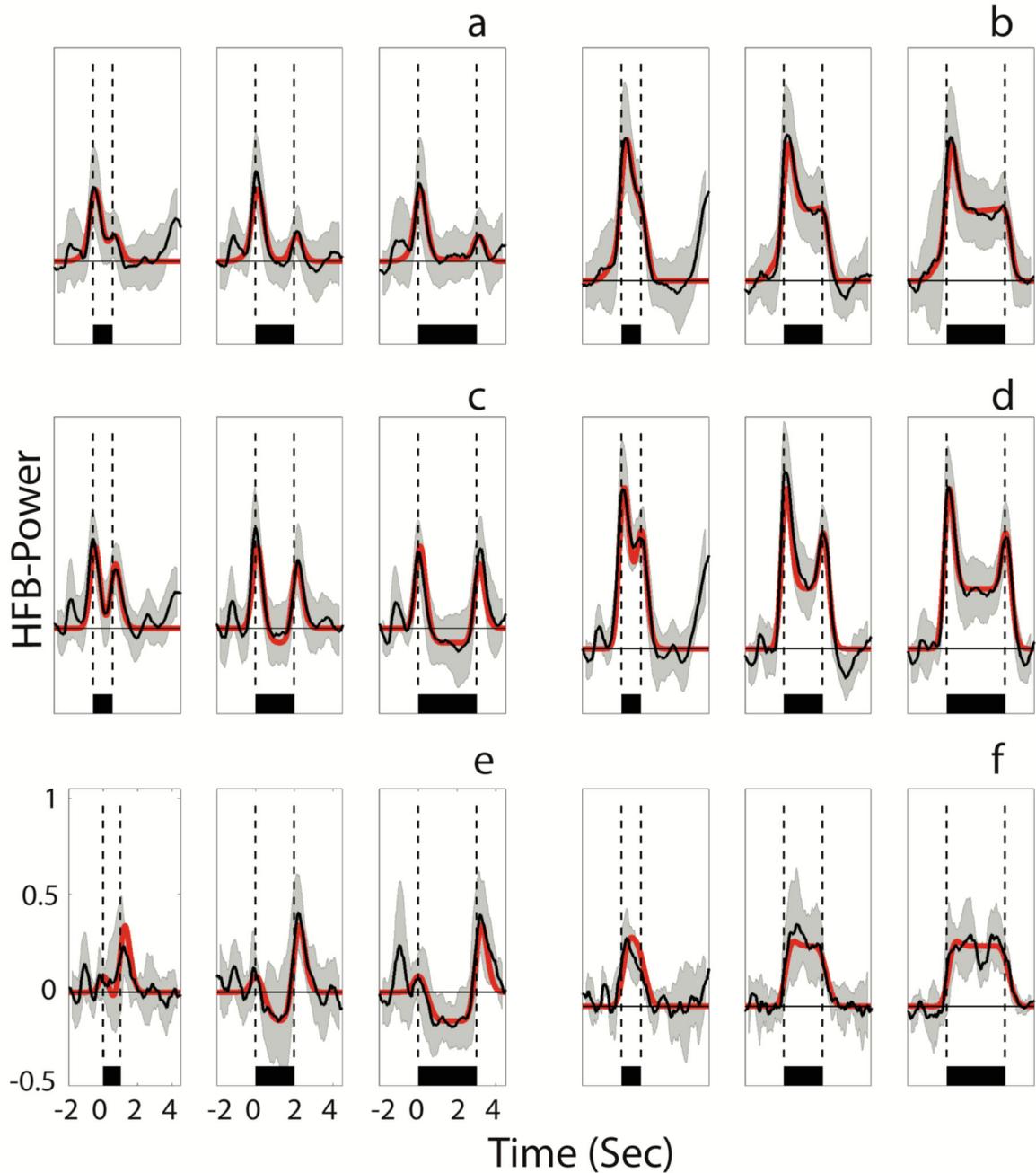
**Fig. 1.**

The regression procedure with the used predictors for the whole time series. The first and third predictor (right panel) were associated with voice onset and offset and were created by convolving an impulse function (left panel) at voice onsets and voice offsets respectively, with a subject specific Gaussian function over the whole time series. The second predictor, which was related to sustained activity during vowel pronunciation, was created by convolving a block function (left) between voice onset and voice offset with the same Gaussian function. Horizontal bars at the top (speech signal) indicate the time periods where subjects spoke. A regression was performed and the signal (gray) and model thereof (black) are shown at the bottom.



**Fig. 2.**

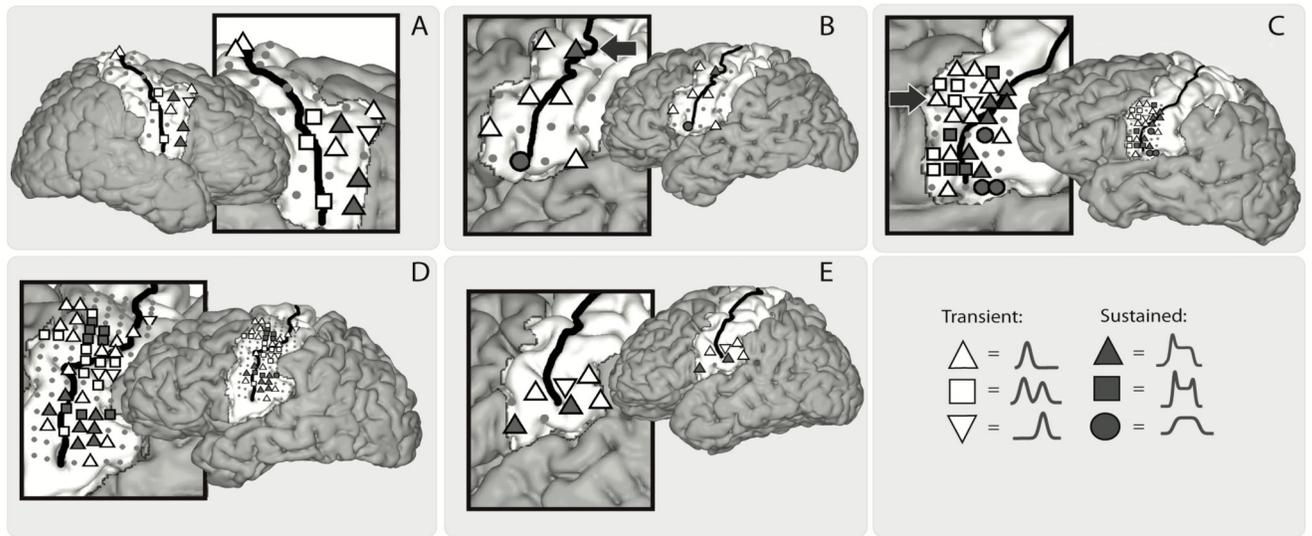
Two examples of the inter-trial variability. (a) Example of an electrode with a sustained response (see arrow in Figure 4b). (b) Example of an electrode with a transient response (see arrow in Figure 4c). On the x-axis, time is indicated with zero being voice onset. On the y-axis, individual 1, 2 and 3 second trials are shown. The color-scale indicates the normalized HFB-power. The vertical black lines indicate voice onset or voice offset. The gray vertical lines indicate the trial duration cue. Note that in the one-second trials, part of the subsequent trial is visible towards the end. Also, note that the trials are aligned relative to voice onset, and that deviations in voice offset from the intended duration were corrected. In (b) twice the number of trials is shown since this subject performed the task twice.



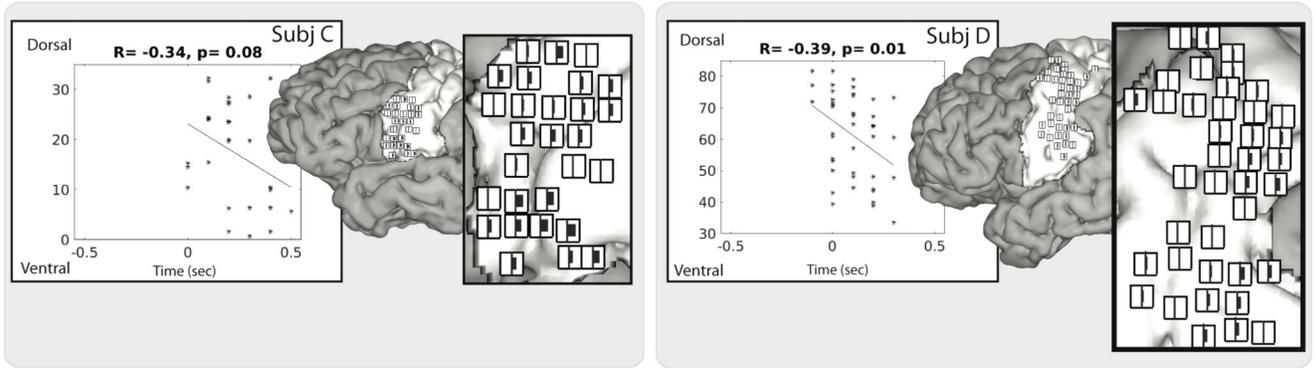
**Fig. 3.**

A visualization of the six response patterns observed in the data, averaged over all subjects. (a) Transient at voice onset, (b) sustained and transient at voice onset, (c) transient at voice onset and offset, (d) sustained and transient at voice onset and offset, (e) transient at voice offset, (f) sustained. In all panels, time is indicated on the x-axis, with zero being voice onset. On the y-axis, the normalized HFB-power is indicated. The mean of all traces that were classified as belonging to a specific class is shown in black, with the standard deviation indicated in light gray. For every electrode, based on the beta values of the three predictors a

best fit model was created (see Figure 1) and the red line shows the mean of all electrode models. The horizontal black bar shows the duration of speech, being 1, 2 and 3 seconds from left to right, respectively. Vertical dashed lines indicate voice onset and voice offset. Note that on average, neural activity already started before voice onset.



**Fig. 4.**  
A visualization of the electrode positions (grey dots) and the response types, per subject (A-E). Symbols indicate the response type of an electrode. Arrows in B and C indicate the electrodes that are shown in Figure 2.



**Fig. 5.**

A visualization of the model shift timing for the subjects with HD-grid implantation (subject C and D). Scatter plots with time in seconds on the x-axis (zero being voice onset), and the ventral-dorsal MRI coordinates on the y-axis, indicate the correlation between model shift (a marker for the response onset) and ventral-dorsal brain position for each electrode. The anatomy plots show the locations of the electrodes and their timing relative to voice onset. The vertical midline represents voice onset and the horizontal bars indicate the model shift (black bar to the right indicates a peak of activity after voice onset, black bar to the left indicates peak activity prior to voice onset, with a larger bar indicating a larger temporal difference with voice onset). Note that the ventral parts are not well covered in subject D while this is the area in subject C with the latest responses. Note also that these timings correspond to timing of the first peak, meaning that a shift time on or near voice onset (vertical midline) reflects a rising activation before voice onset.