UNIL | Université de Lausanne
Faculté de biologie
et de médecine

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Master Thesis in Medicine No
5709

Theta waves in speech perception, imagination and production

Student
Jaquet Florian

Tutor
Prof. Millán José
Defitech Foundation Chair in Brain-Machine Interface
Center for Neuroprosthetics and Institute for Bioengineering
EPFL & WYSS CENTER Geneva

Co-tutor
Martin Stéphanie, PhD
Defitech Foundation Chair in Brain-Machine Interface
Center for Neuroprosthetics and Institute for Bioengineering
EPFL & WYSS CENTER Geneva

Expert
Prof. Clarke Stephanie
Neuropsychology and Neurorehabilitation Clinic & Laboratory for Cognitive Sciences
CHUV

Lausanne, 15.01.2018

Abstract

Language is ingrained in our thought process and consciousness. It takes several forms, of which speech is the most spontaneous. Oral communication is essential to exchange information, emotions, and abstract ideas. Being unable to speak cuts one away from society, causing great suffering for the patient and his relatives as well as a complicating clinical care. People that still have an intact cerebral function but can't communicate, like patients with neurological disease – cerebral palsy, stroke or spinal cord injury and locked-in syndrome for example (2) – would greatly beneficiate from speech decoding devices. However, brain computer interfaces (BCIs) aren't at that stage to this day. Improving our understanding of how our brains treat language as well as refining decoding of speech-related activity will push BCIs into a new era, changing the quality of life of patients suffering from these debilitating conditions for good. To this day, different brain-waves patterns are known (3). Sorted by increasing frequency from delta to gamma, their role starts to be discovered (4, 5). Theta waves – ranging from 4 to 8 Hz – have been shown to be essential for speech intelligibility (6) and to correlate with speech signal envelope (7). In this project, the involvement of theta waves in listening and speech production tasks was explored further. Theta waves signal as well as signal envelope were explored in different speech-related conditions and their potential correlation to speech onset and offset was explored.

Electrocorticogram (ECoG) recording of brain activity of four patients undergoing surgery for severe epilepsy were obtained. ECoG is little invasive and offers an excellent time and space sensitivity. Acquired data encompasses three conditions in which patients listened, imagined, or repeated words or sentences. This data and the concomitant audio recording were analysed using Matlab®. ECoG signal was filtered for theta range (4-8Hz), the envelope was taken, and both signal and envelope were normalised to detect significant changes. Brain heat maps were then computed to assess spatio-temporal changes in theta activity and average changes over the different tasks. Finally, to assess the potential of theta activity as speech onset and offset biomarker – an important feature to improve speech decoding BCIs –, the correlation between theta waves and audio recordings was investigated. A linear regression model linking different inputs (theta signal, theta envelope) to outputs (audio signal, audio envelope, trigger values – "ongoing task markers") was tested.

Results indicate that theta activity and theta envelope is significantly modulated in listening-related tasks. Marked decreases in theta activity and envelope were observed in all experimental set-ups, with focal increases in theta activity in the left primary auditory cortex during listening. Average theta envelope in tasks in which subjects were to listen or repeat words was modulated with regard to baseline in both hemispheres. Theta envelope correlated well to audio envelope and "markers of ongoing task" in these cases but less so when tasks involved imagined and overt speech.

Thus, theta envelope seems to be a good marker of listening tasks but appears to be a poor marker of speech intention.

Key words

Electrocorticogram (ECoG), Language decoding, Theta waves

Table of contents

Introduction

Communication takes many forms, stressing its importance for living organisms. From chemical signals to touch, gesture, or sounds, all convey a wide range of information. However, language could well be a unique human feature (8). It is ingrained in our thought process and consciousness, participating to the formulation, expression, and comprehension of ideas. At an inter-individual level, language relies on both speech and written materials. These allowed humans to develop complex thinking and exchange abstract ideas. However, speech remains the most natural and spontaneous form of expression. As such, being unable to speak makes day-to-day life complicated. Communication can be considerably slowed down and prone to increased misunderstandings. This situation can be a source of great suffering for patients and their relatives as well as a cause of complications in clinical care. Sadly, millions of patients worldwide are suffering from neurological insults like cerebral palsy, stroke, amyotrophic lateral sclerosis or spinal cord injury and locked-in syndrome for example (2, 9). These patients sometimes lose their ability to speak despite potentially fully functional language brain areas. Despite the difficulty of their situation, the technology they can turn to for help is still limited.

To this day, brain computer interfaces (BCIs) can register and use a wide range of cerebral signals to control various types of downstream devices (Figure 1) – from a muscle or a simple cursors on a screen (10) to more complex limb prostheses (11) or robotic apparatus (1). BCIs have however had only limited success with speech decoding so far (12) and allow communication mostly through alternative means. Thus, efforts are still needed to achieve flowing and effortless interactions. Gaining a greater understanding of how our brains treat language as well as improving decoding of speech-related brain activity will push BCIs into a new era and offer better solutions to improve quality of life for patients suffering from these debilitating situations.

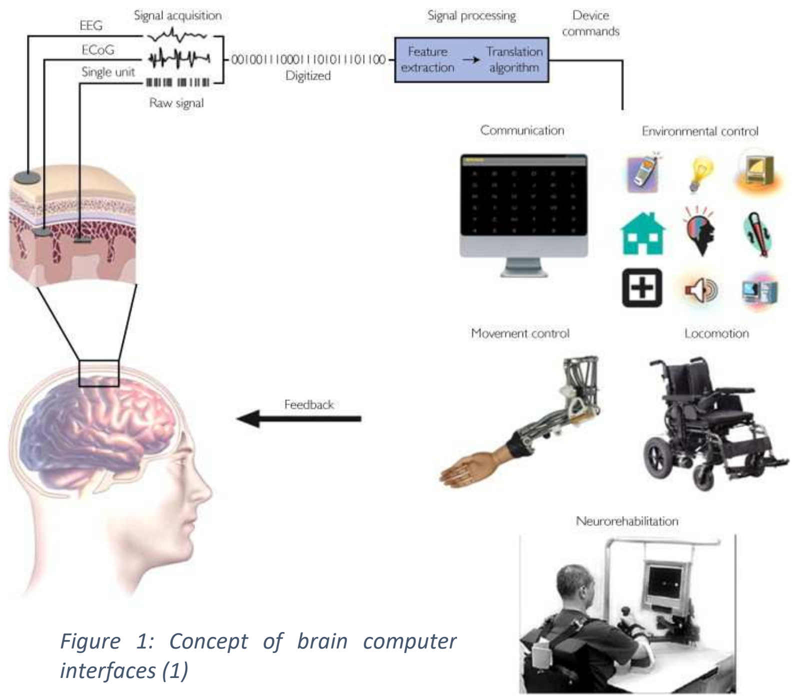

Figure 1: Concept of brain computer interfaces (1)

Numerous areas throughout the brain are related to language, mostly located in the cerebral cortex (Figure 2). The most emblematic ones have been discovered more than a century ago. Named after their discoverers, Broca's and Wernicke's areas are usually found in the left hemisphere. Wernicke's area is located on Brodmann area 22 at the junction between the parietal and temporal lobes, over the posterior superior and middle temporal gyrus as well as the superior temporal sulcus. Broca's area lies in the postero-ventral part of the frontal lobe, on Brodmann areas 44 and 45 (13). Broca's area plays a crucial role in language generation, especially in grammar and syntax. Wernicke's area on the other hand, is important for comprehension. Over the years several additional areas in both hemispheres – e.g. Brodmann area 21 in the left lateral temporal cortex – have been found to play a role in the different aspects of language (14). These aspects include, among others, overt and imagined speech. Speech imagery – the act of "speaking in one's head" without physical speech movements (15), also termed covert or imagined speech – is thought to rely to some extent on the same brain areas as overt – or actual – speech, as this is the case for other imagined experiences like limb movement (16). However, the brain activity pattern in speech imagery is expected to be more complex than mere overt speech without physical movements. Indeed, previous imaging and functional experiments have shown differences between overt and covert speech (17, 18).

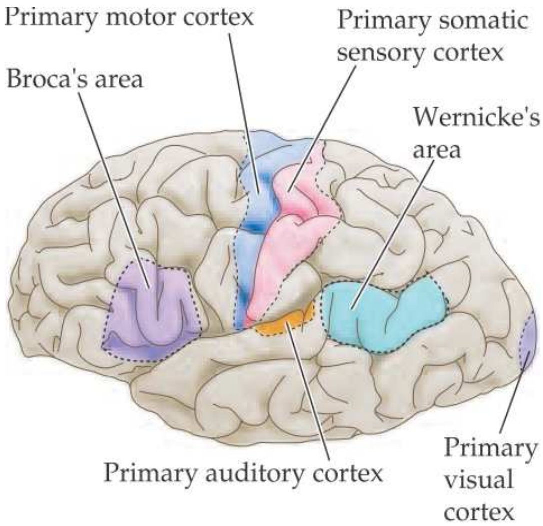

Figure 2 : Principal language associated areas (Purves Neurosciences 3rd Ed)

Investigating those differences between overt and covert speech can be done using different techniques depending on the feature of interest. In the case of brain activity, numerous technologies are available and suitable. For example, imaging techniques such as positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) are often used. However as they reflect vascular and metabolic activity they offer a low temporal resolution (14). They are therefore poorly suited for studying rapid changes in brain activity during language generation. Another popular investigation tool is electroencephalography (EEG). It offers a good temporal resolution but with limited spatial discrimination. EEG is appreciated for its non-invasiveness but is also susceptible to artefacts like electromyographic signals. Yet another existing option is single-neuron recording which has a great spatial resolution. However, it requires numerous electrode implantations and will trigger tissue responses affecting their long-term stability (19). A good compromise combining relatively low invasiveness to an excellent time and space sensitivity is electrocorticography (ECoG). Applied directly on the surface of the brain, electrodes record cortical activity with a great signal-to-noise ratio. Indeed, compared to EEG – in which electrodes are placed on the scalp, ECoG offers an increased signal magnitude, spatial resolution and frequency bandwidth (20). Moreover, ECoG has already successfully been used to decipher and classify brain activity linked to phonemes and words (4, 5, 20) and to assess attention towards speech streams (7). ECoG has also been used to analyse real-life speech related activity (21) and to try to achieve direct speech synthesis (22).

Data obtained from ECoG recordings represents brain waves. These brain waves were arbitrarily separated in frequency bands that range from delta waves (0.5-4 hertz) to gamma waves (30 hertz and above). In between, one can find theta (4-8 hertz), alpha (8-12 hertz) and beta waves (12-30 hertz) (3). All rhythms and bandwidths reflect underlying neurological phenomena. Previous work in our laboratory has focused on the high-gamma wave range known to be associated with spike rate and local field potentials of underlying neuronal populations as well as in tracking rapid neuronal variations during speech production and perception (4, 5). In this work, we focused mainly on theta waves. Theta waves have been linked to will (3) and memory in the hippocampus, but their role in language is mostly unknown. So far, theta waves have been shown to be linked to underlying neuronal firing (3) and to be essential for speech intelligibility (6). Moreover, low frequency waves like theta correlate to speech-envelope time scale fluctuation (7) and are thus involved in attention towards speech. Different features of theta activity start to be investigated. It seems for example that theta waves power band – an indication of relative amplitudes of the signal – diminishes during speech and that its phase and amplitude is modulated during speech (23, 24). Another feature of signal that can be researched is its envelope, a parameter that outlines the extremes of the signal. Martin et al. have for example used the envelope of high gamma ECoG signal for word pair classification (5).

Different types of signal analysis can be performed. One method is to average brain signals over a given period of time, usually linked to a task (e.g. for 1 second during which the subject speaks). This averaged signal can then be compared to the average signal in other conditions. These comparisons are informative on general brain function and are often a first step of analysis, as in this project. However, this misses fine temporal dynamics. Finer investigation of brain activity in complex conditions and naturalistic environments is needed to find functional BCI inputs like speech on-off status. The simplest method to link brain activity to a stimulus or action is to perform a regression. This can be done in both directions and can thus also be used to link a given stimulus to a resulting brain state. As described by Stéphanie Martin (25), the regression process can be broken up into 4 steps. First, features are extracted and sorted into input and output. A typical input would be amplitude in a specific frequency band extracted from a recorded brain activity, while an exemplary output could be concomitant audio signal. A regression model is then estimated using these features, the output being mapped to the input, trying to minimise some error measure (e.g. least squares). Most often, these regression models are linear, meaning that they are constructed as a weighted sum of input features. These models must then be validated (or tested) on data that was not used for fitting (also called training) the model. At this stage, the outputs predicted by the model are compared to the actual measured output and errors measured. Finally, the model can be interpreted, the weights of the model showing the relationship between input and output features, e.g. between brain activity and audio signal. These steps describe a typical encoding model based on a regression and would allow to investigate how the input (e.g. audio signal) influences the output (e.g. brain activity). By reversing input and output, one can build a decoding model. This can then be used to infer the new output (e.g. audio signal) from the new input (e.g. brain activity). It is this decoding ability that can then be used to operate neuroprosthetic devices.

Thus, the scope of this project was to build on current understanding of speech generation and imagination to improve language decoding. Expanding from former studies in our laboratory (4, 5), **this project explored language processing further by investigating the role of theta waves and theta envelope in speech perception, imagination, and production**. Theta waves and envelope might contain additional information or encode different language characteristics than high-gamma waves. This project investigated two condition sets. A set of single conditions encompassing passive listening and active, overt repetition of sentences on the one hand, and a set with a triple condition in which successive passive listening, covert repetition, and overt repetition were performed. In the first part of the project, the neural time course of theta waves and envelope was studied. The spatio-temporal dynamic of theta activity and envelope was mapped on 3-dimensional brain models. In a second part, average theta activity and envelope over the whole simple condition set were projected on 3-dimensional brain models to get an understanding of their role in these situations. Finally, the last part of analysis consisted of computing a linear regression model in both sets of conditions. Indeed, as theta waves are known to track speech envelope (7, 26), we evaluated the opportunity of using this frequency band as a biomarker for the speech state in an on-off manner. This would be a step towards more natural communication than the one presently offered by BCIs.

Materials and methods

The experimental set-up and methods are described in Martin, Brunner (5). Data was provided by the Schalk lab, New-York, USA (www.schalklab.org). Four subjects undergoing surgery for severe epilepsy had pre-surgery ECoG implantation to characterize their epileptic centres. They had 90 electrodes sealed to their scalp. These were grids of platinum-iridium electrodes (4 mm in diameter, 2.3 mm exposed) embedded in silicon and with an inter-electrode distance ranging from 4 to 10 mm. Grid placement and duration of implantation was based solely on clinical needs. ECoG signals were recorded at bedside using seven 16-channel g.USBamp biosignal acquisition devices (g.tec, Graz, Austria) at a sampling rate of 9,600 Hz. Electrode contacts distant from epileptic foci and areas of interest were used for reference and ground. Data acquisition and synchronization with the task presentation were accomplished using BCI2000 software. All electrodes were subsequently downsampled to 1,000 Hz, corrected for DC shifts, and band pass filtered from 0.5 to 200 Hz. Notch filters at 60 Hz, 120 Hz and 180 Hz were used to remove electromagnetic noise. The time series were then visually inspected to remove the intervals containing ictal activity as well as electrodes that had excessive noise (including broadband electromagnetic noise from hospital equipment or poor contact with the cortical surface). Finally, electrodes were re-referenced to a common average. Imagined speech trials were carefully analysed to remove those that were contaminated by overt speech by examining the audio recordings. Overt speech trials that had grammar mistakes were also removed.

In addition to the ECoG signals, the subject's voice was acquired through a dynamic microphone (Samson R21s) that was rated for voice recordings (bandwidth 80–12,000 Hz, sensitivity 2.24 mV/Pa) and placed within 10 cm of the patient's face. We used a dedicated 16-channel g.USBamp to amplify and digitize the microphone signal in synchrony with the ECoG data. Finally, we verified the patient's compliance in the imagined task using an eye-tracker (Tobii T60, Tobii Sweden) to ensure the patient was paying attention to the visual cues given.

Experimental protocols

While under monitoring, patients consented to go through three different experimental settings as follows:

- "Passive listening". In this situation, the patient is exposed to pre-recorded 1.8-second-long sentences that were played upon appearance of a visual cue (a cross in the middle of the screen). Patients had no other task than to listen (Figure 3). These sentences were phonetically transcribed stimuli from the Texas Instruments/Massachusetts Institute of Technology (TIMIT) database (~2s, 16 kHz). Stimuli were presented aurally at the patient's bedside using either external free-field loudspeakers or calibrated ear inserts at approximately 70–80 dB as described by Pasley, David (27).

- "Sentence repetition". The situation was very similar to the one in "Passive listening", with the exception that patients heard a sentence that they were to repeat on appearance of the visual cue.

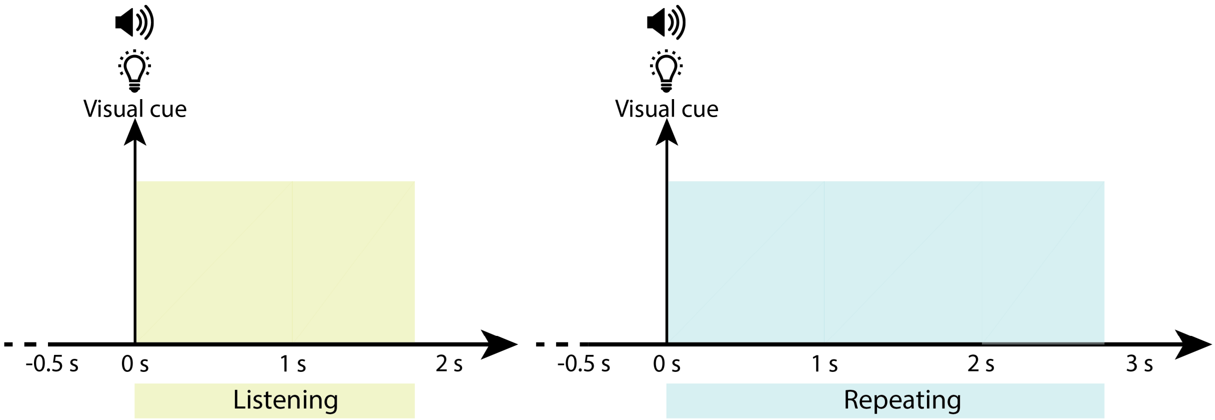

Figure 3: Experimental protocol for passive listening and sentence repetition conditions. They differ in length as a bit more time is given for the patient to repeat the heard stimulus. Trigger values used for regression are 1 during trial (coloured) and 0 otherwise.

The next set-up involved listening, imagined speech and overt repetition.

- "Words 3 conditions". In this case, conditions are identical to those described by Martin, Brunner (4). Each trial started with an auditory word stimulus presented through a loudspeaker (listening condition) – indicating one of 6 individual words (average length = 800ms ± 20ms). Then, 1.5 sec after the auditory cue, a visual cue (a cross displayed for 1 sec) appeared on the screen, at which point the patient had to repeat the word silently – "in his head" (imagined condition). Finally, after a second of blank screen, a second visual cue appeared (a cross again, for 1.5 sec) and the patient had to repeat the word aloud (overt condition) as shown in Figure 4. By pacing the subject, the task was designed to minimize the behavioural variance in producing overt and imagined speech.

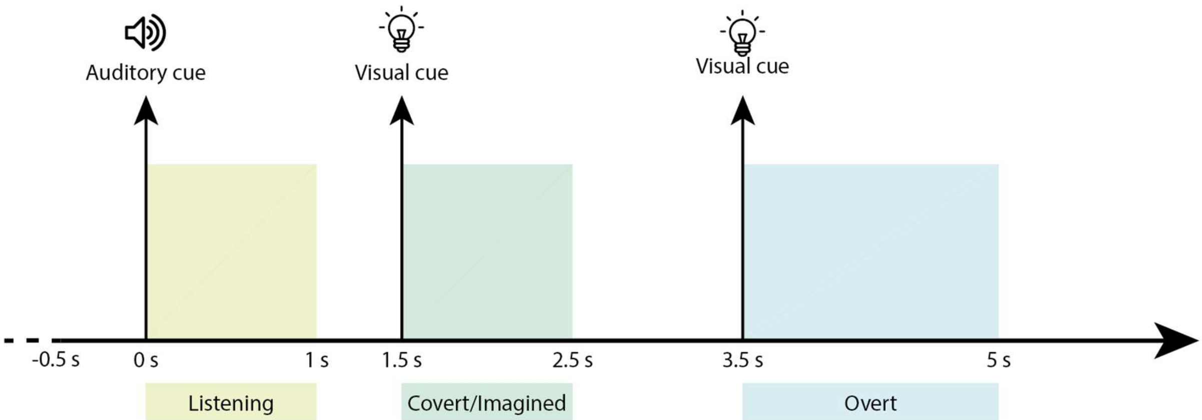

Figure 4: Experimental design of the words 3 condition. The word is played at t0. Duration of the word 800+/-20ms, rounded to one second. At t = 1.5s, a cross is displayed on screen prompting the patient to repeat the word in his head - covertly. Finally, at t = 3.5s, a second visual clue is displayed as the patient is to repeat this same word aloud – overtly. Trigger values used for regression would be 1 in coloured areas and 0 otherwise.

The first pair of conditions, termed "single" conditions, and the second – called "triple" condition, were then analysed separately.

Data analysis

The aim was to extract theta range activity from the data and study its spatio-temporal evolution through the different language related conditions. Moreover, theta envelope was extracted to analyse this feature of theta activity as well. Concomitant audio signal was analysed as well to permit comparison and correlation with given theta activities. Finally, to assess how well theta activity is correlated to different features of speech and listening, a regression with 5-fold cross-validation was computed against the recorded audio signal and its envelope. To investigate the potential of theta waves as a speech on-off marker, this regression was computed as well against trigger values that were 0 outside of trials and 1 during trials. To do so, data from the different experimental conditions was analysed using Matlab R2015a®. Electrodes that had shown epileptic activity at signal inspection were left out of analysis at this stage.

<u>Frequency filtering</u>

ECoG data was filtered for theta wave range (4-8Hz) using a 4th degree 8Hz Butterworth low pass filter, a 4th degree 4Hz Butterworth high pass filter and a notch filter. Signal was analysed in two different forms: filtered signal (thereafter referred to as theta activity) and the envelope of that filtered signal (theta envelope). The later was obtained using the absolute of the Hilbert transform of the original signal. Filtering the whole data at once ensured minimal added noise and aliasing from the filtering process.

Similarly, audio envelope was extracted from the audio signal. Both audio signal and envelope were then downsampled from the initial 16'000Hz to the frequency of the ECoG signal – 1000Hz. This frequency was set at 1000Hz for single electrode activity and spatio-temporal analysis. A frequency of 100Hz was chosen for the regression with cross-validation to reduce computational load.

Once audio and ECoG data was filtered, it was separated and sorted according to the different experimental conditions and trials. The obtained epochs included ECoG and audio signals from 500ms before trial onset to 500ms after the end of trial. These epochs were then averaged over all trials for each electrode. Some trials that had longer or shorter experimental conditions were disregarded at this state to conserve a timely aligned average.

The obtained average signal over all trials was then normalised. Theta activity and theta envelope z-scores were calculated for every condition (see Equation 1 below). To this aim, the mean of theta activity – respectively of theta envelope – over the 500ms pre-trial was calculated for each electrode. This gave a so called "mean baseline". This "mean baseline" was then subtracted from the data. The result was divided by the standard deviation of this "mean baseline". This allowed detection of electrodes with significant signal changes with respect to baseline. A change was said to be significant if the normalised signal was above 1.96 or below -1.96 (i.e. the signal is more than 2 standard deviations away from baseline). Different electrodes and conditions could then be compared.

$$\boldsymbol{Normalised\ signal} = \frac{|\boldsymbol{x}| - \mu(|baseline|)}{\boldsymbol{\sigma}(baseline)} \qquad \boldsymbol{Normalised\ envelope} = \frac{\boldsymbol{x} - \mu(baseline)}{\boldsymbol{\sigma}(baseline)}$$

Equation 1 : signal normalisation calculation. For theta signal, the absolute values were used as the signal oscillates around zero. This was not necessary for the envelope.

Brain heat map

Signal distribution averaged over time was displayed on a brain model using a customised Matlab code developed at the Schalk lab (www.schalklab.org). Average activity over the different experimental conditions was displayed, as well as time series showing the mean brain activity over a 100ms time window at different time points throughout the experimental conditions.

Speech states

Linear regressions and 5-fold cross-validation were used to build decoding models. This aimed at investigating the potential of theta signal and envelope as biomarkers of speech on and off states. Theta signal and theta envelope were used as inputs while audio signal, audio envelope and trigger values were the outputs. These trigger values indicate when a subject is performing a task (trigger value 1) or not (trigger value 0). To reduce computational load, only electrodes with significant theta brain activity changes (as shown by signal normalisation) in simple conditions were used at this stage and both theta and audio signals and envelopes were downsampled to 100Hz

These linear regressions were calculated on the data segmented in 500ms time intervals with a code from Stéphanie Martin and STRFLab MATLAB toolbox (http://strflab.berkeley.edu/). In each case, brain activity and audio recordings were segmented in 5 equal blocs of data. Each one was alternatively used as training bloc to fit the linear regression model and then tested on the remaining blocs. Thus, overall correlation coefficients could be obtained averaging the correlation obtained by all 5 models. Significance of the different calculations was assessed through a Z transformation (similar as the one performed in Equation 1).

Results

Neural time courses

Normalised theta activity as well as theta envelope were computed for all conditions and electrodes. This allowed to detect electrodes with significant changes in activity. The neural time courses of three exemplary electrodes are displayed in Figure 5. Depending on the electrode, significant theta envelope modulation was observed in all different experimental conditions. In Figure 5, each one of the three electrodes displayed has its own condition in which it shows significant changes in signal envelope. Other electrodes have shown significant activity in none, two, or all conditions. In each subject, a large majority of electrodes showed significant theta range activity modulation both in term of signal and signal envelope variation. This first step of analysis confirmed the basic assumption that theta waves are modulated according to task and brain area and that it might thus exhibit some degree of correlation to the audio signal. The few electrodes that didn't display significant theta envelope variation were disregarded for the regression analysis.

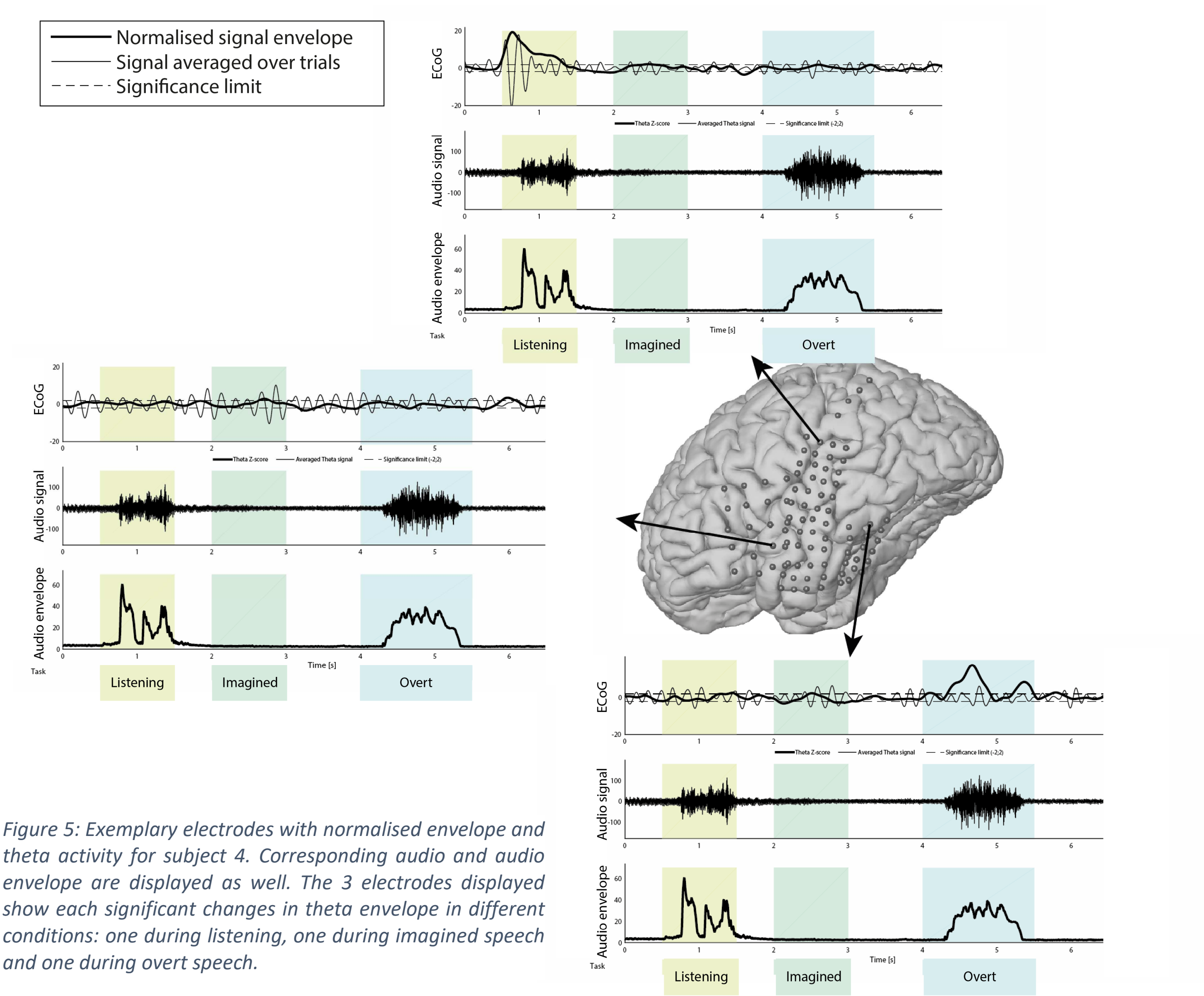

Figure 5: Exemplary electrodes with normalised envelope and theta activity for subject 4. Corresponding audio and audio envelope are displayed as well. The 3 electrodes displayed show each significant changes in theta envelope in different conditions: one during listening, one during imagined speech and one during overt speech.

The correlation of the different patterns of electrode activation to brain anatomy are best seen in the following brain heat maps

Brain heat maps

Time lapse

The brain heat map displayed below show theta activity (Figure 6, see supplementary Figure S1 for increased temporal resolution) or theta envelope (Figure 7, supplementary Figure S2) averaged over 100ms at different time points throughout the triple condition, where patients were to listen, think, and repeat a given word.

Theta activity was in general diminished in all brain areas and conditions (Figure 6). As passive listening begins, all subjects with left hemisphere recordings (1,2 and 4), show a focal increase in theta activity in the auditory cortex (see supplementary Figure S1). This soon vanishes and changes to a marked decrease by the end of passive listening. Other brain areas display a rather stable decrease in theta activity during the listening part (from 0 to 1 second). During the following imagined speech period, no marked changes in theta activity can be seen (from 1.5 to 2.5 seconds). Subjects 1 and 2 seem to have a slight further decrease in theta activity in the auditory cortex area over this period. Subject 2 shows a very restricted temporal theta activity increase at the beginning of the condition. Subject 3 and 4 seem to keep a stable decreased theta activity. The last part, from 3.5 to 5 second, displays brain heat maps during overt speech. Once again, no striking changes are seen. Subject 1 seems to have decreasing theta activity around the temporal sulcus over this overt speech condition. Subject 2 has some focal increases in the inferior temporal lobe and then in the auditory cortex area. Subject 3 expresses once again a stably decreased theta activity. Finally, subject 4's theta activity shows theta activity decreases of varying intensity in the auditory cortex area.

Brain heat map time lapse for theta envelope (Figure 7 and supplementary Figure S2) give a similar impression. Indeed, theta envelope is markedly decreased over the whole triple condition. During the listening part, theta envelope was focally increased in the pre-motor cortex and inferior temporal lobe in subject 2. All subjects exhibited otherwise decreased theta envelope activity. It is even markedly decreased in subject 1 around the lateral sulcus, encompassing the primary auditory cortex and the Wernicke area, at the end of the listening period. This decreased theta envelope was initially not present in the primary auditory cortex in subject 4 but appeared progressively during the listening phase. During imagined speech, theta envelope remained decreased in all subjects. Subjects 1 and 4 show opposite trends, with subject 1 showing milder decreases in the auditory cortex and Wernicke area with imagined speech going on while subject 4 had increasingly strong decrease of activity in those areas. Finally, by the beginning of overt speech, subject 1 and 3 showed increased theta envelope frontally. These increases soon vanished and a general decrease in theta envelope is seen in all subjects, markedly so around the lateral sulcus in subject 1 and 4. The magnitude of these decreases became less important in the post-experimental phase.

To summarise, theta signal and envelope are generally decreased in all brain areas during passive listening, imagined and overt speech. Observed modulations were essentially around the lateral sulcus – in the primary auditory cortex and the Wernicke area – and to a lesser degree and mostly for theta envelope, in the frontal cortex. The most consistent change is the observed increase in theta activity in these lateral sulcus areas during passive listening.

Figure 6: Brain heat map time lapse of theta activity. Mean activity over 100ms is shown every 1 sec for each subject. Diminishing activity is seen in red while increased activity is plotted in blue. Areas in grey have either no activity changes of were not initially covered by the implanted electrode grid.

Figure 7: Brain heat map time lapse of theta envelope. Mean activity over 100ms is shown every 1 sec for each subject. Diminishing activity is seen in red while increased activity is plotted in blue. Areas in grey have either no activity changes or were not initially covered by the implanted electrode grid.

Means over conditions

Brain heat maps were also computed with the mean theta envelope over the whole condition. These display average changes of theta envelope throughout the brain in the different experimental set-ups. Mean theta envelope in the listening and sentence repetition condition is displayed below (Figure 8). Mean theta activity was not shown as it oscillates around zero and thus shows only very limited changes.

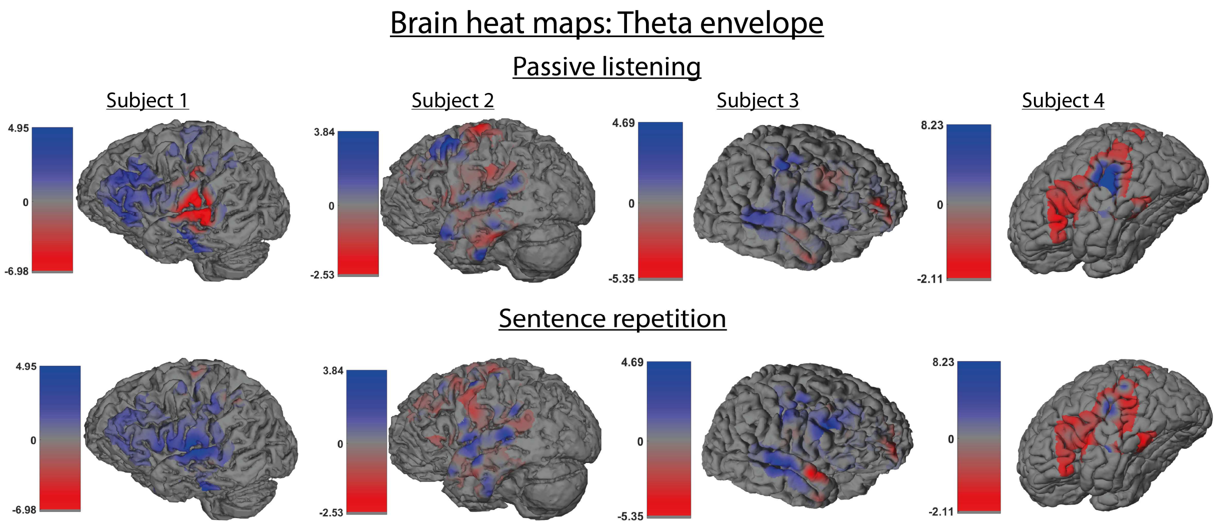

Figure 8: Brain heat maps displaying average theta envelope activity over passive listening and sentence repetition.

The amplitude of changes observed in the mean over a whole condition is far smaller than in the time lapse shown before (watch scale in Figure 8). In passive listening, average theta envelope activity was inconsistent between subjects. Regarding the Wernicke area and left auditory cortex, theta envelope was increased in subjects 2 and 4 while it was markedly decreased in subject 1. Frontal and Broca average theta envelope is conflicting as well, as subject 1 had an increase activity while it was decreased in subject 4. Of note, subject 2 shows an increase of theta envelope in the pre-motor cortex but a decrease in the motor cortex. Finally, changes in the right hemisphere, represented for subject 3, showed an increase of theta envelope in the auditory cortex as well as in the motor cortex with some focal decreases in the frontal cortex.

Marked changes between passive listening and sentence repetition are seen in subject 1. Indeed, in sentence repetition, theta envelope is generally increased, including in Wernicke's area and in the auditory cortex as well. In subject 2, changes are seen as well, although to a smaller degree. Mean theta envelope in the Wernicke area is increased in sentence repetition. A decrease in mean theta envelope can be seen as well in the motor cortex. In subject 4, mean theta envelope during sentence repetition is relatively similar to the one observed over passive listening. The increase of theta envelope is however much more restricted, but still encompasses part of the auditory cortex. In the right hemisphere (subject 3), the only noticeable difference is a focal decrease in theta envelope in the superior frontal part of the temporal lobe.

Regression

Regression with 5-fold cross-validation was performed between the different ECoG and audio signal modalities as well as with triggers values. The calculated regressions yielded numerous significant correlation values despite sometimes sub-optimal calculation. A complete list of regression results can be found in the supplementary materials (Table S1). Regression were not computed with audio signal given its high noise. Moreover, only decoding models were run.

Results are displayed in Figure 9 and in

Table 1 for the highest correlation values. Extended results can be found in supplementary materials, Table S1. A good correlation value was set to be above 0.4 and a decent correlation was considered to be above 0.2. Values below that limit are not shown in Table 1. The best regressions were obtained with theta envelope as an input with trigger values or audio envelope as outputs.

In passive listening, theta envelope was mildly correlated to the played audio envelope (mean over all subjects 0.34 ± 0.03, Z-value 11.33). Theta envelope was strongly correlated to trigger values in all subjects, with a mean over all subjects of 0.49 ± 0.02 (Z-value 24.50).

In sentence repetition, the correlation between theta envelope and audio envelope was not significant while the one with trigger values remained strongly significant in all subjects. The mean over all subjects was 0.51 ± 0.04 (Z-value 12.75).

The remaining results with a significant Z-value (>1.96) in Table 1 were most likely falsely significant by chance as they are single significant values that are not replicated in other subjects.

Table 1: All conditions with correlation values above 0.2 are displayed in the table below. All conditions with a correlation above 0.4 are highlighted in mild green. Significant correlation values are highlighted in dark green. They are only 2, the mean over all subjects for theta envelope with trigger values in passive listening and sentence repetition.

Condition	Input	Output	Subject	Regression		Standard deviation	NaN regresions	Z value
Passive listening	Theta envelope	Audio envelope	Mean	0.34	±	0.03		**11.33**
			subject 2	0.46	±	0.01	0	**46.00**
			subject 3	0.45	±	0.02	0	**22.50**
			subject 4	0.55	±	0.02	0	**27.50**
		Trigger values	Mean	0.49	±	0.02		**24.50**
			subject 1	0.50	±	0.01	0	**50.00**
			subject 2	0.71	±	0.03	0	**23.67**
			subject 3	0.30	±	0.01	0	**30.00**
			subject 4	0.46	±	0.02	0	**23.00**
Sentence repetition	Theta activity	Audio envelope	subject 3	-0.12	±	0.06	1	**2.00**
	Theta envelope	Audio envelope	Mean	0.05	±	0.02		**2.50**
			subject 1	0.46	±	0.02	0	**23.00**
			subject 4	-0.11	±	0.01	0	**11.00**
		Trigger values	Mean	0.51	±	0.04		**12.75**
			subject 1	0.80	±	0.01	0	**80.00**
			subject 2	0.30	±	0.02	0	**15.00**
			subject 3	0.47	±	0.07	0	**6.71**
			subject 4	0.46	±	0.05	0	**9.20**
Words 3 conditions	Theta activity	Audio envelope	subject 4	-0.11	±	0.01	0	**11.00**
	Theta envelope	Audio envelope	subject 2	0.12	±	0.03	0	**4.00**
			subject 4	0.08	±	0.02	0	**4.00**
		Trigger values	Mean	-0.16	±	0.05		**3.20**
			subject 1	0.37	±	0.03	0	**12.33**

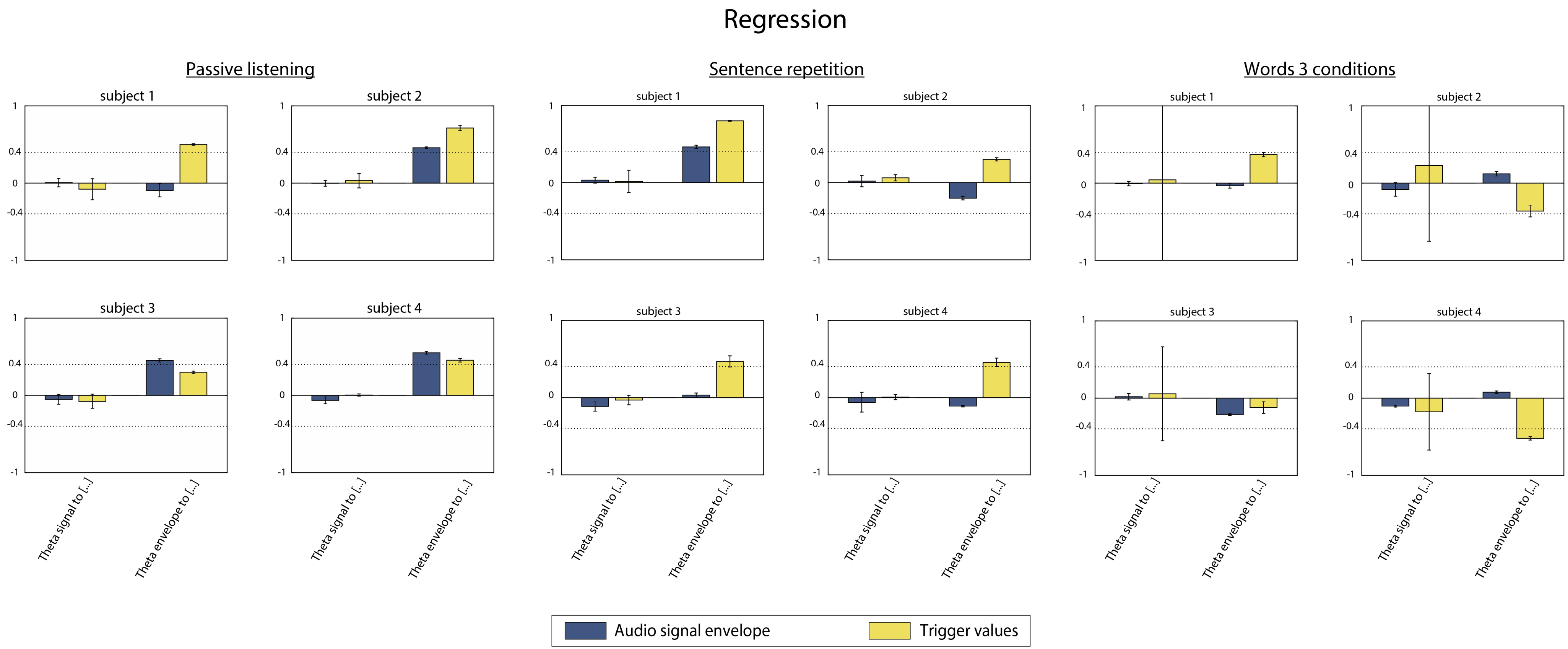

Figure 9: Regression values for all subject and condition investigated. Correlations above 0.4 or smaller than -0.4 were regarded as strong correlations. None of these correlations was significant with a significance level set at p = 0.05. However, for the theta envelope to trigger values correlation, the mean for all subjects was significant in passive listening and sentence repetition.

Discussion

Brain heat maps time lapse

As expected, theta neural time courses were different according to task and location. This was best analysed using brain heat maps displaying changes in theta activity and theta envelope over 3-dimensional brain models. Brain heat maps time lapses of theta activity showed that theta activity was markedly reduced during all tasks. However, at the beginning of the passive listening condition, a focal increase in the left auditory cortex could be seen in all subjects with left hemisphere recordings (Figure S1). This conserved change among subjects was expected given the necessary involvement of the primary cortex in listening tasks. Changes in the Wernicke area were faint, although a further decrease seemed apparent during passive listening and overt speech. This underlines the involvement of the Wernicke area in these tasks. As this decrease happens late in the overt speech period, we could hypothesise that the Wernicke area is activated in a sort of retrocontrol to check whether what the subject just said is in line with what he meant to say.

Theta activity in the frontal, motor and sensory cortices seemed to be rather stable during the whole three conditions. In imagined speech, there were no simultaneous changes in theta activity among all subjects, although a slight transient further decrease in theta activity in the superior temporal cortex, in a region compatible with the primary auditory cortex, might be seen half way through the imagined speech.

Finally, in overt speech, subject 2 shows two short-lived focal increases in theta activity, first in the lower temporal lobe and then at the superior temporal sulcus. In the other subjects, a light modulation around the lateral sulcus could be seen. One would have expected coherent changes among subjects in the Broca area, given its involvement in speech generation, and possibly in some parts of the motor cortex – tongue and lips areas for example. However, theta activity in these area remains stable, with the exception of subject 2, in whom the theta activity becomes smaller by the end of the overt speech condition. The electrode grid in subject 4 does sadly not encompass the parietal part of the motor cortex.

Theta envelope time lapses showed a similar pattern of generally spread decrease. This marked diminution was the rule over most brain areas and conditions. In passive listening, decreases were enhanced over time over the auditory cortices and the Wernicke area, in all subjects but subject 2, in whom focal increases temporally and frontally were seen. This tendency of decreasing activity in the auditory cortices and in the Wernicke areas can be found as well, to some degree, in imagined and overt speech, which might thus be a conserved feature in different language associated tasks. This might be linked to the role of theta waves in working memory and long-range synchronisation already observed in other brain areas (28, 29), two function necessary in all three conditions .

This analysis is limited by the time scale used for the different figures. With one heat map every second (Figure 6 and Figure 7) or even every 500ms (Figure S1 and S2), one could possibly miss fine short-lasting changes. A video analysis might cope with this issue. Another limitation is linked to the experimental set-up and the imagined speech condition. Indeed, despite giving a visual clue to the subject, the temporal experimental frame cannot be strictly control. In other words, it is unknown at which precise moment the subject starts its task and thinks of the given word. This could partially explain the lack of clear dynamics between all subjects in the imagined task. Moreover, the lack of measurable feedback (e.g. an audio output) in the imagined task prevents the verification of a conform execution. As such, side-tracking thoughts cannot be excluded either, although this problem is present in all conditions.

Analysis of brain maps comes with a few minor potential limitations. First, functional regions can only be assumed by comparison of the brain model with generic anatomic and functional brain maps. However, those regions are well defined. A second potential issue is the precision of the 3-dimensional brain model generated and of the projection of the activity over it. This model was validated by the Schalk lab after creation and is precise enough for a visual analysis.

Average brain heat maps

Brain heat maps displaying average theta envelope over whole conditions is one mean to infer theta envelope's reaction to these tasks. As mentioned earlier, only theta envelope brain heat maps made sense as theta activity oscillates around zero. Thus, its average is not very informative. These heat maps differ in general aspects from the time lapse heat maps. Amplitude of changes seen in these heat maps is much smaller. This was to be expected as average values are necessarily smaller than punctual values can be.

Theta envelope in passive listening was inconsistent among subjects, with subject 1 showing a marked decrease around the lateral sulcus while subject 2 and 4 exhibited an increase. Further discording results were seen in the frontal cortex, were subject 4 had a clear decrease in theta envelope while subject 1 showed a convincing increase. Thus, with subject 3 being the only subject with right hemisphere recordings and all left hemisphere subjects showing disparate changes, it is difficult to make definitive conclusions from theta envelope heat maps in passive listening. Surprisingly, no consistent theta envelope changes took place in the auditory cortex and Wernicke area, while it was the only clearly shared increase of activity in the brain heat map time lapse (Figure 7). In general, these grand average heat maps show several areas of average increase in theta envelope, which is surprising given the strong and almost uniform decrease seen otherwise in theta envelope time lapses. As mentioned above, one explanation could be that time lapses, by chance, captured only time points when theta envelope was decreased. However, even with a figure for every 100ms (not shown), the general decrease was still consistent in the time lapses. Another possibility is that the slight changes in the task protocol is behind the observed differences in theta envelope. Passive listening is almost similar, but doesn't involve any memorisation task in the passive listening of the simple condition (Figure 8), while it does in the triple condition as the subject will have to think of this word and repeat it afterwards. That might be a possible explanation for the differences in frontal activity for example.

In sentence repetition, a consistent increase in theta envelope was seen focally around the temporal sulcus in an area compatible with the auditory cortex in all 4 subjects. In this case, theta envelope in this area might be seen as a more reliable marker. This local increase might be linked to working memory processes, like theta waves have been linked to in the hippocampus (3, 30). This could be linked to short term storage activity of the heard word to allow its repetition. Conclusions can't be made from theta envelope changes in other areas as these are inconsistent between subject. We might however infer that theta envelope changes in other brain areas aren't directly linked to the repetition task.

Finally, when comparing theta envelop in passive listening and in sentence repetition, one sees that auditory cortex increases are more consistent in sentence repetition, were all subject displayed increased theta envelope in the auditory cortex. In the Wernicke area, theta envelope is inconsistent between subjects in both conditions. Despite focal changes in theta envelope – mostly in auditory cortex and Wernicke's area – between the two conditions in all three subjects with recordings from the left hemisphere, theta envelope brain-wide does not appear to change markedly.

These means over a whole condition are informative about the overall theta envelope activity, but do not reflect an observed activity at any given time point. They can thus not be used as marker correlated to tasks onset and offset. That is why regressions have been computed as well and are discussed in the next section.

Regression

Regression values were strong between theta envelope and trigger values in passive listening and sentence repetition. All subjects had a quite robust correlation (all above 0.3, means around 0.5!). This is rather surprising given the brain heat maps observed in Figure 8, where no clear consensus in theta envelope seemed to emerge between subjects.

In passive listening, theta envelope correlated well to audio envelope in 3 out of 4 subjects (mean 0.34 +/- 0.03, Z-value 11.33) and trigger values in all subjects (mean 0.49 +/- 0.02, Z-value 24.50). This is in line with prior findings that theta envelope is tracking speech envelope (7, 26). Moreover, this shows that trigger values represent a decent approximation of the presence of speech. Interestingly, it seems that theta envelope from the right hemisphere is well correlated to audio envelope (0.45 +/- 0.02, Z-value 22.50) and triggers as well (0.30 +/- 0.01, Z-value 30.00).

In sentence repetition, the best correlation was seen between theta envelope and trigger values, with a mean correlation of 0.51 (+/- 0.04, Z-value 12.75). This strong correlation, together with the robust correlation seen in passive listening, shows that theta envelope could be a marker of speech related to both listening and repeating. However, a good correlation between theta envelope and trigger values is not very surprising. Indeed, average theta envelope (Figure 8) is relatively similar between the two conditions. As trigger values hide the main difference between the two conditions – namely the audio output –, one could expect similar correlation results. Interestingly, data from subject 3, who had electrodes on the right hemisphere, yielded a robust correlation again. This is interesting in the light of the concept that speech brain areas are not simply concentrated in the left hemisphere. However, no conclusions can be drawn from a single subject observation. However, while theta envelope is not symmetric, its activity in both hemisphere is well correlated to listening and repeating. Theta envelope might thus be a marker of speech related activities. However, to validate this hypothesis, a strong correlation is still required in imagined and overt speech, which were investigated in the triple condition. Of note, while theta envelope was well correlated to audio envelope in passive listening, this was not the case in sentence repetition. Indeed, only one subject had a significant correlation. Moreover, subject 4 had a negative correlation value, indicating that the significance of these correlations between theta and audio envelope is doubtful at best. One possibility for these changes in correlation would be that theta envelope might be linked to listening activities in both conditions but less to the phase of speech production present in the sentence repetition task.

In the triple condition, the best correlation values were as well with theta envelope as an input. Only few were significant, and only one was reasonably robust (subject 1, 0.37 ± 0.03, Z-value 12.33). This was expected given the large-scale decrease in theta envelope seen on brain heat map time lapses (Figure 7 and supplementary Figure S2) and the rather faint changes in theta envelope between the different conditions. Thus, the potential value of theta envelope as a marker for general speech related activities seems to be rather low.

These poorer correlations in triple condition must be linked either to the activity in imagined speech or in overt speech as the first listening period is close to the paradigm in passive listening, task in which correlations were good. It is possible that results in the triple condition are worse due to looser time-locking of the experimental set-up in imagined speech. This explanation would however not hold for overt speech, were the audio output can be verified. Correlation calculation with data covering only the distinct parts of the triple condition might give us further insight. Although only electrodes with significant changes in theta envelope were taken into account for regression calculation, another possibility to increase the correlation of theta envelope with audio output could be to select only a restricted number of electrodes covering specific brain areas (e.g. auditory cortex, Wernicke's area). This would diminish "noise" coming from electrode with unspecific activity. The other possibility is to calculate the regression for each electrode separately, to get a mapping of the areas with the best correlation to the output. Yet another potential improvement of the correlation could be obtained by fine-tuning its calculation. First, correlation accuracy was limited by numerous uncomplete calculations due to non-number elements. Despite debugging efforts, some correlations could still not be calculated (Table S1). Calculation problems consistently involved audio signal or theta activity. As no clear cause could be established so far, improving these calculation is a first step in enhancing regression results. Second, different time-windows could be tested. Only a 500ms time window was used in this project. Although it should give enough time to see changes in a 4-8Hz signal, other time-windows might improve the correlation. Finally, both audio and ECoG signals and envelopes were downsampled to 100Hz to reduce the calculation load. While 100Hz is a sampling rate that should not alter speech nor theta signal and envelope (as 8Hz is largely smaller than a 100Hz), a higher sampling rate could potentially refine the results. This last suggestion would most likely not change the results significantly. Finally, a better marker than trigger values could be found. These indicate crudely the time frame in which the subject acts. Triggers based on audio envelope for example might improve the results, at least in listening and overt speech. Besides improving electrode selection and fine-tuning regression calculation, a last possibility to improve the link between theta waves and speech might be to use a non-linear regression or other more complex machine learning algorithms.

A potential general limitation is that these results were obtained from data of patients suffering from epilepsy. While biased results due to epilepsy are improbable due to strict selection of electrodes that had no epileptic activity on inspection. Moreover, some sentence repetition trials of subject 1 were discarded due to epileptic contamination. Some remaining artefacts interfering with the experiments can't be completely excluded. It is however unlikely that such artefacts would have an impact on the outcomes measured as averages over more than a hundred trials were computed

Overall, these results show a restricted involvement of theta waves in language. It seems that theta activity and mostly theta envelope is an important feature of brain activity in listening related task, as previously shown by Poeppel et al. (7, 26). The importance of theta implication in imagined and overt speech remains uncertain. The use of either theta activity or theta envelope as a marker of speech is at this stage unlikely.

Perspectives

Some hints towards further investigation of theta waves in speech have been given throughout the discussion. An important next step would be improving the regression calculation and refining its spatial. This will represent an important step to understand further the implication of theta activity and envelope in different brain areas during language related tasks. As some changes in theta envelope were shown in this study in the auditory cortex and the Wernicke area in all conditions, refining the regression might also increase the chances of using theta activity as a marker of speech onset and offset. Another possibility to improve decoding from theta activity would be to train the patients while giving them feedback. Although not ideally straight forward for a clinical application, this might increase the yield of relevant features extracted from theta activity. Finally, brain language areas are spread around the brain. Albeit ethically and clinically questionable, recordings from the whole cortex at once might give further insights. In the same line of idea, being able to record from a larger pool of subjects might as well be beneficial. Indeed, some rewiring in epileptic patients might lead to results failing to generalise completely in another population. These last two propositions might become easier to put into action with recording technologies improvements. Indeed, it would be much easier and safer if a non-invasive technique like EEG gains the signal-to-noise power to conduct this kind of experiments.

Furthermore, it would be interesting to extend the exploration of the role of theta waves to other aspects of language. Reading or writing for example were not tested here and must rely on partially overlapping brain areas and activity.

Finally, other frequency bands should be further studied. Gamma and theta bands are probably the most studied so far. It is well possible that frequencies in-between (alpha, beta) or below (delta) convey important features for speech decoding.

To conclude, this study showed that theta activity is increased in the left primary cortex during passive listening tasks and confirmed that theta envelope is well correlated to the audio envelope in listening tasks. However, this study couldn’t establish speech onset and offset markers from theta activity nor theta envelope. We suggest however further experiments that might lead to such a biomarker.

References

1. Shih JJ, Krusienski DJ, Wolpaw JR. Brain-Computer Interfaces in Medicine. Mayo Clinic Proceedings. 2012;87(3):268-79.
2. Smith E, Delargy M. Locked-in syndrome. BMJ (Clinical research ed). 2005;330(7488):406-9.
3. Buzsáki G. Rhythms of the Brain. 198 Madison Avenue, New York, New York 10016: Oxford University Press, Inc.; 2006.
4. Martin S, Brunner P, Holdgraf C, Heinze H-JJ, Crone NE, Rieger J, et al. Decoding spectrotemporal features of overt and covert speech from the human cortex. Frontiers in neuroengineering. 2014;7:14.
5. Martin S, Brunner P, Iturrate I, Millán JdR, Schalk G, Knight RT, et al. Word pair classification during imagined speech using direct brain recordings. Scientific Reports. 2016;6:25803.
6. Luo H, Poeppel D. Phase Patterns of Neuronal Responses Reliably Discriminate Speech in Human Auditory Cortex. Neuron. 2007;54(6):1001-10.
7. Zion Golumbic EM, Ding N, Bickel S, Lakatos P, Schevon CA, McKhann GM, et al. Mechanisms Underlying Selective Neuronal Tracking of Attended Speech at a "Cocktail Party". Neuron. 2013;77(5):980-91.
8. Bolhuis JJ, Tattersall I, Chomsky N, Berwick RC. PLoS Biology. 2014;12(8):1-6.
9. Wolpaw JR, Birbaumer N, McFarland DJ, Pfurtscheller G, Vaughan TM. Brain-computer interfaces for communication and control. Clinical Neurophysiology. 2002;113(6):767-91.
10. Pandarinath C, Nuyujukian P, Blable CH, Sorice BL, Saab J, Willet FR, et al. High performance communication by people with paralysis using an intracortical brain-computer interface. eLIFE. 2017.
11. Tucker MR, Olivier J, Pagel A, Bleuler H, Bouri M, Lambercy O, et al. Control strategies for active lower extremity prosthetics and orthotics: a review. Journal of NeuroEngineering and Rehabilitation. 2015;12(1):1.
12. Guenther FH, Brumberg JS, Wright EJ, Nieto-Castanon A, Tourville JA, Panko M, et al. A Wireless Brain-Machine Interface for Real-Time Speech Synthesis. PLOS ONE. 2009;4(12):e8218.
13. George J. Augustine PD, Dona M. Chikaraishi PD, Michael D. Ehlers MD, Ph.D.,, Gillian Einstein PD, David Fitzpatrick PD, William C. Hall PD, et al. Language and Speech. In: DALE PURVES, GEORGE J. AUGUSTINE, DAVID FITZPATRICK, WILLIAM C. HALL, ANTHONY-SAMUEL LAMANTIA, JAMES O. MCNAMARA, et al., editors. Neuroscience. third edition ed. Sunderland, Massachusetts U.S.A.: Sinauer Associates, Inc.; 2004. p. 637 - 58.
14. Price CJ. A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. NeuroImage. 2012;62(2):816-47.
15. Brigham K, Kumar BVK, editors. Imagined speech classification with EEG signals for silent communication: a preliminary investigation into synthetic telepathy. Imagined speech classification with EEG signals for silent communication: a preliminary investigation into synthetic telepathy; 2010: IEEE.
16. Roth M, Decety J, Raybaudi M, Massarelli R, Delon-Martin C, Segebarth C, et al. Possible involvement of primary motor cortex in mentally simulated movement: a functional magnetic resonance imaging study. Neuroreport. 1996;7(7):1280-4.
17. 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;54(4):2960-72.
18. Geva S, Correia M, Warburton EA. Diffusion tensor imaging in the study of language and aphasia. Aphasiology. 2011;25(5):543-58.
19. Shain W, Spataro L, Dilgen J, Haverstick K, Retterer S, Isaacson M, et al. Controlling cellular reactive responses around neural prosthetic devices using peripheral and local intervention strategies. IEEE Transactions on Neural Systems and Rehabilitation Engineering. 2003;11(2):186-8.

20. Leuthardt EC, Gaona C, Sharma M, Szrama N, Roland J, Freudenberg Z, et al. Using the electrocorticographic speech network to control a brain–computer interface in humans. Journal of Neural Engineering. 2011;8(3):36004.
21. Derix J, Iljina O, Weiske J, Schulze-Bonhage A, Aertsen A, Ball T. From speech to thought: the neuronal basis of cognitive units in non-experimental, real-life communication investigated using ECoG. Frontiers in Human Neuroscience. 2014;8(383).
22. Christian Herff GJ, Lorenz Diener, Jerry Shih, Dean Krusienski, Tanja Schultz. Towards direct speech synthesis from ECoG: A pilot study. 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2016. 2016-October: Institute of Electrical and Electronics Engineers Inc.; 2016. p. 1540--3.
23. Hermes D, Miller KJ, Vansteensel MJ, Edwards E, Ferrier CH, Bleichner MG, et al. Cortical theta wanes for language. NeuroImage. 2014;85:738-48.
24. Gross J, Hoogenboom N, Thut G, Schyns P, Panzeri S, Belin P, et al. Speech Rhythms and Multiplexed Oscillatory Sensory Coding in the Human Brain. PLoS Biology. 2013;11(12).
25. Martin S. Understanding and Decoding Imagined Speech using Electrocorticographic Recordings in Humans [Programme Doctoral en Neurosciences]. Lausanne: Ecole Polytechnique Fédérale; 2017.
26. Giraud A-LL, Poeppel D. Cortical oscillations and speech processing: emerging computational principles and operations. Nature neuroscience. 2012;15(4):511-7.
27. Pasley BN, David SV, Mesgarani N, Flinker A, Shamma SA, Crone NE, et al. Reconstructing Speech from Human Auditory Cortex. PLOS Biology. 2012;10(1):e1001251.
28. Long LL, Bunce JG, Chrobak JJ. Theta variation and spatiotemporal scaling along the septotemporal axis of the hippocampus. Frontiers in Systems Neuroscience. 2015;9:37.
29. Daume J, Gruber T, Engel AK, Friese U. Phase-Amplitude Coupling and Long-Range Phase Synchronization Reveal Frontotemporal Interactions during Visual Working Memory. The Journal of Neuroscience. 2016;37(2):313-22.
30. Vertes RP. Hippocampal theta rhythm: A tag for short-term memory. Hippocampus. 2005;15(7):923-35.

Supplementary materials

Detailed brain heat map time lapses and complete regression results can be found below.

Figure S1: Brain heat map time lapse for theta activity. One heat map represents the average activity over the 50ms before and 50ms after it is displayed. One heat map is shown every 500ms.

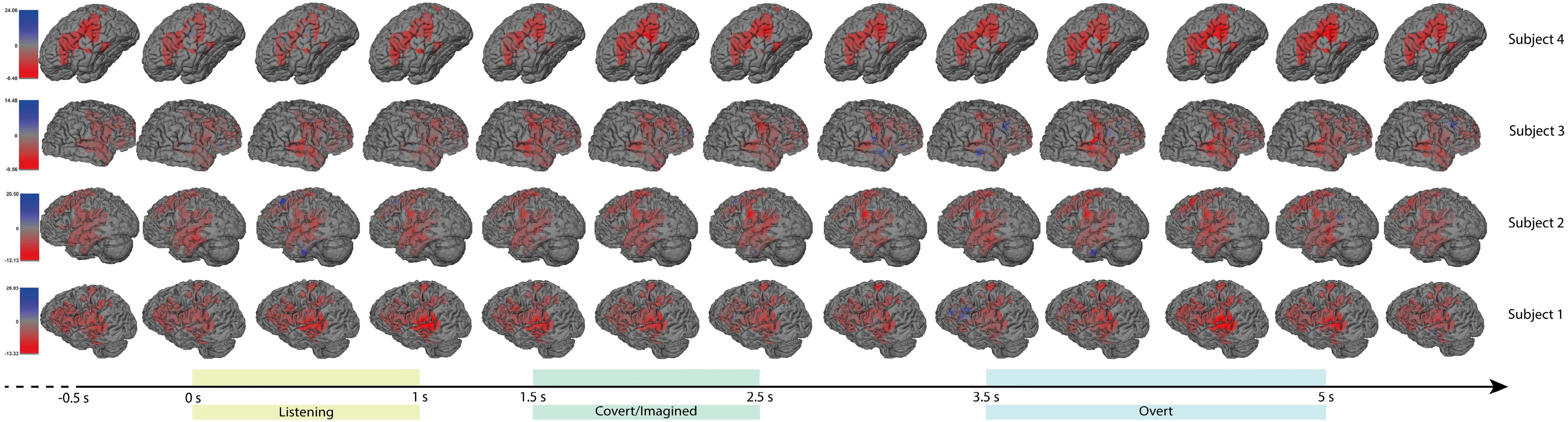

Figure S2: Brain heat map time lapse for theta envelope. One heat map represents the average envelope over the 50ms before and 50ms after it is displayed. One heat map is shown every 500ms.

Table S1: Extended regression results for passive listening, sentence repetition and words 3 conditions. Regressions with an absolute correlation value above 0.2 or indicated with light green or blue. Strong correlation values (above 0.4) are highlighted in darker green or blue. The number of unsuccessful cross-validations is given by the "NaN regressions" column. Finally, the Z transformation was used to assess the significance of the different correlations. Values over 1.96 are estimated to be significant and highlighted in green.

Condition	Input	Output	Subject	Regression		Standard deviation	NaN regresions	Z value
Passive listening	Theta activity	Audio envelope	Mean	-0.03	±	0.05		0.6
			subject 1	0.01	±	0.05	0	0.2
			subject 2	0.00	±	0.04	1	0
			subject 3	-0.05	±	0.06	0	0.8333
			subject 4	-0.06	±	0.04	0	1.5
		Audio signal	Mean	0.04	±	0.01		4
			subject 1	0.00	±	0	4	x
			subject 2	0.00	±	0.02	1	0
			subject 3	NaN	±	NaN	5	x
			subject 4	0.12	±	0.01	3	12
		Trigger values	Mean	-0.03	±	0.08		0.375
			subject 1	-0.08	±	0.14	0	0.5714
			subject 2	0.03	±	0.09	1	0.3333
			subject 3	-0.08	±	0.09	2	0.8889
			subject 4	0.00	±	0.01	0	0
	Theta envelope	Audio envelope	Mean	0.34	±	0.03		11.333
			subject 1	-0.09	±	0.08	0	1.125
			subject 2	0.46	±	0.01	0	46
			subject 3	0.45	±	0.02	0	22.5
			subject 4	0.55	±	0.02	0	27.5
		Audio signal	Mean	0.00	±	0.03		0
			subject 1	0.04	±	0.06	1	0.6667
			subject 2	-0.01	±	0.01	3	1
			subject 3	-0.08	±	0.04	0	2
			subject 4	0.04	±	0	4	x
		Trigger values	**Mean**	**0.49**	**±**	**0.02**		24.5
			subject 1	0.50	±	0.01	0	50
			subject 2	0.71	±	0.03	0	23.667
			subject 3	0.30	±	0.01	0	30
			subject 4	0.46	±	0.02	0	23

Condition	Input	Output	Subject	Regression		Standard deviation	NaN regresions	Z value
Sentence repetition	Theta activity	Audio envelope	Mean	-0.03	±	0.07		0.42857
			subject 1	0.03	±	0.04	0	0.75
			subject 2	0.02	±	0.07	1	0.28571
			subject 3	-0.12	±	0.06	1	2
			subject 4	-0.06	±	0.13	0	0.46154
		Audio signal	Mean	-0.03	±	0.01		3
			subject 1	NaN	±	NaN	5	x
			subject 2	0.01	±	0.02	1	0.5
			subject 3	-0.10	±	0	4	x
			subject 4	-0.01	±	0.02	3	0.5
		Trigger values	Mean	0.01	±	0.07		0.14286
			subject 1	0.02	±	0.14	0	0.14286
			subject 2	0.06	±	0.04	0	1.5
			subject 3	-0.03	±	0.06	0	0.5
			subject 4	0.01	±	0.03	1	0.33333
	Theta envelope	Audio envelope	Mean	0.05	±	0.02		2.5
			subject 1	0.46	±	0.02	0	23
			subject 2	-0.20	±	0.02	0	10
			subject 3	0.03	±	0.03	0	1
			subject 4	-0.11	±	0.01	0	11
		Audio signal	Mean	0.02	±	0.04		0.5
			subject 1	NaN	±	NaN	5	x
			subject 2	0.01	±	0.04	1	0.25
			subject 3	0.00	±	0	3	x
			subject 4	0.03	±	0.07	0	0.42857
		Trigger values	**Mean**	**0.51**	**±**	**0.04**		12.75
			subject 1	0.80	±	0.01	0	80
			subject 2	0.30	±	0.02	0	15
			subject 3	0.47	±	0.07	0	6.71429
			subject 4	0.46	±	0.05	0	9.2

Condition	Input	Output	Subject	Regression		Standard deviation	NaN regresions	Z value
Words 3 conditions	Theta activity	Audio envelope	Mean	-0.04	±	0.04		1
			subject 1	-0.01	±	0.03	0	0.33333
			subject 2	-0.08	±	0.09	2	0.88889
			subject 3	0.02	±	0.04	0	0.5
			subject 4	-0.11	±	0.01	0	11
		Audio signal	Mean	0.00	±	0.03		0
			subject 1	0.01	±	0.01	0	1
			subject 2	-0.01	±	0	4	x
			subject 3	-0.04	±	0.07	2	0.57143
			subject 4	0.02	±	0.04	2	0.5
		Trigger values	Mean	0.04	±	0.83		0.04819
			subject 1	0.04	±	1.26	3	0.03175
			subject 2	0.23	±	0.98	3	0.23469
			subject 3	0.06	±	0.61	1	0.09836
			subject 4	-0.18	±	0.5	3	0.36
	Theta envelope	Audio envelope	Mean	-0.01	±	0.02		0.5
			subject 1	-0.03	±	0.03	0	1
			subject 2	0.12	±	0.03	0	4
			subject 3	-0.21	±	0.01	0	21
			subject 4	0.08	±	0.02	0	4
		Audio signal	Mean	0.02	±	0		x
			subject 1	0.00	±	0.01	2	0
			subject 2	0.02	±	0	4	x
			subject 3	0.04	±	0	4	x
			subject 4	NaN	±	NaN	5	x
		Trigger values	Mean	-0.16	±	0.05		3.2
			subject 1	0.37	±	0.03	0	12.3333
			subject 2	-0.36	±	0.08	0	4.5
			subject 3	-0.12	±	0.08	0	1.5
			subject 4	-0.52	±	0.02	0	26