MVPA analysis of intertrial phase coherence of neuromagnetic responses to words reliably classifies multiple levels of language processing in the brain

https://doi.org/10.1523/ENEURO.0444-18.2019

Cite as: eNeuro 2019; 10.1523/ENEURO.0444-18.2019
Received: 7 November 2018
Revised: 30 May 2019
Accepted: 28 June 2019

This Early Release article has been peer-reviewed and accepted, but has not been through the composition and copyediting processes. The final version may differ slightly in style or formatting and will contain links to any extended data.

Alerts: Sign up at www.eneuro.org/alerts to receive customized email alerts when the fully formatted version of this article is published.

Copyright © 2019 Jensen et al.
This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license, which permits unrestricted use, distribution and reproduction in any medium provided that the original work is properly attributed.
MVPA analysis of intertrial phase coherence of neuromagnetic responses to words reliably classifies multiple levels of language processing in the brain.

MVPA analysis of language processing in the brain

Mads Jensen¹, Rasha Hyder¹, Yury Shtyrov¹²
¹: Center of Functionally Integrative Neuroscience (CFIN), Department of Clinical Medicine, Aarhus University, Aarhus, Denmark
²: Laboratory of Behavioural Neurodynamics, St. Petersburg State University, St. Petersburg, Russia

RH and YS designed research, RH performed research, MJ analysed data, MJ and YS wrote paper.

Correspondence should be addressed to Mads Jensen
Center of Functionally Integrative Neuroscience (CFIN), Nørrebrogade 44, build. 10G, 4th DK-8000 Aarhus, Denmark
Email: mads@cfin.au.dk

6. Number of figures: 2
7. Number of tables: 4
9. Number of words for Abstract: 246
10. Number of words for Significance Statement: 118
11. Number of words for Introduction: 1793
12. Number of words for Discussion: 1754

Acknowledgements
We thank Andreas Højlund, Karen Østergaard, Christelle Gansonre, Alina Leminen, Eino Partanen, Nikola Vukovic, Lilli Kimppa, Christopher Bailey, and Britta Westner for their help at different stages of this work

Conflict of Interest: Authors report no conflict of interest.

Funding sources: Danish Council for Independent Research (DFF 6110-00486), Lundbeck Foundation (R164-2013-15801; R140-2013-12951), Aarhus University, Institute for Clinical Medicine, Central Jutland Region, MINDLab infrastructure, (14.W03.31.0010)
Abstract

Neural processing of language is still among the most poorly understood functions of the human brain, whereas a need to objectively assess the neurocognitive status of the language function in a participant-friendly and noninvasive fashion arises in various situations. Here, we propose a solution for this based on a short task-free recording of MEG responses to a set of spoken linguistic contrasts. We used spoken stimuli that diverged lexically (words/pseudowords), semantically (action-related/abstract) or morphosyntactically (grammatically correct/ungrammatical). Based on beamformer source reconstruction we investigated inter-trial phase coherence (ITPC) in five canonical bands (alpha, beta, and low, medium and high gamma) using multivariate pattern analysis (MVPA). Using this approach, we could successfully classify brain responses to meaningful words from meaningless pseudowords, correct from incorrect syntax, as well as semantic differences. The best classification results indicated distributed patterns of activity dominated by core temporofrontal language circuits and complemented by other areas. They varied between the different neurolinguistic properties across frequency bands, with lexical processes classified predominantly by broad gamma, semantic distinctions – by alpha and beta, and syntax – by low gamma feature patterns. Crucially, all types of processing commenced in a near-parallel fashion from ~100 ms after the auditory information allowed for disambiguating the spoken input. This shows that individual neurolinguistic properties take place simultaneously and involve overlapping yet distinct neuronal networks that operate at different frequency bands. This brings further hope that brain imaging can be used to assess neurolinguistic processes objectively and noninvasively in a range of populations.

Significance Statement

In an MEG study that was optimally designed to test several language features in a non-attentive paradigm, we found that, by analysing cortical source-level inter-trial phase
coherence (ITPC) in five canonical bands (alpha, beta, and low, medium and high gamma) with machine-learning classification tools (multivariate pattern analysis, MVPA), we could successfully classify meaningful words from meaningless pseudowords, correct from incorrect syntax, and semantic differences between words, based on passive brain responses recorded in a task-free fashion. The results show different time courses for the different processes that involve different frequency bands. It is to our knowledge the first study to simultaneously map and objectively classify multiple neurolinguistics processes in a comparable manner across language features and frequency bands.

**Introduction**

A neuropsychological assessment of the neurological and/or cognitive status of a subject or patient is a routine required in a variety of situations. This typically involves elaborate behavioural tests aimed, for instance, at evaluating the extent of developmental disorders, assessing the level of neurodegenerative impairment, neurological damage after a head injury, screening for hearing loss, etc. In order to perform such tests successfully, the subject, on the one hand, must typically have a reasonably clear understanding of what they have to do in a particular procedure; on the other hand, they must also be able to communicate their responses to a given task, by giving e.g. a manual, facial or oral response. This raises the problem of uncooperative subjects, such as those suffering from neurological/organic brain disorders or mental illnesses that cause patients to become unresponsive. A brain-damaged person may be unable to respond verbally because of a collateral lesion-related injury (such as akinetic mutism, paralysis, aphasia). A young child with a developmental disorder may be unable or unwilling to communicate their reaction overtly. A locked-in patient, although not
necessarily unconscious, is unable to respond due to the total loss of their motor control function (Laureys, Owen, & Schiff, 2004; Ragazzoni, Grippo, Tozzi, & Zaccara, 2000). On the other hand, an undetected language impairment could have drastic consequences for (mis)diagnosis of language-unrelated disturbances (Majerus, Bruno, Schnakers, Giacino, & Laureys, 2009). A range of such situations is wide and they create a substantial challenge for diagnosis and assessment of performance, development or recovery in various groups. Clearly, techniques that could reveal the brain processing of language without relying on the individual’s overt behaviour would be helpful in such cases.

Fast-paced development of non-invasive neuroimaging techniques in recent years has given hopes of assessing the brain status of various cognitive functions objectively, even when overt response may not be possible. Language is a complex phenomenon that, to make a coherent whole, requires multiple information types involving different specific features and properties that rapidly unfold in time with a fine temporal structure and remarkable speed (Friederici, 2002, 2011; Shtyrov, 2010). This, in turn, suggests that for the capturing of cerebral processing of linguistic information, a temporally-resolved neuroimaging method is needed that could faithfully track the dynamic neural activity during language comprehension (e.g. Hagoort, 2008). At present, two main techniques are able to provide such high temporal resolution: magneto- and electroencephalography (MEG, EEG), both capable of registering mass neural activity on a millisecond scale. Whereas the two methods are highly similar, MEG has a certain advantage in the ease of modelling underlying cortical activity, owing to unhindered spreading of magnetic fields (but not electric currents) through the head and skull tissues (Baillet, Mosher, & Leahy, 2001).

A body of electrophysiological research done in EEG and MEG using various linguistic
materials has provided a rich picture of linguistic processing in the brain. Perhaps the most
well-known neurophysiological response to language stimuli is the so-called N400 (Kutas &
Hillyard, 1980), usually seen as an index of (lexico-)semantic processes and peaking at ~400
ms. Syntactic (grammatical) processing has been associated with an early ELAN (early left
anterior negativity) response reflecting the stimulus’ grammaticality from ~100 ms
(Friederici, Pfeifer, & Hahne, 1993; Neville, Nicol, Barss, Forster, & Garrett, 1991), as well
as later frontal negativities (LAN) with longer latencies (Gunter, Friederici, & Schriefers,
2000; Münte, Heinze, Matzke, Wieringa, & Johannes, 1998) and the P600, a late positive
shift (Osterhout, Holcomb, & Swinney, 1994).

Most of these responses (except, to an extent, ELAN) still require at least a degree of overt
attention or even focused task performance from the subject. In terms of assessing linguistic
processing in a more task-free fashion, a number of studies have attempted to use the so-
called passive paradigms, in which the subjects are presented with linguistic contrasts without
having to perform an overt task and are usually distracted from the auditory input by a video
or another unrelated activity (Näätänen, Paavilainen, Rinne, & Alho, 2007; Pulvermüller &
Shtyrov, 2006). A series of studies using this approach established MEG/EEG correlates of
automatic linguistic access including phonological, lexical, semantic and syntactic levels of
information processing (Shtyrov, 2010). These have shown that information-specific
linguistic activations can be recorded noninvasively without an explicit task or even an
instruction to focus on the speech input. For instance, meaningful native words show stronger
activation in the core language system than meaningless acoustically similar pseudowords,
which putatively indicates activation of word-specific memory traces (Pulvermüller et al.,
2001). Furthermore, these activations show both semantic specificity (e.g. by involving motor
cortex activation and deactivation when presenting action-related words (Hauk, Shtyrov, &
Pulvermüller, 2008; Moseley, Pulvermüller, & Shtyrov, 2013; Shtyrov, Butorina, Nikolaeva, & Stroganova, 2014) and sensitivity to grammatical properties (ELAN-like activity in response to syntactic violations (Hasting, Kotz, & Friederici, 2007).

Such studies of neurolinguistic processing typically focus on one single process (e.g. syntax or semantics) at a time. For this approach to be more practically-oriented, it would seem essential to develop a task-free paradigm that can assess multiple types of linguistic information processing in a single short participant-friendly session. Furthermore, previous studies have mostly investigated neurolinguistic processes using ERP/ERFs - that is, evoked activity in a rather narrow frequency band (typically 1-30 Hz), - and it would thus appear advantageous to open up the frequency spectrum to maximise possibilities for registering brain reflections of language. Oscillations in different frequency bands have been shown to reflect multiple cognitive processes and are considered to be a vehicle for neural communication (Fries, 2005; Varela, Lachaux, Rodriguez, & Martinerie, 2001). Different frequency bands have been ascribed different functional roles in the neocortical information processing (Cannon et al., 2013; Sauseng & Klimesch, 2008), where higher bands could be related to local computations (Buzsáki & Schomburg, 2015; Buzsáki & Silva, 2012) and lower bands to longer-range connectivity (Von Stein & Sarnthein, 2000). That different neurolinguistic processes involve different frequency bands have been shown in a range of studies. For instance, Bastiaansen & Hagoort (2006) argued that retrieval of lexical information and unification of semantic and syntactic information are underpinned by different oscillatory networks, while Kösem and van Wassenhove (2017) suggested that θ oscillations (3-8 Hz) are related to acoustic properties. Teng, Tian, Rowland, & Poeppel (2017) shows that humans can track spoken language dynamics at both θ and γ oscillations. Using electrocorticography (ECoG), Towle and colleagues (2008) have shown an increase in
high $\gamma$ (70-100 Hz) in relation to hearing a word compared to a tone. Luo and Poeppel (2007) showed that phase information in the oscillatory dynamics can track and discriminate spoken sentences. However, although the interest in time-frequency analysis of electrophysiological data has been steadily rising in recent years, language-related oscillatory dynamics still remains relatively unexplored, with most studies in the field focussing on evoke potentials/fields instead.

In an attempt to close the gap between these research strands and at the same time bridge this research with applied/clinical needs, we set out to combine these different approaches and designed a simple short paradigm that simultaneously includes lexical, semantic and syntactic contrasts, in order to assess different levels of linguistic processing in a single session. This paradigm is tested here by recording MEG responses in a sample of healthy adult participants as a first step to establishing its applicability. The participants were presented with spoken stimuli, which were either meaningless pseudowords or meaningful words with different semantics (action-related verb vs. concrete visual noun) and that either could be syntactically correct or included a morphosyntactic stem-affix violation. These were presented without any stimulus-related task, while the subjects’ attention was diverted away from the auditory stimulus stream.

Furthermore, in order not to restrict our results to narrow-band evoked activity, we analysed activation in a wide range of frequency bands, from $\alpha$ to high $\gamma$. We focussed on the phase part of the oscillatory activity, as phase synchrony has been theorised to be directly related to neural computations (Fries, 2005; Lopes da Silva, 2006; Palva, Palva, & Kaila, 2005; Varela et al., 2001). In order to quantify coordinated neural phase activity related to neurolinguistic processing, we determined phase synchrony by analysing Inter-Trial Phase Coherence (ITPC).
of the MEG responses. ITPC is a measure of how aligned the phase angles over individual trials are. It therefore provides information about mass-synchronised neural activation (and in this sense is to a degree similar to ERP/ERF) without restricting it to a specific power peak or band. ITPC can inform us about the phase synchrony over trials: i.e., if ITPC is high, it means that the phases become aligned when performing a particular perceptual or cognitive computation. We investigated the ITPC in different frequency bands allowing for a comprehensive assessment of the neurolinguistic dynamics. To gain anatomic specificity, single-subject MRIs were used to model cortical source activity, and ITPC values were calculated in individual source space.

Given the vast amount of information about neural activity created by analysing the ITPC across both time and frequency bands, we used Multivariate Pattern Analysis (MVPA) for an unbiased statistical assessment of the data (Bzdok, Altman, & Krzywinski, 2018), where by ‘unbiased’ we mean that the researcher does not have to choose when and where to test for differences. MVPA (see Methods for more detail) is a machine learning technique based on predicting data rather than on parameter estimation (as used in traditional factorial analyses, e.g. ANOVA or t-tests). The MVPA algorithm will first try to extract a pattern from a subset of the data, which can then be used to predict new, previously unseen data. This allows for a data-driven analysis without a priori hypotheses about spatial or temporal location of the signal of interest. For instance, we might allow the decoding algorithm to train itself on MEG recording trials of meaningful vs. meaningless word form stimuli (i.e. find specific patterns of features in the data most valuable for detecting word-pseudoword differences) and then ask whether it can correctly predict the same distinction in a different data subset; this can be done both within and across subjects. Crucially, such an algorithm, if successful, may in principle be used to estimate data predictability in one subject (e.g.
patient) from another one (e.g. healthy norm) and thus determine both normal functionality
and functional abnormalities, or assess the same or different states of a particular processing
system in the same individual.

In sum, we present here a short (< 30 min) task-free paradigm in which auditory linguistic
stimuli are presented to MEG subjects, without explicitly requiring their attention, and the
resulting data are analysed using MVPA-based machine learning algorithm to classify brain
responses in different frequency bands as reflecting meaningful lexical input as well as its
semantic and syntactic properties.

Methods

Participants

The experiment was conducted according to the principles of the Helsinki Declaration and
was approved by Central Jutland Region Committee on Health Research Ethics. MEG data
were acquired in seventeen healthy right-handed (handedness assessed using Oldfield (1971))
native speakers of Danish (age range 18–27 years, 12 females) with normal hearing and no
record of neurological impairments. All participants gave written consent before the start of
the experiment and received payment for their participation.

Stimuli

Since we wished to address a range of different neurolinguistic processes – those at lexical,
semantic and syntactic levels – we chose stimulus items which could enable us to contrast a
combination of different linguistic phenomena while controlling for acoustic features (see
Fig. 1 for examples of the stimuli used). To this end, we followed a previously suggested
strategy (Gansonre, Højlund, Leminen, Bailey, & Shtyrov, 2018) and selected a set of spoken Danish-language stimuli which (i) belonged to different lexical and semantic categories (action-related verb, abstract verb, object-related noun and meaningless pseudoword), (ii) were close in terms of phonology so we could compare them directly with minimal acoustic/phonetic confounds, and which (iii) could be modified morpho-syntactically in the exact same way and nonetheless exhibit different linguistic properties (i.e., grammatically correct vs. incorrect) such that we could test the very same contrasts in different linguistic contexts.

These requirements led to the choice of four main base stimuli: bide ([biːðə], to bite), gide

Figure 1

(A) Examples of spectrograms of spoken stimuli used in the experiment. (B) Examples of waveforms plotted on top of each other.

These requirements led to the choice of four main base stimuli: bide ([biːðə], to bite), gide...
([giːdə], *to bother), mide ([miːdə], a mite), *nide ([niːdə], *pseudoword). These were presented as such. Note that they have identical CVCV phonological structure and only differ in the first consonant. The second syllable [ðə], which allows recognition of the lexical items, is the same across all items. To ensure that the full recognition of each particular word form in the restricted experimental context is only possible at the second syllable, we also included, in a 1:1 ratio with all other stimuli, all four first syllables in isolation: [biː], [giː], [miː] and [niː:]. These served as fillers to ensure identical acoustic divergence point across the four types, to be used for time-locking brain activity to, and were not analysed as such.

The above quadruplet provided us with a way to address both lexical and semantic contrasts. By estimating the brain activity elicited by the same word-final syllable [ðə], we could compare, on the one hand, word vs. meaningless pseudoword activation, putatively indicating lexical access, which we expected to find its reflection in an automatic activation of the core left temporo-frontal language system (Tyler & Marslen-Wilson, 2008). On the other hand, by comparing action vs. non-action items we could address semantically-specific aspects of these activations. Previous EEG, MEG and fMRI research has indicated automatic involvement of the brain’s motor system in the comprehension of action-related verbs (Pulvermüller, 2005; Pulvermüller & Fadiga, 2010); we therefore expected more pronounced centro-frontal activity for the action verb bide, but not the concrete noun mide.

Based on the above base forms, we produced further stimuli that included a balanced morphosyntactic contrast. We took advantage of Danish morphology and the fact that the morphemes -(e)t and -(e)n can be used to express the past participle of verbs and definiteness on common nouns. This enabled us to compare the inflected items based on their syntactic congruence or incongruence, e.g. -n in miden vs. *giden, and -t in gidet vs. *mider (where *
indicates a violation of the stem/affix syntactic agreement). Note that each of these pairs have identical codas (t/n) that lead to grammatical/morphosyntactic violation in a counterbalanced fashion: each of them is correct in combination with one but not the other stem. These were presented, in equal proportion along with the other stimuli above, to make sure syntactic properties are only recognised at the very last consonant. To balance for these acoustic modifications, we also included similar items based on other forms (bide[n/t] and nide[n/t], all meaningless), which were used to make a balanced design, but not analysed as such.

The stimuli were made based on a digital recording of a male native speaker of Danish in an anechoic chamber (recording bandwidth: 44k Hz, 16bit, stereo). The first and second syllables of four CVCV stimuli were recorded independently, in order to avoid possible co-articulation effects, and cross-spliced together, such that the second syllables were physically identical across all items. The second syllable commenced at 300 ms after the onset of the first one, and this was the earliest time (the so-called disambiguation point) when any lexical or semantic effects could be expected in the MEG data.

To produce the morphosyntactic items, ending in [t] or [n], we cross-spliced recordings of these two morphemes onto the four main stems, in order to obtain words either violating or respecting rules of Danish morphology such that the exact same phonemes completed syntactic or asyntactic forms in a counterbalanced fashion. These morphemes became distinct at 408 ms after the word onset, and this was therefore the earliest time any morphosyntactic contrasts could affect the brain responses.

The sounds were matched for loudness, with a 1.93 dB drop between the first and the second
syllables so that our stimuli sounded as natural as possible and normalised to have identical power (measured as root-mean-square, RMS). All sound editing was done using Adobe Audition CS6 software (Adobe Inc., San Jose, CA).

In sum, the stimulus set included four CV syllables, four CVCV stems, four CVCV+[t] and four CVCV+[n] forms, all strictly controlled for phonological and acoustic properties. These were combined, in a pseudorandom fashion, in a single auditory sequence ensuring that the stimuli’s lexical, semantic and syntactic properties were available at stringently defined times.

Procedure

The MEG recording was conducted in an electromagnetically shielded and acoustically attenuated room (Vacuum Schmelzer GmbH, Hanau, Germany). During the recordings, participants were instructed to focus on watching a silent film and to pay no attention to the sounds. The auditory stimuli were controlled using Neurobehavioral Systems Presentation v16 (www.neurobs.com) and presented through in-ear-tubes (Etymotic ER-30) binaurally at 50 dB above individual auditory threshold.

All sixteen stimuli were presented equiprobably in a single data acquisition session interspersed in a continuous manner, with 100 pseudorandom repetitions of each stimulus resulting in 1600 epochs in total. The stimulus onset was fixed at 1000 ms (The fixed ISI was based on previous studies of automatic neurolinguistics processing using non-attend designs, which served as the starting point for the current paradigm). The total recording time was 28 minutes.
MEG data were acquired with an Elektra Neuromag Triux MEG setup (Elekta Neuromag Oy, Helsinki, Finland), with 102 magnetometers and 204 planar gradiometers; for eye movement and heartbeat artifact detection 2 bipolar EOG and 1 bipolar ECG recordings were taken. Cardinal landmarks and additional head points were digitised using a Polhemus FASTRAK setup (Polhemus, Vermont, USA). Data were recorded at 1000 Hz, a high pass filter of 0.1 Hz and low pass of 330 Hz were applied online. Head position and head movements were continuously tracked using four Head Position Indicator Coils (HPIs). The participants were lying still on a non-magnetic patient bed, with their head as close to the top of the helmet as possible, the MEG dewar being in supine position.

**Data preprocessing**

All data were preprocessed using MNE-python version 0.16 software package (Gramfort et al., 2013). First, continuous data were bandpass filtered from 1 to 95 Hz, downsampled to 500 Hz, and epoched into single-trial epochs of 1000-ms duration, starting 100 ms before and ending 900 ms after stimulus onset. Bad channels were automatically detected and interpolated, epochs with excessive bad channels discarded and outlier trials removed using an automatic approach (as implemented in Autoreject utility, see Jas, Engemann, Bekhti, Raimondo, & Gramfort, 2017, for details). On average per subject there were 23.24 bad channels (median: 32, std: 10.93) and 12.9 (std: 18.66) bad epochs. No signal-space separation transformation (SSS, also known as maxfiltering) was applied at any stage of the preprocessing. Thereby cleaned epoch data were bandpass-filtered into five frequency bands (Dalal et al., 2011): $\alpha$ within 8-12 Hz, $\beta$ (13-30 Hz), $\gamma$-low (30-45 Hz), $\gamma$-medium (55-70 Hz) and $\gamma$-high (70-90 Hz).
Source reconstruction

For each participant, a T1 and T2 structural magnetic-resonance images (MRIs) were obtained using a Siemens Tim Trio 3T MR scanner. The images were segmented with separate surfaces created for the grey matter, inner skull and skin using SimNIBS utility (Thielscher, Antunes, & Saturnino, 2015). For each subject, individual 3-layer boundary element model (BEM) was calculated together with individual forward models. A common template grey matter surface was created by averaging all of the study participants, using FreeSurfer software (Dale, Fischl, & Sereno, 1999).

Source reconstruction was carried out using an LCMV beamformer (Van Veen, Van Drongelen, Yuchtman, & Suzuki, 1997) using planar gradiometer data and previously developed (Westner, 2017; Westner & Dalal, 2017) Hilbert beamformer. The decision to use only gradiometers was chosen as mixing channel types is not trivial due to magnetometers and planar gradiometers producing values of different scales, furthermore planar gradiometers are less sensitive to external magnetic sources and have a better signal-to-noise ratio compared to magnetometers (Hari, Joutsiniemi, & Sarvas, 1988). First, for each frequency band of interest ($\alpha$, $\beta$, $\gamma$-low, $\gamma$-medium, and $\gamma$-high), the epochs were bandpass-filtered for each subject without subtracting the evoked signal from the single trials, as we were interested in investigating the complete information contained in the responses time-locked to the auditory stimuli. Secondly, an adaptive filter was created by combining responses to all the stimuli in the paradigm, using a 3-layer BEM and a cortically constraint source space. The adaptive filter was computed using a data covariance matrix based on all time points of the bandpass-filtered epochs; the covariance matrix was not regularized prior to inversion. Source orientation was optimised by using the orientation of maximum signal power (cf. Sekihara & Nagarajan, 2008). Neural Activity Index (NAI, cf. Sekihara & Nagarajan, 2008), which incorporates the weight normalisation using a unit-gain...
beamformer, was selected as the output value. Third, after the adaptive filters were created, the single-trial epoched data were Hilbert-transformed and the adaptive filter was applied to the complex data, providing a source reconstruction of the Hilbert-transformed single-trial data. Lastly, we calculated Inter-trial Phase Coherence (ITPC) of thereby obtained single-trial source space data. This was done for each time point and in each source space location independently using the equation below:

$$\text{ITPC}_{tf} = \left| n^{-1} \sum_{r=1}^{n} e^{ik_{tfr}} \right|$$

where $n$ is the number of trials, $e^{ik}$ provides a complex polar representation of the phase angle $k$ on trial $r$ for the time-frequency point $tf$, where frequency is the frequency band (for a thorough review, see Cohen, 2014, Chapter 19, esp. pp. 244-245). This resulted in a single ITPC time course for each point in the source space for each subject.

After the source space ITPC data were calculated for each subject, the individual data were morphed onto a common template surface (5124 vertices) created as the mean over all individual participants for data standardisation across the group. Finally, the data were smoothed with a 10 ms rolling window mean in temporal dimension for each source independently.

**Multivariate pattern analysis**

For each participant, the morphed ITPC time series for each point in the source space was extracted based on the contrast in question. Common for all the Multivariate Pattern Analysis (MVPA) was the classification over time, i.e. for each time point a classifier pipeline was applied giving a classification score over time. We used the entire cortical source space for the classification at each time point. As it was comprised of 5124 vertices, it gave 5124
The classifier pipeline was constructed in MNE-python using scikit-learn utility (Pedregosa et al., 2011) and composed of three steps: First, the features were standardised (z-scored); the standardisation was done across all vertices in the source space at each time point independently using the training set and then applied to the test set. Second, feature selection was done using cross-validated Lasso model (stratified folds, n=4) to create a sparse feature space that was adaptive for each time point, i.e. allowing for number of relevant features to be different for different time points. Lastly, a logistic regression (C=1) was used to classify the two contrasts, and Receiver Operating Characteristic Area Under the Curve (ROC-AUC) was used as the classification score. The pipeline was applied across subjects, meaning that we obtained a decoding score over time for all participants and, hence, we only looked for effects that could hold across the subjects in the tested population.

MVPA can easily overfit data (i.e. become biased towards a specific response) and, in order to prevent that from happening, cross-validation was used for the pipeline. Simply put, cross-validation makes use of all the data by first splitting the data into two smaller data sets. One set, called the training set, is used for standardising the features and fitting (training) the MVPA model. The other data set, called the test set, is then used to test accuracy of the fitted model (i.e. trained model) by trying to predict the class of the new label and, by comparing the prediction to the actual class of the test set, we can calculate the ROC-AUC for the model. By creating new training and test sets from the full data set, it is possible to use all the data for both training and testing (see Varoquaux et al. 2017 for more details and strategies for cross-validating brain imaging data).
All the steps in the pipeline were cross-validated with stratified folds (n=5). Stratified folds imply that ratio of classes in the all the data is maintained in all the cross-validated folds; e.g. in the lexical condition there are 3 real words and one pseudoword for each participant, so in the stratified folds the ratio 3:1 would remain such that even when shuffled there will always be a 3:1 ratio. For the semantic condition and morphosyntax comparison the ratio was 1:1, i.e. 50%. We had 17 participants in the study and it was, therefore, impossible to have an equal number in each of the cross-validation folds and to make cross-validation splits that have the exact same number of participants in all folds and keep the ratio the same. However, by using stratified folds we kept the balance between the folds as equal as possible.

To test for statistical significance of the classification we used permutation tests (Ojala & Garriga, 2010). A permutation test was performed for each time point independently. First, the arrangement of the labels was shuffled, e.g. action verb and object noun in the semantic condition, such that the data might be from an action verb but the label tell the MVPA algorithm that it is an object noun. By repeatedly shuffling the labels and running the classification (n=2000) we can build a null distribution of random ROC-AUC scores. This null distribution is interpreted as the distribution of what the ROC-AUC score could be just by chance. Hence, we can then assess the classification score we had from the real labels by comparing to the null distribution. If number of random scores that are better than the actual classification score is less than or equal 5%, the classification score is said to be better than chance, where 5% is the α level choosen.

To optimise the computational time, only classification scores that surpassed the threshold
(calculated, for each frequency band independently, as the mean score of the baseline plus 1.5 standard deviations of the baseline) were tested for statistical significance. So, for each time point where the ROC-AUC score was above the threshold we made a permutation test (\( n = 2000 \)). As the minimum p-value that can be obtained with a permutation test is dependent on the number of permutations run, the p-value is defined as \( p_u = \frac{(b+1)}{(m+1)} \), where \( b \) is the number of times the permutation was equal or more extreme than the observed classification value and \( m \) is the number of permutations, and 1 is added to \( b \) and to \( m \) to ensure that the estimated p-value is not zero and to avoid division by zero (see Puolivali, Palva, & Palva (2018); and Phipson & Smyth (2010)). So, the p-value in a permutation test does not behave in the same way as a p-values in a parametric test, it is rather a description of observed probability that the given classifier is better than chance. Since we perform a series of permutations, i.e. one for each time point, we need to consider the multiple comparison problem, i.e. the problem that some tests will be significant purely by chance and not be a true effect. We opted to apply cluster threshold correction, i.e. there needs to be a continuous range of significant time points (at least 10 ms or longer) to accept it as a non-random effect.

The 10 ms threshold is based on previous language and auditory research (Edmonds & Krumbholz, 2014; Hagoort & Brown, 2000; Wang, Zhu, & Bastiaansen, 2012), typically using 10 ms as the shortest bin duration to test. It should be noted that only two of the thirteen clusters reported as significant at 10 ms, the rest are longer, ranging from 12 ms to 42 ms.

For each linguistic contrast we fitted the classification pipeline independently. For the lexical contrast, we tested if we could classify lexical features, i.e. if the participant heard a real word vs. an acoustically similar pseudoword (see the Stimuli section above). In the semantic contrast, we tested classification of the action verb vs. object noun. In the syntactic contrast, we tested whether we could correctly classify ungrammatical vs. grammatically correct items.
(irrespective of their acoustic features, i.e. *midet, *giden vs. miden, gidet). Each contrast was tested from the relevant disambiguation point (DP: 300 ms for lexis and semantics and 408 ms for morphosyntax) through the end of the epoch. Latencies reported in the Results section below are measured relative to DP.

**Results**

**Lexical contrast**

For the lexical contrast (see Figures 2, 3, 4, Table 1) we found that it was possible to significantly decode words vs. pseudowords in the β, γ-low, γ-medium and γ-high bands. The earliest significant decoding was achieved in the γ-medium band already at 62-76 ms after the disambiguation point. The pattern of classifier-selected features in the left hemisphere included the frontal lobe (BA 44, border of BA 6 / BA 8, BA 4), parietal BA 7, junction of BA 7 / 39, and BA 39/19. In the right hemisphere, features in temporal lobe BA 22 and frontal BA 9, BA4 were selected. The highest ROC-AUC score was in the γ-low band at 224 ms after DP (ROC-AUC: 94.35%, SD: 4.5%). The cluster including this AUC-ROC peak spanned from 222 ms to 238 ms and involved a broad pattern of features including, in the left hemisphere, temporal lobe (BA 22), frontal areas (BA 11, border of BA10/47, BA44 dorsal/posterior), and parietal areas (BA 40, BA 7 as well as BA 3/1/2). In the right hemisphere, it included the temporal lobe (BA 22, BA 23, BA 43), frontal areas (BA 11, border of BA 9/10/46, posterior dorsal BA44), and parietal areas (BA 39, BA 40, BA 7, BA 1/2). No significant classification results were obtained in the α-band.
Figure 2

*Top: heatmap of significant clusters across three linguistic contrasts, five frequency bands and time. Lexical condition in blue colours, semantic in green, and syntax in red. Bottom: surface topography of significant effects. For all conditions, colours go from lighter to darker as latency becomes longer.*

Table 1: Table of significant clusters in the lexical condition sorted by time from the divergence point. Peak is highest ROC-AUC scores of the cluster. Peak std is the standard
deviation of cross-validation folds for the peak ROC-AUC score. Peak time is the time of the peak from DP. Cluster start is the start time of the cluster from DP. Cluster end is the end time of the cluster from DP. Cluster length is the length of the cluster. Cluster mean is the mean ROC-AUC score of the cluster. Cluster std is the standard deviation of the Cluster mean across cross-validation folds.

Figure 3
Figure 3. (A) Model patterns: interpreting coefficients of a machine learning model is not trivial and a high coefficient value does not necessitate a high signal value in the MEG data (see Haufe et al., 2014 for details). “Model patterns” are a way to highlight the signal in a
neurophysiological sensible way that is directly interpretable compared to the raw coefficients (Haufe et al., 2014). We show top and bottom 5% of the patterns in the \( \gamma \)-low band from 222 to 238 ms. Blue colours are areas of activation able to predict real words and yellow/red are areas used to predict pseudo word. (B) Average top and bottom 5% of ITPC difference; blue colours indicate higher ITPC for real words and yellow/red indicate higher ITPC for pseudo word \( \gamma \)-low band from 222 to 238 ms. (C) Average ITPC over time; solid lines are the average of the selected features, dashed lines are the average of all vertices in the source space. Time zero is the divergence point, when stimuli could be recognised from the available acoustic information.

Figure 4. Heatmap of ROC-AUC scores for all bands in lexical condition. Note that chance in this condition is 75%. Time is relative to DP

Semantic contrast

In the semantic contrast (see figures 2, 5, 6, Table 2), the peak classification score was found in the \( \alpha \)-band at 106 ms after DP (AUC-ROC: 91.11%, SD: 12.96%). The cluster including
the peak AUC-ROC was between 80-122 ms after the onset of the second syllable disambiguating the semantics of the particular form. The cluster comprised a pattern of features including both temporal (BA 22) and frontal areas (BA 9, BA 44, BA 45, BA46) of the left hemisphere. In the right hemisphere, in addition to the temporal (BA 21, BA 22, BA 38, BA 42) and the frontal lobe (BA 6, BA 9, BA 11, BA 44), it also involved parietal areas (BA 7 and BA 2).

Table 2

<table>
<thead>
<tr>
<th>band</th>
<th>peak (%)</th>
<th>peak std (%)</th>
<th>Peak time (ms)</th>
<th>Cluster start (ms)</th>
<th>Cluster end (ms)</th>
<th>Cluster length (ms)</th>
<th>Cluster mean (%)</th>
<th>Cluster std (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>α</td>
<td>91,11</td>
<td>12,96</td>
<td>106</td>
<td>80</td>
<td>122</td>
<td>42</td>
<td>68,21</td>
<td>9,77</td>
</tr>
<tr>
<td>β</td>
<td>75,00</td>
<td>19,08</td>
<td>138</td>
<td>134</td>
<td>146</td>
<td>12</td>
<td>69,62</td>
<td>5,25</td>
</tr>
<tr>
<td>γ-low</td>
<td>84,58</td>
<td>9,01</td>
<td>224</td>
<td>220</td>
<td>230</td>
<td>10</td>
<td>71,85</td>
<td>7,06</td>
</tr>
<tr>
<td>β</td>
<td>70,00</td>
<td>13,43</td>
<td>256</td>
<td>254</td>
<td>264</td>
<td>10</td>
<td>66,53</td>
<td>3,37</td>
</tr>
<tr>
<td>α</td>
<td>85,83</td>
<td>3,74</td>
<td>584</td>
<td>574</td>
<td>590</td>
<td>16</td>
<td>71,71</td>
<td>7,62</td>
</tr>
</tbody>
</table>

Table 2: Table of significant clusters in the Semantic condition sorted by time from the divergence point. See the legend of table 1 for an explanation of the columns.
Figure 5. (A) Model patterns (see also Figure 4 legend): top and bottom 5% of the patterns in

A: Model patterns: action verb vs object noun

B: ITPC contrast between action verb vs object noun

C: Time course of average ITPC

Cluster: 80 - 122 ms
Cluster: 134 - 146 ms
Cluster: 220 - 230 ms
Cluster: 254 - 264 ms
Cluster: 574 - 590 ms
The results of MVPA classification of grammatically correct vs. incorrect inflections (see figure 2, 7, 8, table 3) showed the peak ROC-AUC score in the γ-low band at 90 ms after the syntactic DP (ROC-AUC: 69.22%, SD: 13.05%). The significant classification cluster on the left hemisphere included the temporal lobe (BA 21, BA 22, BA 37, BA 41), frontal areas (BA 4, BA 9 BA 10, BA 11), as well as parietal areas (BA 40/7/2), and occipital area BA 19. In the
right hemisphere, it involved temporal (BA 21, 22, 38, 42), frontal (BA 44, 11, 4, 6, 9),
parietal (BA 39, BA 40, 2, 1, 7), and occipital (BA 18, 9) areas.

Table 3

<table>
<thead>
<tr>
<th>Syntax</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>peak (%)</td>
<td>peak std (%)</td>
</tr>
<tr>
<td>J</td>
<td>J-LOW</td>
</tr>
<tr>
<td>J</td>
<td>J-LOW</td>
</tr>
</tbody>
</table>

Table 3: Table of significant clusters in the syntax condition sorted by time from the divergence point. See the legend of table 1 for an explanation of the columns.
Figure 7. (A) Model patterns: top and bottom 5% of the patterns in the y-low band from 84 to 98 ms. Blue colours are areas used to predict correct syntax and yellow/red are areas used to predict incorrect syntax. (B) Average top and bottom 5% of ITPC difference. Blue colours indicate higher ITPC for correct syntax and yellow/red indicate higher ITPC y-low band from 84 to 98 ms (C) Average ITPC over time; solid lines are the average of the selected
features, dashed lines are the average of all vertices in the source space. Time is relative to the divergence point.

![Figure 8. Heatmap of ROC-AUC scores for all bands in syntax condition.](image)

**Classification of amplitude data**

We also tested ad hoc if we could decode language properties using amplitude data. We only found significant clusters in semantic in the θ-medium band from 500 ms to 512 ms after DP. For the morphosyntactic condition we found that we could classify correctly in the θ-high bands (from 114 ms to 128 ms after DP) and β band (from 248 ms to 262 ms after DP). As we were not able to decode all conditions successfully, we have left out the amplitude results from further discussion.

**Discussion**

In this study, we suggested and tested a paradigm which, in the absence of focused attention
on the auditory input or any explicit stimulus-focussed behavioural task, addressed different levels of neurolinguistic information processing in the brain using a carefully crafted set of spoken stimuli with strictly controlled acoustic/psycholinguistic contrasts. We registered the brain’s activity elicited by these speech stimuli using high-density whole-head MEG set-up, and analysed it using machine learning-based multivariate pattern analysis techniques applied to inter-trial phase coherence in a range of frequency bands, at the level of cortical sources calculated using individual MRIs. The results indicated that, by using this approach, we were able to successfully classify lexical, semantic and morphosyntactic contrasts (see Table 4 for an overview of results). These effects were exhibited in different frequency bands and at different times. Below, we will briefly discuss these results in more detail.

Table 4

<table>
<thead>
<tr>
<th>Condition</th>
<th>Band</th>
<th>Peak score</th>
<th>Peak std</th>
<th>Peak time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lexical condition</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>α</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>β</td>
<td>85.97</td>
<td>14.95</td>
<td>538</td>
</tr>
<tr>
<td></td>
<td>γ-low</td>
<td>94.35</td>
<td>4.50</td>
<td>224</td>
</tr>
<tr>
<td></td>
<td>γ-medium</td>
<td>88.53</td>
<td>6.43</td>
<td>66</td>
</tr>
<tr>
<td></td>
<td>γ-high</td>
<td>87.88</td>
<td>9.29</td>
<td>358</td>
</tr>
<tr>
<td>Semantic condition</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>α</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>β</td>
<td>75.00</td>
<td>19.08</td>
<td>138</td>
</tr>
<tr>
<td></td>
<td>γ-low</td>
<td>84.58</td>
<td>9.01</td>
<td>224</td>
</tr>
<tr>
<td></td>
<td>γ-medium</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>γ-high</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Syntax condition</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>α</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>β</td>
<td>69.22</td>
<td>13.05</td>
<td>90</td>
</tr>
<tr>
<td></td>
<td>γ-low</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>γ-medium</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>γ-high</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 4: Table of peak scores for each bands and condition, dash (-) indicate no significant cluster. Peak score is highest ROC-AUC scores of the cluster. Peak std is the standard deviation of cross-validation folds for the peak ROC-AUC score. Peak time is the time of the peak from DP.
Lexical contrast

In the lexical contrast, we found clusters of significant decoding performance (with a peak classification score of ~94%) over the entire duration of the analysis epoch starting from ~60 ms after the divergence point. These predominantly occurred in the $\gamma$ range (including all three sub-bands tested) and were underpinned by activity in bilateral temporo-parietal and frontal clusters. This early onset of lexical effects here is in line with previous ERF and ERP studies that used similar non-attend auditory designs (MacGregor, Pulvermüller, van Casteren, & Shtyrov, 2012; Pulvermüller & Shtyrov, 2006; Shtyrov & Lenzen, 2017) and showed that the brain responds differently to meaningful words vs. meaningless pseudoword stimuli from ~50-80 ms after the acoustic input allows to identify the lexicality of stimuli, with additional lexicality-driven activity spanning across peaks until ~400 ms. Such ERP/ERF results (Shtyrov, Kujala, & Pulvermüller, 2010; see also, e.g. Shtyrov, Pihko, & Pulvermüller, 2005) have reported similar perisylvian configuration of bilateral source activations; importantly, here we find them for higher-frequency phase coherence values, not reported previously.

Increased activity in the $\gamma$ band (quantified as spectral power or as event-related synchronisation) has previously been linked to lexical access possibly underpinning synchronisation of neural elements making up distributed word memory circuits (Lutzenberger, Pulvermüller, & Birbaumer, 1994; Pulvermüller, Preisсл, Lutzenberger, & Birbaumer, 1996; see also, e.g., Tavabi, Embick, & Roberts, 2011). What our findings suggest is that this process may involve multiple synchronisation steps expressed as phase resetting (an important mechanisms in information processing, Canavier (2015) ) at different times,
These previous studies have typically found a single power peak in low \( \gamma \) dynamics, whereas we here find a series of activations at frequencies up to 90 Hz, potentially reflecting the specific advantages of machine learning techniques in classifying distributed clusters of activity. We also found a single significant \( \beta \) cluster at the end of the epoch (>500 ms post-DP). This is fully in line with previous research highlighting the role of \( \beta \) oscillations in lexico-semantic storage and processing (e.g. Bakker, Takashima, van Hell, Janzen, & McQueen, 2015; Brennan, Lignos, Embick, & Roberts, 2014), although our data suggest that this \( \beta \) activity is secondary in relation to the almost immediate phase resetting in the \( \gamma \) band. Note that while a number of above studies have also identified \( \theta \)-band activity in relation to some aspects of speech processing (e.g. syllable tracking (Luo & Poeppel, 2007, 2012)), we optimised our recording for time (with a view of potential applied use) that led to short baselines not suitable for \( \theta \) analysis. There is, however, compelling evidence of a relation between \( \theta \) and \( \gamma \)-band activity (Canolty et al., 2006), and future research using similar paradigms could investigate whether the current ITPC findings may also have a counterpart in the \( \theta \) range.

**Semantic contrast**

The semantic contrast between the action verb and non-action noun has indicated activity from \(~100\) ms after the divergence point in \( \alpha \), \( \beta \) and, to a smaller degree, in a lower \( \gamma \) range, with peak classification scores of \(~91\), 75 and 85\%, respectively. This mostly involved temporo-frontal cortices, largely overlapping with the core language systems, but importantly, indicated elevated frontal involvement, including clusters of activity in inferior-frontal (BA44, 45) and motor (BA6) areas, which is compatible with the motor system involvement in action word processing, posited previously (Pulvermüller et al., 2006). Previous EEG and
MEG research into the brain basis of action-related semantics has found ERF and ERP correlates between 80-200 ms indicating near-immediate and largely automated activation of the motor strip in action word comprehension (Shtyrov et al., 2014; Shtyrov, Hauk, & Pulvermüller, 2004). In line with this previous research, the features indicated here by the semantic contrast did not include many in the temporal lobe but mostly in the inferior-frontal and prefrontal areas in the left and right hemispheres. While the timing of this activation was generally similar – and thus largely parallel – to the lexical processes above, the spectral composition was different, with significant ITPC findings in a lower frequency range.

Changes in $\alpha$ and especially $\beta$ power have previously been reported as related to single word semantics, including action word semantics in particular (Bakker et al., 2015; Vukovic & Shtyrov, 2014). What we show here is that the phase resetting likely linked to these changes can reliably classify words with different meaning. In oscillatory space our results are in line with those previously reported by Mamashli and colleagues (Mamashli, Khan, Obleser, Friederici, & Maess, 2018) who investigated functional connectivity across regions of interest and found activity in the $\alpha$, $\beta$ and low $\gamma$ bands, while Haarman & Camaron (2005) reported a change of coherence in the 10 – 14 Hz range while retaining a sentence.

**Syntax contrast**

Successful classification of the morphosyntactic contrast was achieved in the $\gamma$-low range exclusively and was found to commence at ~100 ms after the syntactic divergence point. While, in terms of the absolute stimulus timing, this difference was later than the lexical and semantic findings above, it is important to note the syntactic disambiguation in the stimulus itself was also possible at a later time, as it was determined by the final consonant (n/t). Thus, in terms of the relative timing, the syntactic properties appear to be assessed roughly in parallel to other tested features, overall in line with the view positing near-simultaneous onset of neurolinguistics processing of different information types (Hagoort, 2004; Marslen-
Wilson, 1987). The timing is in line with the findings of ELAN literature that suggested syntactic parsing to sometimes start as early as 50 ms in an automatic fashion (Herrmann, Maess, & Friederici, 2011). Findings of links between γ band activity and syntactic parsing have been reported in the literature previously (Lewis, Wang, & Bastiaansen, 2015), including morphosyntactic processing of the kind broadly similar to that required in comprehension of complex words used here (Levy, Hagoort, & Démonet, 2014). Importantly, we used a fully balanced set of contrasts, with highly matched word stimuli in which morphosyntactic (in)correctness was carried by physically the same phonemes, i.e. [n] and [t] equally employed in both sets. This activation was underpinned by broad temporo-frontal networks in both hemispheres, in line with previous literature (Bozic, Fonteneau, Su, & Marslen-Wilson, 2015; Friederici, 2011; Hickok & Poeppel, 2007). We have also found activity in parietal and occipital areas, not commonly reported in syntax studies and therefore requiring validation in future research; notably, the values here do not reflect absolute activity (e.g. activity) as such but rather reliable classification of activation, however small it may be. Some studies investigating oscillatory neural activity in relation to syntactic processing have found that there is a link between α power and syntactic properties of phrases (e.g. Leiberg, Lutzenberger, & Kaiser, 2006). While we do not find any similar activity in the α range for our morphosyntactic condition, that may be due to the use of single word stimuli with morphosyntactic modifications (rather than phrases with more elaborate syntax) in our paradigm.

Of the three linguistic contrasts tested, the morphosyntactic one had the lowest decoding percentage (peak at ~69%). One possible reason for this is that it used two acoustically different stimuli, words ending in [n] and [t], with the grammatical correctness independent of the word ending. The different acoustic properties — later onset of [t] than [n] and different amplitude envelopes of the two — may have smeared the effect in time leading to
poorer (but nevertheless significant) classification results. Future studies could use other contrasts and different languages in order to improve classification results.

**MVPA of MEG ITPC data as a tool for objective assessment of neurolinguistics processes**

Previous literature has shown that passive paradigms can be used to investigate language processing (Gansonre et al., 2018; Shtyrov et al., 2012), an important first step towards assessing participants and patient groups that have difficulties responding verbally or in other behavioural ways. However, the next necessary step is to optimise the analysis of data obtained in such paradigms. Important issues arise when trying to automate MEG analysis both at preprocessing stage and at the level of identifying significant effects and differences either within or between groups. What we present here is a tentative proposal to solve these issues. By using the automated data cleanup protocol combined with a single-trial beamformer reconstruction, we reduce the number of manual steps needed; and further, by applying MVPA we avoid having to a priori select time and ROIs for the statistical assessment of the difference between groups. The focus on oscillatory dynamics and ITPC provides us the ability to assess activity in different frequency bands simultaneously, resulting in a more detailed picture of the neural activity related to neurolinguistic processing. Previously, such analyses have been hampered by the number of tests that are needed to statistically evaluate the differences across groups/conditions. In order to solve this problem, we used MVPA, as this approach makes use of the powerful machine learning techniques capable of assessing the effects statistically in an unbiased fashion. So, by combining ITPC and MVPA we have the possibility to assess data in a detailed and yet exploratory manner.

Common to the effects across all conditions is that they last a relatively short time compared
to the length of the trials. It is worth remembering that we are looking for clusters of time
where we can decode a difference in the signals that are compared, which may help explain
temporally sparse results. It is also worth noting that one cannot from a significant cluster
alone conclude what exact processes underpin it. However, comparing the ITPC data and the
patterns of the MVPA models together may provide an overall picture of the activity
underlying the process at hand.

Conclusions

Using a passive paradigm, we probed several different neurolinguistic properties. Separating
the data into different frequency bands and looking at ITPC, we found that, by using MVPA,
we could classify lexical, semantic, and syntactic information processing. The best
classification results varied between the different neurolinguistic properties both in time and,
importantly, frequency bands, with lexical processes classified predominantly by broad \( \gamma \),
semantic distinctions – by \( \alpha \) and \( \beta \), and morphosyntax – by low \( \gamma \) feature patterns. Crucially,
all types of processing commenced in a near-parallel fashion from \( \sim 100 \) ms after the auditory
information allowed for disambiguating the spoken input. This shows that individual
neurolinguistic processes take place near-simultaneously and involve overlapping yet
distinct neuronal networks that operate at different frequency bands. Further investigations
are needed to understand the precise relation of the time courses, frequency bands, neuronal
substrates and neurolinguistic properties, and to test the applicability of this approach to
detecting linguistic anomalies in various populations.

Competing interests

The authors declare that no competing interests exist.
References


sentence processing classes: Evidence from event-related brain potentials.  
Journal of Cognitive Neuroscience, 3(2), 151–165.


https://doi.org/10.1523/JNEUROSCI.4250-04.2005


https://doi.org/10.2202/1544-6115.1585


https://doi.org/10.1038/nrn2811


auditory processing: Oscillatory neural entrainment reveals segregation of function at different scales. *PLOS Biology, 15*(11), e2000812. https://doi.org/10.1371/journal.pbio.2000812


A: Model patterns: real words vs pseudo word

B: ITPC contrast between real words and pseudo word

C: Time course of average ITPC
A: Model patterns: action verb vs object noun

B: ITPC contrast between action verb vs object noun

C: Time course of average ITPC

- Cluster: 80 - 122 ms
- Cluster: 134 - 146 ms
- Cluster: 220 - 230 ms
- Cluster: 254 - 264 ms
- Cluster: 574 - 590 ms
A: Model patterns of activity related to predicting correct and incorrect syntax

B: ITPC contrast of correct and incorrect syntax

C: Time course of average ITPC