



Alpha and theta brain oscillations index dissociable processes in spoken word recognition



Antje Strauß^{a,*}, Sonja A. Kotz^{b,c}, Mathias Scharinger^a, Jonas Obleser^a

^a Max Planck Research Group "Auditory Cognition", Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstraße 1A, 04103 Leipzig, Germany

^b Research Group "Subcortical contributions to comprehension", Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

^c School of Psychological Sciences, The University of Manchester, Manchester, UK

ARTICLE INFO

Article history:

Accepted 3 April 2014

Available online 18 April 2014

Keywords:

Time–frequency analysis

Lexical decision

EEG

Source localization

Slow neural oscillations

ABSTRACT

Slow neural oscillations (~1–15 Hz) are thought to orchestrate the neural processes of spoken language comprehension. However, functional subdivisions within this broad range of frequencies are disputed, with most studies hypothesizing only about single frequency bands. The present study utilizes an established paradigm of spoken word recognition (lexical decision) to test the hypothesis that within the slow neural oscillatory frequency range, distinct functional signatures and cortical networks can be identified at least for theta- (~3–7 Hz) and alpha-frequencies (~8–12 Hz). Listeners performed an auditory lexical decision task on a set of items that formed a word–pseudoword continuum: ranging from (1) real words over (2) ambiguous pseudowords (deviating from real words only in one vowel; comparable to natural mispronunciations in speech) to (3) pseudowords (clearly deviating from real words by randomized syllables). By means of time–frequency analysis and spatial filtering, we observed a dissociation into distinct but simultaneous patterns of alpha power suppression and theta power enhancement. Alpha exhibited a parametric suppression as items increasingly matched real words, in line with lowered functional inhibition in a left-dominant lexical processing network for more word-like input. Simultaneously, theta power in a bilateral fronto-temporal network was selectively enhanced for ambiguous pseudowords only. Thus, enhanced alpha power can neurally 'gate' lexical integration, while enhanced theta power might index functionally more specific ambiguity–resolution processes. To this end, a joint analysis of both frequency bands provides neural evidence for parallel processes in achieving spoken word recognition.

© 2014 Elsevier Inc. All rights reserved.

Introduction

Accumulating evidence shows that speech comprehension is more completely described by not only looking at evoked but also induced components of the electrophysiological brain response (Ghitza, 2011; Giraud and Poeppel, 2012). Besides research concerning the phase (for review see Peelle and Davis, 2012), also power changes of transient slow oscillations have been found to determine language processes (Bastiaansen et al., 2008; Hald et al., 2006; Meyer et al., 2013; Obleser and Weisz, 2012). However, a functional differentiation between alpha (~8–12 Hz) and theta oscillations (~3–7 Hz), even though previously put forward (e.g., Klimesch, 1999; Roux and Uhlhaas, 2014; for current debate in audition see e.g., Weisz et al., 2011), remains to be shown for speech processing (e.g. an open issue in Obleser and Weisz, 2012; Tavabi et al., 2011).

Generally, alpha oscillations are the predominant rhythm in ongoing neuronal communication and therefore observable in diverse cognitive functions such as auditory processing (sometimes labeled 'tau'; Lehtelä et al., 1997; Tavabi et al., 2011; Hartmann et al., 2012), attention

(Klimesch, 2012), working memory (e.g., Meyer et al., 2013; Obleser et al., 2012; Wilsch et al., 2014), or decision making (Cohen et al., 2009). A tentative theoretical account on the role of alpha oscillatory activity has only been put forward recently (Jensen and Mazaheri, 2010; Klimesch, 2012; Klimesch et al., 2007a): functional inhibition. In fact, most of the above-cited data are compatible with increased needs for inhibition of concurrent, task-irrelevant, or task-detrimental neural activity. Also, direct evidence for alpha-mediated inhibition of local neural activity, as expressed in spiking (Haegens et al., 2011) or gamma-band activity (Roux et al., 2013; Spaak et al., 2012), has been provided.

To this end, first evidence has shown that greater alpha suppression post-stimulus is associated with more effective language processing: alpha oscillations in response to single words were found to be suppressed as a function of intelligibility of acoustically degraded words (Obleser and Weisz, 2012). This is in line with the inhibitional account meaning that alpha power remains high when the language processing network is inhibited, the crucial mechanism for the present study.

In contrast to functional inhibition across a range of general cognitive functions plausibly associated with alpha, theta oscillations in human EEG have been related more consistently to episodic memory (e.g., Hanslmayr et al., 2009), sequencing of memory content (e.g., Lisman and Jensen, 2013; Roux and Uhlhaas, 2014), and matching of

* Corresponding author.

E-mail address: strauss@cbs.mpg.de (A. Strauß).

new information with memory content (e.g., Klimesch, 1999). Moreover, neural periodic reactivation of information held in human short-term memory has been directly related to theta-timed oscillatory cycles (Fuentemilla et al., 2010). Such ‘replay’ of sensory evidence in order to arrive at accurate lexical decisions might be decisive in the present design, especially when input is somewhat ambiguous as outlined below.

Interestingly, theta power enhancement has been observed in a series of language- or speech-specific effects. For example, semantic violations more than world knowledge violations drive theta enhancement during sentence processing (Hagoort et al., 2004; Hald et al., 2006); also, the retrieval of lexico-semantic information (Bastiaansen et al., 2008) as well as the increasing intelligibility of acoustically degraded words (Obleser and Weisz, 2012) lead to theta enhancement. Note that in the latter study, the alpha suppression reported above was directly proportional to theta enhancement. These results tie theta enhancements in language paradigms to the neural re-analysis of difficult-to-interpret stimulus materials.

In the present study, we want to dissociate neural oscillatory dynamics in the alpha and theta frequency bands in order to link them to segregable functions in spoken word recognition. As a control, however, we also extracted event-related potentials (ERPs) because its N400 component in particular has proven to be a robust index of ‘wordness’ (Chwilla et al., 1995; Desroches et al., 2009; Friedrich et al., 2009; Laszlo et al., 2012; for review see Friederici, 1997; Van Petten and Luka, 2012). Larger N400 amplitudes, elicited by unexpected (Connolly and Phillips, 1994; Kutas and Hillyard, 1980; Strauß et al., 2013), infrequent words (Dufour et al., 2013; Rugg, 1990; Van Petten and Kutas, 1990), or pseudowords (Friedrich et al., 2006), compared to high-probable or high-frequent real words, have mostly been associated with increased neural processing effort in matching the input signal to items in the mental lexicon. We hope to shed new light on this matching process by investigating alpha and theta oscillations which are framed in terms of inhibition and replay.

We designed an auditory lexical decision task where a word-pseudoword continuum would induce a stepwise reduction in lexical accessibility (‘wordness’). Additionally, ambiguous stimuli would evoke a task-dependent conflict (task: ‘Is it a word (yes/no)?’) and call for re-evaluation of the auditory input. First, we hypothesize that a neural ‘wordness’ effect should be observable in the alpha band, with less alpha power when auditory input approximates real words held in the mental lexicon. This effect should be prominent in brain areas associated with lexical processes (e.g., left middle temporal gyrus; Kotz et al., 2002; Minicucci et al., 2013) and would characterize alpha as a signature of enabling lexical integration. Second, we hypothesize that the power of theta oscillations with their ascribed functionality in memory and lexico-semantics would vary with the need for resolving ambiguity.

Altogether, our focus on dissociable slow neural oscillations and their corresponding functional roles during spoken word recognition allows us to contribute to long-standing debates on whether recognition is best conceived as serial, feed-forward mechanisms (Norris et al., 2000) or as parallel, interacting processes (Marslen-Wilson, 1987; McClelland and Elman, 1986). Importantly, time-frequency analyses of on-going EEG activity are ideally suited to extract potentially parallel cognitive processes.

Methods

Participants

Twenty participants (10 female, 10 male; 25.6 ± 2.0 years, $M \pm SD$) took part in an auditory electroencephalography (EEG) experiment. All of them were native speakers of German, right-handed, with normal hearing abilities, and reported no history of neurological or language-related problems. They gave their informed consent and received

financial compensation for their participation. All procedures were approved of by the ethics committee of the University of Leipzig.

Stimuli

Adapted from Raettig and Kotz (2008), stimuli were 60 three-syllabic, concrete German nouns (termed ‘real’, e.g., ‘Banane’ [banana]). For the ‘ambiguous’ condition, we exchanged the core vowel of the second syllable (e.g., ‘Banene’). Finally for the ‘pseudoword’ condition, we scrambled syllables across words (concrete and abstract, see below), while keeping their position-in-word fixed (e.g., ‘Bapossner’). Note that there was a fourth condition with 60 three-syllabic, abstract German nouns not relevant for the current analyses which was necessary to maintain an equal ratio of words and pseudowords. These were considered as fillers and not analyzed further. Previous studies used word-like stimuli in order to investigate lexicality effects on phoneme discrimination (Connine and Clifton, 1987; Frauenfelder et al., 1990; Wurm and Samuel, 1997). An important difference to these studies is that we created a distribution of formant distances between real word vowels and their pseudoword equivalents. For illustration purposes, these difference can be quantified by calculating the Euclidian distance of the first three formants for each vowel pair (Obleser et al., 2003): Distances ranged from 200 Hz (/ε/ → /i/, Geselle → Gesille) to 2100 Hz (/o:/ → /i:/, Kommode → Kommode). The majority (approximately one third) of vowel pairs were 600 to 1000 Hz apart from each other (/ə/ → /ɔ/, Batterie → Battorie). Therefore, exchanging a vowel here means that stimuli were lexically but not phonetically ambiguous which calls for ambiguity resolution processes on a decisional rather than a perceptual level (for discussion see Norris et al., 2000). However, we show with this acoustic analysis that lexical ambiguity necessarily corresponds to variance in acoustic input.

Importantly, we controlled for equal ratio of stress patterns across conditions, because in unstressed syllables formant distance decreases, which raises perceptual confusions and task difficulty. The substitution of the vowel marked the deviation point to any existing German word but at the same time did not violate German phonotactic rules. The same holds true for clear pseudowords even though deviation points were not as exactly timed as in the ambiguous condition and alternated between the first and second phoneme of the second syllable. Please note that ambiguous stimuli had only one real word neighbor whereas clear pseudowords might have evoked several real word associations.

All words and pseudowords were spoken by a trained female speaker and digitized at 44.1 kHz. Post-editing included down-sampling to 22.05 kHz, cutting at zero crossings closest to articulation on- and off-sets, and RMS normalization. In sum, the experimental corpus consisted of 240 stimuli with a mean length of $754.2 \text{ ms} \pm 83.5 \text{ ms}$ ($M \pm SD$).

Experimental procedure

In an electrically shielded and sound-proof EEG cabin, participants were instructed to listen carefully to the words or word-like stimuli and to perform a lexical decision task.

Fig. 1A shows the detailed trial timing. After each stimulus, a delayed prompt indicated that a response should be given via button press (‘Yes’/‘No’) to answer whether or not a German word had been heard. The response delay was introduced in order to gain longer trial periods free of exogenous components (due to the visual prompt) or artifacts (i.e., button press), which are required for a clean time-frequency estimation and source localization of oscillatory activity. The button assignment (left/right) was counterbalanced across participants such that 10 participants used their left and the other 10 their right index finger for the ‘Yes’ response. Accuracy scores (percentage correct) and reaction times were acquired. Subsequently, in order to better control for eye-related EEG activity, an eye symbol marked the time period during which participants could blink. Duration of blink break and onset of the next stimulus were jittered to avoid a contingent negative variation.

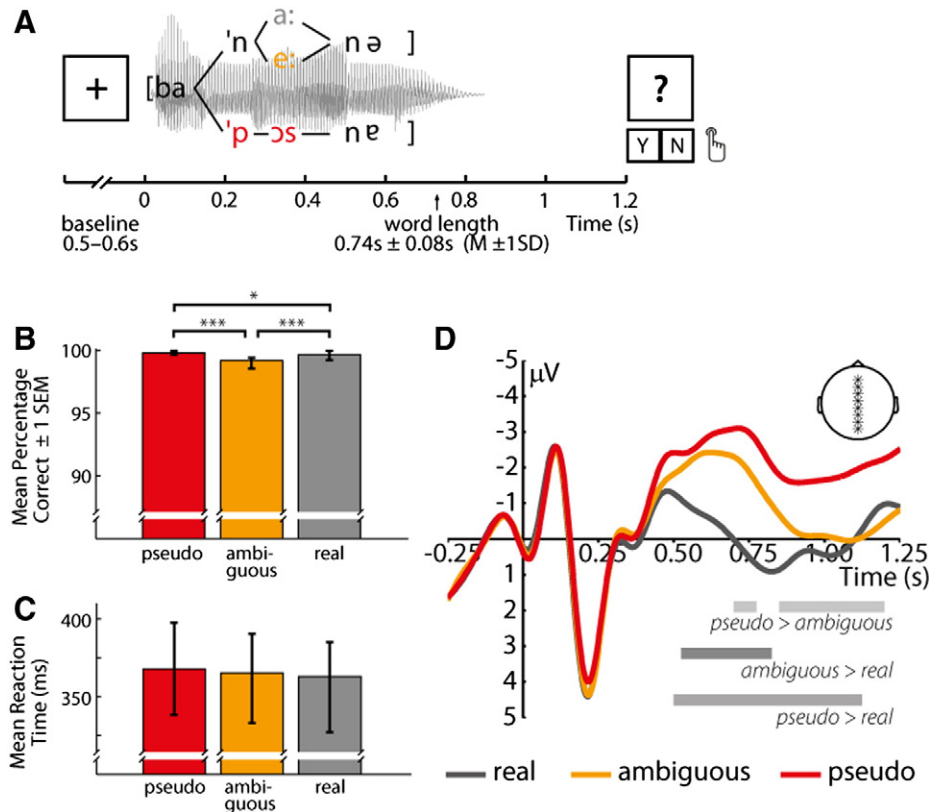


Fig. 1. Study design and behavioral measures. A. Stimulus design and schematic time course of one trial. Stimuli were three syllabic German nouns ('real'), 'ambiguous' pseudowords (one vowel exchanged), and clear 'pseudowords' (scrambled syllables across items). B. Accuracy. Mean percentage correct ± 1 between-subjects standard error of the mean. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. C. Reaction times. Mean reaction times relative to the prompt, ± 1 between-subjects standard error of the mean. D. Event-related potentials. Grand average over midline electrodes. Gray shaded bars indicate statistical differences.

Prior to the experiment there was a short familiarization phase. It consisted of 10 trials taken from Raettig and Kotz (2008) which had similar manipulations but were not used in the present experiment. Then each participant listened to all 240 stimuli. Listeners paused at their own discretion after blocks of 60 trials. The overall duration of the experimental procedure was about 30 min.

Each participant obtained an individually pseudo-randomized stimulus sequence. Note that the order of occurrence for a given ambiguous pseudoword (e.g., 'Banene') and its real word complement (e.g., 'Banane') was counterbalanced across participants in order to control for facilitated word recognition due to ordering effects. As a constraint to pseudo-randomization, their sequential distance was kept maximal (i.e., ~120 other items in between).

Electroencephalogram acquisition

The electroencephalogram (EEG) was recorded from 64 Ag–AgCl electrodes positioned according to the extended 10–20 standard system on an elastic cap with a ground electrode mounted on the sternum (Oostenveld and Praamstra, 2001). The electrooculogram (EOG) was acquired at a horizontal (left and right eye corner) and a vertical (above and below left eye) line. All impedances were kept below 5 k Ω . Signals were referenced against the left mastoid and digitized online with a sampling rate of 500 Hz, with a pass-band of DC to 140 Hz.

Individual electrode positions were determined after EEG recording with the Polhemus FASTRAK electromagnetic motion tracker (Polhemus, Colchester, VT, USA) for more precise source reconstructions.

Data analysis: event-related potentials

Data pre-processing and analysis was done offline by using the open source Fieldtrip toolbox for Matlab™, which is developed at the F.C. Donders Centre for Cognitive Neuroimaging in Nijmegen, Netherlands (Oostenveld et al., 2011). Data were re-referenced to linked mastoids and band-pass filtered from 0.1 Hz to 100 Hz. To reject systematic artifacts, independent component analysis was applied and components were rejected according to the 'bad component' definition by Debener et al. (2010). Remaining artifacts were removed when the EOG channels exceeded $\pm 60 \mu\text{V}$ for frequencies between 0.3 and 30 Hz, which led to whole trial exclusion (3.6 ± 5.3 trials per participant). Resulting clean data were used for subsequent analyses.

To extract event-related potentials (ERPs), epochs were low-pass filtered using a 6th order Butterworth filter at 15 Hz, baseline-corrected (baseline -0.2 to 0 s), and then averaged over trials per condition. As in previous studies (Oleser and Kotz, 2011; Strauß et al., 2013), auditory evoked potentials were considered to be strongest over midline electrodes (FPz, AFz, Fz, FCz, Cz, CPz, Pz, POz, Oz), which were defined as a region of interest (ROI) for the ERP analysis, best capturing the dynamics of the N400 component. On the ERP amplitudes, we performed a time series analysis (49 consecutive steps of 50 ms width, windows overlap by 25 ms thereby covering a time range from 0 to 1.25 s) using repeated measures ANOVA with the factor of wordness (pseudo, ambiguous, real). We assessed p values with Greenhouse–Geisser-corrected degrees of freedom. If p values survived false discovery rate (FDR) correction for multiple comparisons (i.e., time windows), post-hoc t tests for pairwise comparisons were performed within these time windows.

Data analysis: time–frequency representations

In order to obtain time–frequency representations (TFRs), clean data were re-referenced to average reference. This is important for comparability with source analysis since the forward model needs a common average reference as well. For power estimates of non-phase-locked oscillations, Morlet wavelets were used on single trial data in 20-ms steps from -700 to 2100 ms with a frequency-specific window width (linearly increasing from 2 to 12 cycles for frequencies logarithmically-spaced from 3 to 30 Hz). Single trials were subsequently baseline-corrected (against the mean of a -500 to 0 ms pre-stimulus window of all trials) and submitted to a multi-level or ‘random effects’ statistics approach (for application to time–frequency data see e.g., [Obleser et al., 2012](#); [Henry and Obleser, 2012](#)). On the first or individual level, massed independent samples regression coefficient t tests with condition as dependent variable and contrast weights as independent variable (chosen correspondently to our effects of interest, see below) were calculated. Uncorrected regression t values and β values were obtained for all time–frequency bins. According to our hypotheses, our effects of interest were a ‘wordness’ effect, namely a linear trend [pseudo > ambiguous > real], but also a stimulus-specific or ‘ambiguity’ effect [ambiguous > (pseudo, real)].

On the second or group level, the β values were tested against zero in a one tailed dependent sample t test. A Monte–Carlo non-parametrical permutation method (1000 randomizations) as implemented in the Fieldtrip toolbox estimated type I-error controlled cluster significance probabilities ($\alpha < 0.05$).

To evaluate the influence of baseline correction, we repeated first and second level statistics on absolute power estimates (skipping single trial baseline correction).

Source localization of time–frequency effects

Source localization for resulting clusters followed the Fieldtrip protocol on source reconstruction using beamformer techniques (e.g., [Haegens et al., 2010](#); [Medendorp et al., 2007](#); [Obleser and Weisz, 2012](#); [Obleser et al., 2012](#)). In short, an adaptive spatial filter (DICS–Dynamic Imaging of Coherent Sources; [Gross et al., 2001](#)) based on the cross-spectral density matrix was built by estimating the single trial fast Fourier transformation of time windows and smoothed frequencies of interest (TOI and FOI) using a set of Slepian Tapers ([Mitra and Pesaran, 1999](#)). TOI and FOI were determined according to cluster results in sensor space but computational considerations were also taken into account (more time and frequency smoothing allows better spatial estimation). For theta, estimates were centered around 4.5 Hz (± 2.5 Hz smoothing) and covered a 700 ms time window from 500 to 1200 ms, thus, three theta cycles and three tapers were used. For alpha (10 Hz ± 2 Hz smoothing), a 700 ms time window was defined centered around 1000 ms, which covers approximately seven alpha cycles and results in two tapers.

For source localization, the individual EEG electrode locations were first co-registered to the surface of a standard MRI template (by applying rigid-body transformations using the `ft_electroderealign` function). By co-registering to this template, a realistically shaped three-layer boundary elements model (BEM) provided by the Fieldtrip toolbox ([Oostenveld et al., 2003](#)) based on the same template was used. We were then able to calculate individualized forward models (i.e., lead fields) based on individual electrode positions and a standard head model for a grid with 1 cm resolution. Using the cross-spectral density matrices and the individual lead fields, a spatial filter was constructed for each grid point, and the spatial distribution of power was estimated for each condition in each subject. A common filter was constructed from all baseline and post-trigger segments (i.e., based on the cross-spectral density matrices of the combined conditions). Subject- and condition-specific solutions were spatially normalized to MNI space and averaged across subjects, and then displayed on an MNI template

(using SPM8). [Fig. 2](#) (column 4) shows the result of cluster-based statistical tests (essentially the same tests as used for the electrode-level data before) that yielded voxel clusters for covariation of source power with the alpha and theta effect, respectively. This was mainly done for illustration purposes, and unlike the tests for channel–time–frequency clusters in sensor space, no strict cluster-level thresholding was applied. We plotted t values on a standard MR template, and MNI coordinates mentioned in [Fig. 2](#) column 5 refer to brain structures that showed local maxima of activation.

In order to visualize the specificity of the neural networks for either alpha or theta frequency range oscillations, we calculated an index using the t values of the wordness t_{α} and the ambiguity t_{θ} -effect and divided their difference by their sum:

$$i_{\text{cst}} = \frac{t_{\theta} - t_{\alpha}}{|t_{\theta}| + |t_{\alpha}|} \quad (1)$$

The index has been calculated only for those grid points which exceeded the critical value of $t_{19} = 1.7291$ in the source space solution. As such, only areas are highlighted which either show an alpha (blue) or theta (red) effect. This resulted in a descriptive source map as shown in [Fig. 3](#). Values around zero indicate non-dominance for either network (i.e. green in the figure).

Results

Highly accurate performance

The performance of the lexical decision task after each trial revealed high accuracy overall (>95% in each condition, see [Fig. 1B](#)). Nevertheless, an ANOVA with the three-level factor wordness was significant ($F_{2,38} = 28.54$, $p < 0.001$) with lowest accuracy for ambiguous pseudowords (ambiguous vs. real: $t_{19} = -4.16$, $p < 0.001$; ambiguous vs. pseudo: $t_{19} = -8.01$, $p < 0.001$). Highest accuracy was found for proper pseudowords (vs. real: $t_{19} = 2.18$, $p < 0.05$), indicating some confusion of ambiguous pseudowords with real words. Since the response was prompted with delay, effects on reaction time were neither expected nor found ($F_{2,38} = 1.582$, $p = 0.221$, see [Fig. 1C](#)).

Sequential effects of word-pseudoword discrimination in event-related potentials

Overall, the ERPs over midline electrodes show the typical pattern of an N1–P2 complex followed by a later N400-like deflection in all conditions (see [Fig. 1D](#)).

Binning the ERP in 50 ms time windows with 25 ms overlap and testing for condition differences (repeated measures ANOVA, threefold factor wordness) showed no differences in amplitude before 500 ms post stimulus onset: there were no differences in the N1 or P2 ($F < 1$). The repeated measures ANOVA yielded significantly different amplitudes from 0.5 to 1.2 s (mean $F_{2,38} = 13.19$, $p < 0.01$ after FDR correction). Furthermore, post-hoc t tests on the ERP amplitudes confirmed a regrouping of conditions over time: pseudowords differed from real words over the whole time course (pseudo > real from 0.5 to 1.125 s, mean $t_{19} = -4.62$, $p_{\text{mean}} < 0.01$); ambiguous stimuli initially differed from real words (ambiguous > real from 0.525 to 0.825 s, mean $t_{19} = -4.27$, $p_{\text{mean}} < 0.01$), but regrouped with real words later, differing from proper pseudowords (pseudo > ambiguous from 0.85 to 1.2 s, mean $t_{19} = 3.1$, $p_{\text{mean}} < 0.01$; [Fig. 1D](#), gray-shaded inlay).

Differential signatures of wordness in time–frequency data

As seen in the grand average TFRs in [Fig. 2](#) top row, frequencies of the theta range (3–7 Hz) were enhanced, first phase-locked to stimulus onset around 200 ms, and, with markedly decreased phase-locking, from 400 to 1000 ms after stimulus onset. In contrast, alpha power

(8–12 Hz) was suppressed during the whole time course of a trial with the lowest power around 800 ms.

For assessing relative power changes, a multi-level statistics approach was chosen as described in the **Methods** section. A linear contrast was set on the first level for testing the wordness effect [real > ambiguous > pseudo]. On the second-level, the cluster permutation test, testing the first-level betas against zero, revealed one positive cluster ($T_{\text{sum}} = 8319.8$; $p < 0.05$) covering mainly lower- and

mid-alpha frequencies (peak at 9.3 Hz and 0.88 s; **Fig. 2** bottom rows). In general broadly distributed, the cluster showed the largest statistical differences over the left frontal and right and left central electrodes (**Fig. 2** bottom rows fourth column). Extracted power values from the cluster (8–12 Hz, 0.88 ± 0.06 s) confirmed significant differences between all three conditions (post-hoc paired t tests: real vs. ambiguous: $t_{19} = 2.32$, $p < 0.05$; real vs. pseudo: $t_{19} = 4.66$, $p < 0.01$; ambiguous vs. pseudo: $t_{19} = -2.09$, $p < 0.05$). When using absolute

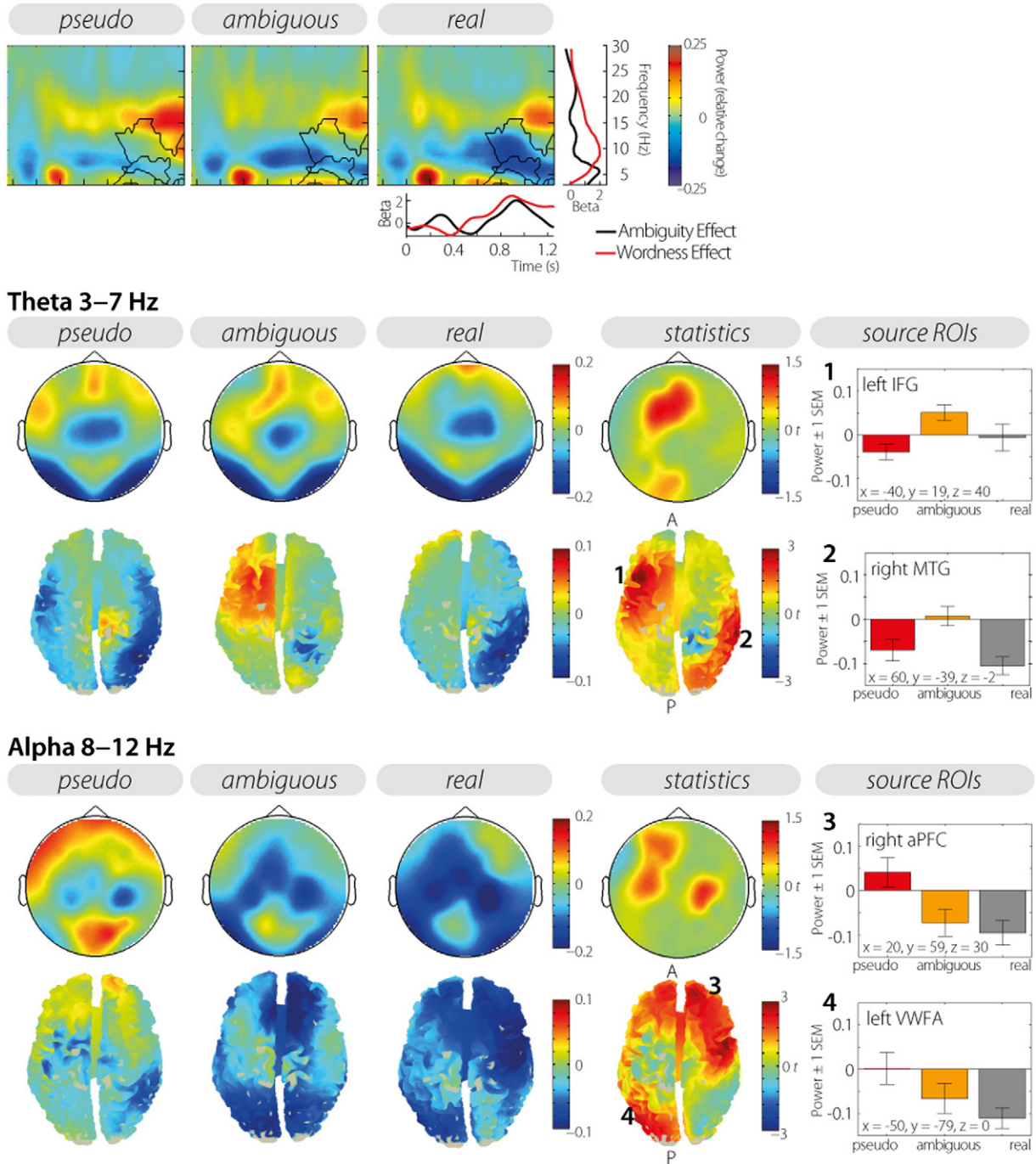


Fig. 2. Time–frequency representations in sensor and source space. Top row shows the grand average of time–frequency power changes relative to a 500 ms pre-stimulus baseline over all electrodes for the three conditions separately: from left to right for clear pseudowords, ambiguous pseudowords, and real words. Black contours mark cluster boundaries. Middle rows show scalp topographies for relative power changes in the theta band (4.5 ± 2.5 Hz, 500–1200 ms, in dependence on the time and frequency window used for the source localization) and below their correspondent source projection. Bottom rows show scalp topographies and source projections for relative alpha power changes (10 ± 2 Hz, 1000 ± 350 ms, in dependence on the time and frequency window used for the source localization). Fourth column depicts statistical differences. Fifth column are bar graphs extracted from source peaks in the left IFG and the right MTG for theta, and left VWFA and right aPFC for alpha, respectively.

power, the positive cluster ($T_{\text{sum}} = 39,928$; $p < 0.001$) showed a similar distribution over frequency and time with peak effects at 10.7 Hz and 0.9 s over left anterior electrodes.

Interestingly, testing the ambiguity effect [ambiguous > (pseudo, real)] using the same statistical approach revealed one positive cluster ($T_{\text{sum}} = 8134.6$; $p < 0.05$) in the theta frequency range (peak at 5.2 Hz, 0.94 s; Fig. 2 middle rows). Scalp topographies suggested two foci, one at the left-central anterior electrodes and the other at the parietal electrodes. Further, post-hoc paired t tests on power values extracted from the cluster (3–7 Hz, 0.88–1.1 s) confirmed that pseudowords and real words did not differ from each other ($t_{19} = 1.72$, $p < 0.1$) in the theta frequency range. Testing the absolute theta power, a comparable positive cluster was identified ($T_{\text{sum}} = 17,919$; $p < 0.01$) with the highest effect size at 5.5 Hz and 0.92 s but with a slightly shifted topography that overlaps at the left anterior electrodes but additionally emphasizes the right temporal areas.

Source localization of alpha and theta power changes

With respect to scalp topography (Fig. 2 bottom rows), alpha oscillations appeared to be distributed broadly over the scalp with a central focus and exhibited less power with increasing wordness. Following from the single conditions' source projections, source estimation of the alpha-driven wordness effect revealed peak activation in BA 9, right dorsolateral prefrontal cortex ($t_{19} = 3.04$; MNI = [10, 57, 40]). The cluster ($T_{\text{sum}} = 1152.4$; $p < 0.05$) extended into the right primary somatosensory areas (BA 3), premotor cortex (BA 6), and motor cortex (BA 4), but also into the bilateral ventral and dorsal anterior cingulate cortex (BA 24/32), and the right inferior prefrontal gyrus (BA 47), including pars triangularis (BA 45). A second peak was found in the left occipital temporal cortex ($t_{19} = 2.88$; MNI = [−50, −79, 0]) and extending into BA 37 (fusiform gyrus) and BA 20/21 (inferior and middle temporal gyrus).

For theta power changes, the spreading of power change on the scalp (Fig. 2 middle rows) suggested at least two generators: one with left frontal and one with right parietal origin, which had the highest relative power increase for ambiguous stimuli. Accordingly, two peak activations were found in one trend-level cluster ($T_{\text{sum}} = 341.9$; $p = 0.067$) for the ambiguity effect in the theta range. The first peak activation was found left anteriorly in BA 44 (pars opercularis; $t_{19} = 3.18$; MNI = [−40, 19, 40]). It extends to BA 9/46, left dorsolateral prefrontal cortex, and BA 6, premotor cortex. The second local peak was found right posteriorly in the middle temporal gyrus ($t_{19} = 3.01$; MNI = [60, −39, −2]), extending into inferior temporal gyrus (BA 20), fusiform gyrus (BA 37), supramarginal gyrus (BA 40), and posterior STG (BA 22).

Two separate networks disclosed by an alpha–theta index

Calculating the alpha–theta index as shown in Fig. 3 reveals that three of the four identified source peaks are selective for either the alpha-indexed lexical integration or the theta-indexed ambiguity resolution. Notably, the left IFG shows equally strong effects of alpha and theta activities as indicated by index values around zero.

Discussion

In order to functionally dissociate slow neural oscillations contributing to speech processing, we set up an auditory EEG study using a well-established lexical decision paradigm. Simultaneously, the data speak to theoretical controversies concerning spoken word recognition models (e.g., McClelland and Elman, 1986) by applying time–frequency analysis and revealing parallel processes of lexical integration and ambiguity resolution. Notably, alpha suppression, scaling with wordness and hence more akin to the N400, can be considered as a marker of ease in lexical integration, while theta enhancement marks the re-evaluation

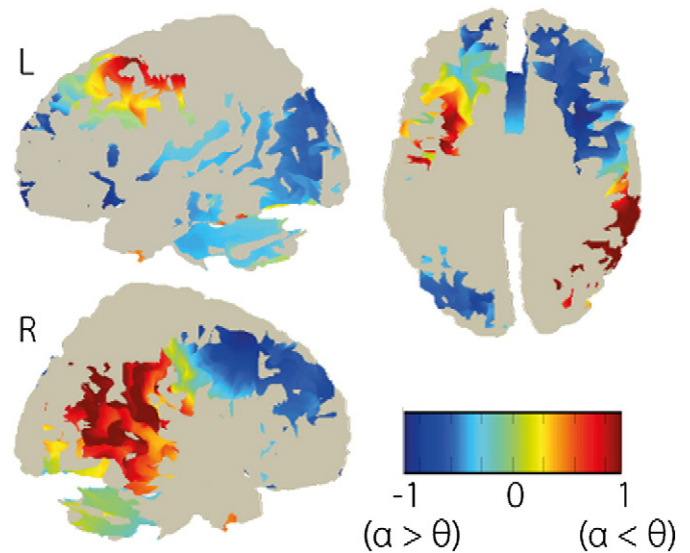


Fig. 3. Alpha–theta index. The index compares the theta effect (A) and the alpha effect (B) per source space grid point. The index has been calculated for grid points only which exceeded the critical value of $t_{19} = 1.7291$ such that only areas are highlighted which either show an alpha (blue) or theta (red) effect. Areas with index values around zero (green) show equal sensitivity to both effects, e.g., left frontal regions.

of the available sensory evidence. Generators of the alpha suppression effect were part of a left temporo–occipital and right frontal network. Oppositely, generators of the theta effect were localized in the left frontal and right middle temporal regions.

As we discuss below in further detail, the analysis of different oscillatory frequency bands disclosed the parallel maintenance of lexical and prelexical word versus pseudoword features in different brain regions and frequency ranges. To this end, time–frequency analysis is an important tool to inform discussions on sequential versus parallel processes in word recognition (e.g., Marslen-Wilson, 1987; for discussion see Norris et al., 2000).

Wordness effect in the alpha band

In line with previous findings (Obleser and Weisz, 2012), alpha power showed the greatest suppression for real words compared to the lowest suppression (or even enhancement) for clear pseudowords. Interestingly, ambiguity leads to sub-optimal lexical integration (Friedrich et al., 2006; Proverbio and Adorni, 2008) and seems to be expressed in a state of intermediate alpha power.

Two (related) theoretical framings are relevant for this effect of wordness observed in the alpha frequency range. On the one hand, it has been emphasized that parieto–occipital alpha power reflects an inhibitory mechanism, with particular relevance for working memory and selective attention tasks (Foxe et al., 1998; Klimesch et al., 2007a). On the other hand, recent findings provide more direct evidence for an influence of alpha oscillations on the timing of neural processing: Haegens et al. (2011) could show that better discrimination performance can be traced back to neuronal spiking in sensorimotor regions, which depends on the alpha rhythm not only in terms of power (firing is highest during alpha suppression) but also in terms of phase (firing is highest at the trough of a cycle; see also Spaak et al., 2012). Supporting the view put forward by Hanslmayr et al. (2012; high alpha oscillatory power mirroring reduced Shannon entropy and flow of information), Haegens et al. (2011) also found low spike-firing rates during periods of strong alpha coherence, for example, during the baseline, as opposed to the stimulus period. Both frameworks converge on predicting that low alpha power can serve as a marker of successful lexical integration.

An open issue is the potential contribution of the visual 'what'-pathway to the alpha effect observed here. Particularly the left temporo–occipital

source localization peak suggests involvement of visual fields. This might be due to the fact that we used concrete nouns, which are by definition easily imaginable in comparison to the less imaginable pseudowords (for review see Binder et al., 2009). Note, however, that in a previous fMRI study using highly similar manipulations (Raettig and Kotz, 2008), no such effects even in the contrast of concrete versus abstract nouns were found. Nevertheless, the visual word form area has been found in auditory lexical decision tasks before and has been attributed to the literacy of participants (Dehaene and Cohen, 2011; Dehaene et al., 2010). Binder et al. (2006) gathered evidence that this area is especially sensitive to sublexical bigram frequency—a pivotal element of our study design. The argument of suppressed alpha power allowing lexical integration laid out above would also hold for such a traditionally more reading-related brain area.

Ambiguity effect in the theta band

Contrary to a previous study by Obleser and Weisz (2012), theta power did not scale linearly with difficulty of word processing (if defined as difficulty of lexical access). In particular, Obleser and Weisz (2012) found higher theta power for higher intelligibility, whereas in our case theta power was highest for the ambiguous (i.e. the most difficult) case. The data provided by Obleser and Weisz (2012) suggest that sufficient spectral information is needed to enable linguistic processes or lexical evaluation, which is reflected in increasing theta power. Our data extend this view by adding ambiguity on a lexical level which requires additional lexical re-evaluation. Future research needs to clarify whether these two factors, spectral detail and lexical ambiguity, might interact. Nevertheless, both results together support our interpretation of theta oscillations subserving a language-related but task-dependent mechanism and are in line with previous studies associating theta enhancement with lexico-semantic processing (Bastiaansen et al., 2008; Hagoort et al., 2004; Hald et al., 2006; Peña and Melloni, 2012).

Interestingly, a recent opinion paper by Roux and Uhlhaas (2014) suggests that theta oscillations may be involved in the phonological loop (Baddeley, 2003). The link to the phonological loop as a concept of linguistic short-term memory speaks in favor of our interpretation where lexical re-evaluation is achieved by replay of sensory evidence (Fuentemilla et al., 2010).

Furthermore, increased prefrontal theta power has been found in response to other types of ambiguous stimuli as well, and therefore might not be tied to the language domain. Specifically, increased mid-frontal theta activity has been reported in studies investigating the ambiguity induced response conflict (Cavanagh et al., 2009; Cohen and Donner, 2013; Hanslmayr et al., 2008) and episodic memory retrieval (Ferreira et al., 2014; Staudigl et al., 2010). Although these studies differ markedly with regard to several aspects, they all share the need for processing an ambiguous stimulus. It thus appears possible that enhanced theta oscillations during ambiguous word processing reflects enhanced conflict monitoring due to the co-activated real word ('Banane' co-activates 'Banane').

We localized the enhanced theta activity in a bilateral fronto-temporal network with peak activity in the left inferior frontal gyrus (IFG, BA 44) and the right middle temporal gyrus (MTG). Their contributions, though, to the proposed interpretation of replay need to remain speculative. Instructively, a right hemispheric advantage in tracking spectral information has been shown (Obleser et al., 2008; Scott et al., 2009; Zatorre and Belin, 2001; for review see Price, 2012) which converges with the fact that vowel differences (our crucial manipulation) are primarily spectral differences. More specifically, Carreiras and Price (2008) found in accordance with our results increased activation of right hemispheric areas when manipulating vowels. Combining both ideas, Zaehle et al. (2008) could show that the analysis of prelexical segments with respect to their spectral characteristics involved bilateral MTG activation.

The left IFG, however, has been associated with a variety of linguistic processes (see Binder et al., 2009 for a meta-analysis). The unfortunate vagueness of EEG source localization limits functional dissociations which have been assigned to different subregions of the left IFG. Still, left IFG as a whole plays a role when monitoring auditory input (e.g., Giraud et al., 2004; Obleser and Weisz, 2012; Zatorre et al., 1996). Other terms such as 'auditory search', 'auditory attention', or 'auditory short-term memory' have been used to describe this function. This speaks in favor of our interpretation of auditory re-evaluation.

One might argue that our task was too easy to require top-down or re-evaluative processes. This relates to the ongoing psycholinguistic discussion whether replay or any feedback loop is really necessary in word recognition (McClelland et al., 2006; Norris et al., 2000). Since our stimuli were not phonetically ambiguous (see description in the 'Stimuli' section), no perceptual confusion occurred which would have required replay (Frauenfelder et al., 1990; Ganong, 1980; Newman et al., 1997; Norris, 2006; Wurm and Samuel, 1997). However, stimuli were lexically ambiguous which led to decisional conflicts and required ambiguity resolution processes. Recall that we introduced manipulations not before the second syllable. The third (and final) syllable, however, either continued the wordness violation (clear pseudoword) or created a lexically ambiguous case by resuming to the initially pre-activated cohort (ambiguity). Mattys (1997) summarizes evidence that retrograde information, i.e. provided after the deviation point, can influence the decision on the identity of a stimulus. This may increase reaction times (Goodman and Huttenlocher, 1988; Taft and Hambly, 1986), implying some re-evaluative processes. We therefore suggest that prelexical information were maintained and replayed in order to resolve decisional ambiguity.

In sum, we argue for a theta-tuned network which is co-activating the left IFG and the right MTG in order to replay lexico-semantic information for task-relevant ambiguity resolution.

Relationship of evoked potentials and induced oscillations

So far, studies analyzing the ERP have related the N400 to effortful processing, for example when mapping the phonological form and meaning of pseudowords, compared to real words, onto a stored representation in the mental lexicon (Friedrich et al., 2006). Recent accounts more rooted in the predictive coding framework of cortical functional organization (e.g., Summerfield and Egner, 2009) may describe the N400 as a marker of the mismatch between what is predicted and what is perceived (Lau et al., 2009, 2013). While we cannot distinguish between these explanations in a context-free setup using single word stimuli, our data more importantly show parallels in the pattern of the N400 changes over midline electrodes and the pattern of alpha oscillatory changes. Contrary to the effort- or predictive coding-hypothesis, the inhibition theory for alpha oscillations would then imply that lexical processing takes place for real words, and must be inhibited for pseudowords.

Notably, only analyzing the ERP would have led to the view that lexico-semantic integration in ambiguous pseudowords can be accomplished in the same way as their real word analogs. The regrouping of N400 deflections over time would have suggested a sequential change in processing strategy: first, ambiguous stimuli were analyzed in the same way as proper pseudowords, but from 850 ms onwards no difference between ambiguous and real word stimuli was discernible. Thus, the conclusion derived from ERPs only would have been a sequential process of lexical access. Such time-frequency decompositions of the ERP as demonstrated here may help in the future to resolve inconsistencies in the N400 literature and its generating brain structures (Halgren et al., 2002; Khateb et al., 2010).

By looking additionally at oscillatory activity, which arguably constitutes the ERP activity to large extents (Hanslmayr et al., 2007; Klimesch et al., 2007b; Makeig et al., 2004; Mazaheri and Jensen, 2008; Min et al.,

2007), parallel neural processes become discernible. The data suggest a combination of lexical integration and ambiguity resolution processes: wordness violations are detected (N400) and maintained (alpha power), but also re-evaluated retrieving stimulus-specific information (i.e., enhanced power of theta oscillations for ambiguous stimuli).

Conclusion

Time–frequency decomposition functionally separates parallel contributions of theta and alpha oscillations to speech processing, thereby fruitfully extending current frameworks based on evoked potentials. The data presented here provide evidence that lexical as well as prelexical information are maintained in spoken word recognition. The observed specificity, with theta bearing relevance to stimulus-specific, lexico-semantic processes and alpha reflecting more general inhibitory processes (thereby gating lexical integration), is a promising starting point for future studies on speech comprehension in more demanding circumstances such as peripheral hearing loss and/or noisy environments. The data furthermore shed light onto the neural bases of the lexical decision task that has been in use for decades. In sum, this approach allows for a refinement of neural models describing the complex nature of spoken word recognition.

References

- Baddeley, A., 2003. Working memory and language: an overview. *J. Commun. Disord.* 36 (3), 189–208.
- Bastiaansen, M.C.M., Oostenveld, R., Jensen, O., Hagoort, P., 2008. I see what you mean: theta power increases are involved in the retrieval of lexical semantic information. *Brain Lang.* 106 (1), 15–28.
- Binder, J.R., Medler, D.A., Westbury, C.F., Liebenthal, E., Buchanan, L., 2006. Tuning of the human left fusiform gyrus to sublexical orthographic structure. *NeuroImage* 33 (2), 739–748.
- Binder, J.R., Desai, R.H., Graves, W.W., Conant, L.L., 2009. Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cereb. Cortex* 19 (12), 2767–2796.
- Carreiras, M., Price, C.J., 2008. Brain activation for consonants and vowels. *Cereb. Cortex* 18 (7), 1727–1735.
- Cavanagh, J.F., Cohen, M.X., Allen, J.J.B., 2009. Prelude to and resolution of an error: EEG phase synchrony reveals cognitive control dynamics during action monitoring. *J. Neurosci.* 29 (1), 98–105.
- Chwilla, D.J., Brown, C.M., Hagoort, P., 1995. The N400 as a function of the level of processing. *Psychophysiology* 32 (3), 274–285.
- Cohen, M.X., Donner, T.H., 2013. Midfrontal conflict-related theta-band power reflects neural oscillations that predict behavior. *J. Neurophysiol.* 110 (12), 2752–2763.
- Cohen, M.X., Elger, C.E., Fell, J., 2009. Oscillatory activity and phase-amplitude coupling in the human medial frontal cortex during decision making. *J. Cogn. Neurosci.* 21 (2), 390–402.
- Connine, C.M., Clifton, C., 1987. Interactive use of lexical information in speech perception. *J. Exp. Psychol. Hum. Percept. Perform.* 13 (2), 291–299.
- Connolly, J.F., Phillips, N.A., 1994. Event-related potential components reflect phonological and semantic processing of the terminal word of spoken sentences. *J. Cogn. Neurosci.* 6 (3), 256–266.
- Debener, S., Thorne, J., Schneider, T.R., Viola, F.C., 2010. Using ICA for the Analysis of Multi-channel EEG Data. Oxford University Press pp. 121–134.
- Dehaene, S., Cohen, L., 2011. The unique role of the visual word form area in reading. *Trends Cogn. Sci.* 15 (6), 254–262.
- Dehaene, S., Pegado, F., Braga, L.W., Ventura, P., Nunes Filho, G., Jobert, A., Dehaene-Lambert, G., Kolinsky, R., Morais, J., Cohen, L., 2010. How learning to read changes the cortical networks for vision and language. *Science* 330 (6009), 1359–1364.
- Desroches, A.S., Newman, R.L., Joannisse, M.F., 2009. Investigating the time course of spoken word recognition: electrophysiological evidence for the influences of phonological similarity. *J. Cogn. Neurosci.* 21 (10), 1893–1906.
- Dufour, S., Brunellière, A., Frauenfelder, U.H., 2013. Tracking the time course of word-frequency effects in auditory word recognition with event-related potentials. *Cogn. Sci.* 37 (3), 489–507.
- Ferreira, C.S., Marful, A., Staudigl, T., Bajo, T., Hanslmayr, S., 2014. Medial prefrontal theta oscillations track the time course of interference during selective memory retrieval. *J. Cogn. Neurosci.* 26 (4), 777–791.
- Foxe, J.J., Simpson, G.V., Ahlfors, S.P., 1998. Parieto-occipital approximately 10 Hz activity reflects anticipatory state of visual attention mechanisms. *Neuroreport* 9 (17), 3929–3933.
- Frauenfelder, U.H., Segui, J., Dijkstra, T., 1990. Lexical effects in phonemic processing: facilitatory or inhibitory? *J. Exp. Psychol. Hum. Percept. Perform.* 16 (1), 77–91.
- Friederici, A.D., 1997. Neurophysiological aspects of language processing. *Clin. Neurosci.* 4 (2), 64–72.
- Friedrich, C.K., Eulitz, C., Lahiri, A., 2006. Not every pseudoword disrupts word recognition: an ERP study. *Behav. Brain Funct.* 2, 36.
- Friedrich, C.K., Schild, U., Röder, B., 2009. Electrophysiological indices of word fragment priming allow characterizing neural stages of speech recognition. *Biol. Psychol.* 80 (1), 105–113.
- Fuentemilla, L., Penny, W.D., Cashdollar, N., Brunzeck, N., Düzel, E., 2010. Theta-coupled periodic replay in working memory. *Curr. Biol.* 20 (7), 606–612.
- Ganong, W.F., 1980. Phonetic categorization in auditory word perception. *J. Exp. Psychol. Hum. Percept. Perform.* 6 (1), 110–125.
- Ghitza, O., 2011. Linking speech perception and neurophysiology: speech decoding guided by cascaded oscillators locked to the input rhythm. *Front. Psychol.* 2, 130.
- Giraud, A.-L., Poeppel, D., 2012. Cortical oscillations and speech processing: emerging computational principles and operations. *Nat. Neurosci.* 15 (4), 511–517.
- Giraud, A.-L., Kell, C., Thierfelder, C., Sterzer, P., Russ, M.O., Preibisch, C., Kleinschmidt, A., 2004. Contributions of sensory input, auditory search and verbal comprehension to cortical activity during speech processing. *Cereb. Cortex* 14 (3), 247–255.
- Goodman, J.C., Huttenlocher, J., 1988. Do we know how people identify spoken words? *J. Mem. Lang.* 27 (6), 684–698.
- Gross, J., Kujala, J., Hamalainen, M., Timmermann, L., Schnitzler, A., Salmelin, R., 2001. Dynamic imaging of coherent sources: studying neural interactions in the human brain. *Proc. Natl. Acad. Sci. U. S. A.* 98 (2), 694–699.
- Haegens, S., Nacher, V., Luna, R., Romo, R., Jensen, O., 2011. α -Oscillations in the monkey sensorimotor network influence discrimination performance by rhythmical inhibition of neuronal spiking. *Proc. Natl. Acad. Sci. U. S. A.* 108 (48), 19377–19382.
- Haegens, S., Osipova, D., Oostenveld, R., Jensen, O., 2010. Somatosensory working memory performance in humans depends on both engagement and disengagement of regions in a distributed network. *Hum. Brain Mapp.* 31 (1), 26–35.
- Hagoort, P., Hald, L., Bastiaansen, M., Petersson, K.M., 2004. Integration of word meaning and world knowledge in language comprehension. *Science* 304 (5669), 438–441.
- Hald, L.A., Bastiaansen, M.C.M., Hagoort, P., 2006. EEG theta and gamma responses to semantic violations in online sentence processing. *Brain Lang.* 96 (1), 90–105.
- Halgren, E., Dhond, R.P., Christensen, N., Van Petten, C., Marinkovic, K., Lewine, J.D., Dale, A.M., 2002. N400-like magnetoencephalography responses modulated by semantic context, word frequency, and lexical class in sentences. *NeuroImage* 17 (3), 1101–1116.
- Hanslmayr, S., Klimesch, W., Sauseng, P., Gruber, W., Doppelmayr, M., Freunberger, R., Pecherstorfer, T., Birbaumer, N., 2007. Alpha phase reset contributes to the generation of ERPs. *Cereb. Cortex* 17 (1), 1–8.
- Hanslmayr, S., Pastötter, B., Bäuml, K.-H., Gruber, S., Wimber, M., Klimesch, W., 2008. The electrophysiological dynamics of interference during the Stroop task. *J. Cogn. Neurosci.* 20 (2), 215–225.
- Hanslmayr, S., Spitzer, B., Bäuml, K.-H., 2009. Brain oscillations dissociate between semantic and nonsemantic encoding of episodic memories. *Cereb. Cortex* 19 (7), 1631–1640.
- Hanslmayr, S., Staudigl, T., Fellner, M.-C., 2012. Oscillatory power decreases and long-term memory: the information via desynchronization hypothesis. *Front. Hum. Neurosci.* 6, 74.
- Hartmann, T., Schlee, W., Weisz, N., 2012. It's only in your head: expectancy of aversive auditory stimulation modulates stimulus-induced auditory cortical alpha desynchronization. *NeuroImage* 60 (1), 170–178.
- Henry, M.J., Obleser, J., 2012. Frequency modulation entrains slow neural oscillations and optimizes human listening behavior. *Proc. Natl. Acad. Sci. U. S. A.* 109 (49), 20095–20100.
- Jensen, O., Mazaheri, A., 2010. Shaping functional architecture by oscillatory alpha activity: gating by inhibition. *Front. Hum. Neurosci.* 4, 186.
- Khateb, A., Pegna, A.J., Landis, T., Mouthon, M.S., Annoni, J.-M., 2010. On the origin of the N400 effects: an ERP waveform and source localization analysis in three matching tasks. *Brain Topogr.* 23, 311–320.
- Klimesch, W., 1999. EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain Res. Brain Res. Rev.* 29 (2–3), 169–195.
- Klimesch, W., 2012. Alpha-band oscillations, attention, and controlled access to stored information. *Trends Cogn. Sci.* 16 (12), 606–617.
- Klimesch, W., Sauseng, P., Hanslmayr, S., 2007a. EEG alpha oscillations: the inhibition-timing hypothesis. *Brain Res. Rev.* 53 (1), 63–88.
- Klimesch, W., Sauseng, P., Hanslmayr, S., Gruber, W., Freunberger, R., 2007b. Event-related phase reorganization may explain evoked neural dynamics. *Neurosci. Biobehav. Rev.* 31 (7), 1003–1016.
- Kotz, S.A., Cappa, S.F., von Cramon, D.Y., Friederici, A.D., 2002. Modulation of the lexical-semantic network by auditory semantic priming: an event-related functional MRI study. *NeuroImage* 17 (4), 1761–1772.
- Kutas, M., Hillyard, S.A., 1980. Reading senseless sentences: brain potentials reflect semantic incongruity. *Science* 207 (4427), 203–205.
- Laszlo, S., Stites, M., Federmeier, K.D., 2012. Won't get fooled again: an event-related potential study of task and repetition effects on the semantic processing of items without semantics. *Lang. Cogn. Process.* 27 (2), 257–274.
- Lau, E., Almeida, D., Hines, P.C., Poeppel, D., 2009. A lexical basis for N400 context effects: evidence from MEG. *Brain Lang.* 111 (3), 161–172.
- Lau, E.F., Holcomb, P.J., Kuperberg, G.R., 2013. Dissociating N400 effects of prediction from association in single-word contexts. *J. Cogn. Neurosci.* 25 (3), 484–502.
- Lehtelä, L., Salmelin, R., Hari, R., 1997. Evidence for reactive magnetic 10-Hz rhythm in the human auditory cortex. *Neurosci. Lett.* 222 (2), 111–114.
- Lisman, J.E., Jensen, O., 2013. The θ - γ neural code. *Neuron* 77 (6), 1002–1016.
- Makeig, S., Debener, S., Onton, J., Delorme, A., 2004. Mining event-related brain dynamics. *Trends Cogn. Sci.* 8 (5), 204–210.
- Marslen-Wilson, W.D., 1987. Functional parallelism in spoken word-recognition. *Cognition* 25 (1–2), 71–102.
- Mattys, S., 1997. The use of time during lexical processing and segmentation: a review. *Psychon. Bull. Rev.* 4 (3), 310–329.

- Mazaheri, A., Jensen, O., 2008. Asymmetric amplitude modulations of brain oscillations generate slow evoked responses. *J. Neurosci.* 28 (31), 7781–7787.
- McClelland, J.L., Elman, J.L., 1986. The TRACE model of speech perception. *Cogn. Psychol.* 18 (1), 1–86.
- McClelland, J.L., Mirman, D., Holt, L.L., 2006. Are there interactive processes in speech perception? *Trends Cogn. Sci.* 10 (8), 363–369.
- Medendorp, W.P., Kramer, G.F.L., Jensen, O., Oostenveld, R., Schoffelen, J.-M., Fries, P., 2007. Oscillatory activity in human parietal and occipital cortex shows hemispheric lateralization and memory effects in a delayed double-step saccade task. *Cereb. Cortex* 17 (10), 2364–2374.
- Meyer, L., Obleser, J., Friederici, A.D., 2013. Left parietal alpha enhancement during working memory-intensive sentence processing. *Cortex* 49 (3), 711–721.
- Min, B.-K., Busch, N.A., Debener, S., Kranczioch, C., Hanslmayr, S., Engel, A.K., Herrmann, C.S., 2007. The best of both worlds: phase-reset of human EEG alpha activity and additive power contribute to ERP generation. *Int. J. Psychophysiol.* 65 (1), 58–68.
- Minicucci, D., Guediche, S., Blumstein, S.E., 2013. An fMRI examination of the effects of acoustic-phonetic and lexical competition on access to the lexical-semantic network. *Neuropsychologia* 51 (10), 1980–1988.
- Mitra, P.P., Pesaran, B., 1999. Analysis of dynamic brain imaging data. *Biophys. J.* 76 (2), 691–708.
- Newman, R.S., Sawusch, J.R., Luce, P.A., 1997. Lexical neighborhood effects in phonetic processing. *J. Exp. Psychol. Hum. Percept. Perform.* 23 (3), 873–889.
- Norris, D., 2006. The Bayesian reader: explaining word recognition as an optimal Bayesian decision process. *Psychol. Rev.* 113 (2), 327–357.
- Norris, D., McQueen, J.M., Cutler, A., 2000. Merging information in speech recognition: feedback is never necessary. *Behav. Brain Sci.* 23 (3), 299–325 (discussion 325–370).
- Obleser, J., Kotz, S.A., 2011. Multiple brain signatures of integration in the comprehension of degraded speech. *NeuroImage* 55 (2), 713–723.
- Obleser, J., Weisz, N., 2012. Suppressed alpha oscillations predict intelligibility of speech and its acoustic details. *Cereb. Cortex* 22 (11), 2466–2477.
- Obleser, J., Elbert, T., Lahiri, A., Eulitz, C., 2003. Cortical representation of vowels reflects acoustic dissimilarity determined by formant frequencies. *Brain Res. Cogn. Brain Res.* 15 (3), 207–213.
- Obleser, J., Eisner, F., Kotz, S.A., 2008. Bilateral speech comprehension reflects differential sensitivity to spectral and temporal features. *J. Neurosci.* 28 (32), 8116–8123.
- Obleser, J., Wöstmann, M., Hellbernd, N., Wilsch, A., Maess, B., 2012. Adverse listening conditions and memory load drive a common alpha oscillatory network. *J. Neurosci.* 32 (36), 12376–12383.
- Oostenveld, R., Praamstra, P., 2001. The five percent electrode system for high-resolution EEG and ERP measurements. *Clin. Neurophysiol.* 112 (4), 713–719.
- Oostenveld, R., Stegeman, D.F., Praamstra, P., van Oosterom, A., 2003. Brain symmetry and topographic analysis of lateralized event-related potentials. *Clin. Neurophysiol.* 114 (7), 1194–1202.
- Oostenveld, R., Fries, P., Maris, E., Schoffelen, J.-M., 2011. FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput. Intell. Neurosci.* 2011, 156869.
- Peelle, J.E., Davis, M.H., 2012. Neural oscillations carry speech rhythm through to comprehension. *Front. Psychol.* 3, 320.
- Peña, M., Melloni, L., 2012. Brain oscillations during spoken sentence processing. *J. Cogn. Neurosci.* 24 (5), 1149–1164.
- Price, C.J., 2012. A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. *NeuroImage* 62 (2), 816–847.
- Proverbio, A.M., Adorni, R., 2008. Orthographic familiarity, phonological legality and number of orthographic neighbours affect the onset of ERP lexical effects. *Behav. Brain Funct.* 4, 27.
- Raettig, T., Kotz, S.A., 2008. Auditory processing of different types of pseudo-words: an event-related fMRI study. *NeuroImage* 39 (3), 1420–1428.
- Roux, F., Uhlhaas, P.J., 2014. Working memory and neural oscillations: alpha-gamma versus theta-gamma codes for distinct WM information? *Trends Cogn. Sci.* 18 (1), 16–25.
- Roux, F., Wibral, M., Singer, W., Aru, J., Uhlhaas, P.J., 2013. The phase of thalamic alpha activity modulates cortical gamma-band activity: evidence from resting-state MEG recordings. *J. Neurosci.* 33 (45), 17827–17835.
- Rugg, M.D., 1990. Event-related brain potentials dissociate repetition effects of high- and low-frequency words. *Mem. Cogn.* 18 (4), 367–379.
- Scott, S.K., Rosen, S., Beaman, C.P., Davis, J.P., Wise, R.J.S., 2009. The neural processing of masked speech: evidence for different mechanisms in the left and right temporal lobes. *J. Acoust. Soc. Am.* 125 (3), 1737–1743.
- Spaak, E., Bonnefond, M., Maier, A., Leopold, D.A., Jensen, O., 2012. Layer-specific entrainment of gamma-band neural activity by the alpha rhythm in monkey visual cortex. *Curr. Biol.* 22 (24), 2313–2318.
- Staudigl, T., Hanslmayr, S., Bäuml, K.-H.T., 2010. Theta oscillations reflect the dynamics of interference in episodic memory retrieval. *J. Neurosci.* 30 (34), 11356–11362.
- Strauß, A., Kotz, S.A., Obleser, J., 2013. Narrowed expectancies under degraded speech: revisiting the N400. *J. Cogn. Neurosci.* 25 (8), 1383–1395.
- Summerfield, C., Egner, T., 2009. Expectation (and attention) in visual cognition. *Trends Cogn. Sci.* 13 (9), 403–409.
- Taft, M., Hambly, G., 1986. Exploring the cohort model of spoken word recognition. *Cognition* 22 (3), 259–282.
- Tavabi, K., Embick, D., Roberts, T.P.L., 2011. Word repetition priming-induced oscillations in auditory cortex: a magnetoencephalography study. *Neuroreport* 22 (17), 887–891.
- Van Petten, C., Kutas, M., 1990. Interactions between sentence context and word frequency in event-related brain potentials. *Mem. Cogn.* 18 (4), 380–393.
- Van Petten, C., Luka, B.J., 2012. Prediction during language comprehension: benefits, costs, and ERP components. *Int. J. Psychophysiol.* 83 (2), 176–190.
- Weisz, N., Hartmann, T., Müller, N., Lorenz, I., Obleser, J., 2011. Alpha rhythms in audition: cognitive and clinical perspectives. *Front. Psychol.* 2, 73.
- Wilsch, A., Henry, M.J., Herrmann, B., Maess, B., Obleser, J., 2014. Alpha oscillatory dynamics index temporal expectation benefits in working memory. *Cereb. Cortex*.
- Wurm, L.H., Samuel, A.G., 1997. Lexical inhibition and attentional allocation during speech perception: evidence from phoneme monitoring. *J. Mem. Lang.* 36, 165–187.
- Zaehle, T., Geiser, E., Alter, K., Jancke, L., Meyer, M., 2008. Segmental processing in the human auditory dorsal stream. *Brain Res.* 1220, 179–190.
- Zatorre, R.J., Belin, P., 2001. Spectral and temporal processing in human auditory cortex. *Cereb. Cortex* 11 (10), 946–953.
- Zatorre, R.J., Meyer, E., Gjedde, A., Evans, A.C., 1996. PET studies of phonetic processing of speech: review, replication, and reanalysis. *Cereb. Cortex* 6 (1), 21–30.