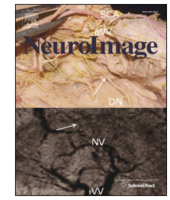
NeuroImage xxx (2014) xxx–xxx

Contents lists available at ScienceDirect

NeuroImage

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



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

Antje Strauβ [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, Stephanstraβe 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 xxxx

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.

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, 2007b): 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.,

* Corresponding author.
E-mail address: strauss@cbs.mpg.de (A. Strauβ).

http://dx.doi.org/10.1016/j.neuroimage.2014.04.005
1053-8119/© 2014 Elsevier Inc. All rights reserved.

Lisman and Jensen, 2013; Roux and Uhlhaas, 2014), and matching of 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; Strauβ 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 ± 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 → Kommide). The majority (approximately one third) of vowel pairs were 600 to 1000 Hz apart from each other (/ / → /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 offsets, and RMS normalization. In sum, the experimental corpus consisted of 240 stimuli with a mean length of 754.2 ms ± 83.5 ms (M ± 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.

Please cite this article as: Strauβ, A., et al., Alpha and theta brain oscillations index dissociable processes in spoken word recognition, NeuroImage (2014), http://dx.doi.org/10.1016/j.neuroimage.2014.04.005