EJN European Journal of Neuroscience FENS Federation of European Neuroscience Societies

European Journal of Neuroscience, pp. 1–13, 2017 doi:10.1111/ejn.13748

The neural oscillations of speech processing and language comprehension: state of the art and emerging mechanisms

Lars Meyer

Department of Neuropsychology, Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstraße 1A, 04103 Leipzig, Germany

Keywords: chunking, entrainment, memory, predictive coding

Abstract

Neural oscillations subserve a broad range of functions in speech processing and language comprehension. On the one hand, speech contains—somewhat—repetitive trains of air pressure bursts that occur at three dominant amplitude modulation frequencies, physically marking the linguistically meaningful progressions of phonemes, syllables and intonational phrase boundaries. To these acoustic events, neural oscillations of isomorphous operating frequencies are thought to synchronise, presumably resulting in an implicit temporal alignment of periods of neural excitability to linguistically meaningful spectral information on the three low-level linguistic description levels. On the other hand, speech is a carrier signal that codes for high-level linguistic meaning, such as syntactic structure and semantic information—which cannot be read from stimulus acoustics, but must be acquired during language acquisition and decoded for language comprehension. Neural oscillations subserve the processing of both syntactic structure and semantic information. Here, I synthesise a mapping from each linguistic processing domain to a unique set of subserving oscillatory mechanisms—the mapping is plausible given the role ascribed to different oscillatory mechanisms in different subfunctions of cortical information processing and faithful to the underlying electrophysiology. In sum, the present article provides an accessible and extensive review of the functional mechanisms that neural oscillations subserve in speech processing and language comprehension.

Overview

This review summarises the current research into neural oscillations as subserving both lower-level functions of speech processing and higher-level functions of language comprehension. Throughout the article, the dichotomy of speech processing vs. language comprehension is used to capture an essential difference: speech processing describes the segmentation of speech into linguistically meaningful units, as well as the identification of these units—based on temporal and spectral cues that are present in the physical world and can be recognised by the auditory system. In contrast, language comprehension describes the decoding of the meaning of words and combinations of words, such as phrases and sentences. Meaning is not present in the physical world; rather the meaning of speech sounds must be acquired first through contextual association during language acquisition, to be recognised and composed in the course of language comprehension.

Accordingly, this review comes in two parts. The first part, on speech processing, illustrates the popular view that neural oscillations subserve the segmentation and identification of more or less discrete phonological units across a range of oscillatory bands. The operation frequency of these bands is thought to be stimulus-bound, corresponding to the occurrence frequency of phonological units in the acoustic spectrum. The second part, on language comprehension, illustrates the major functions that neural oscillations subserve in higher-level linguistic processing. Here, oscillations at the frequency where multiword groups can be decoded from speech have been proposed to functionally support the internal generation of syntactic structures. In addition, oscillations functionally subserve memory-related processes required for, and potentially specific to, the internal computation of sentence-level relational meaning, as well as the cumulative prediction and contextual interpretation of upcoming words.

It is important to note that the proposed links between specific oscillatory frequency bands and specific subdomains of speech processing and language comprehension do not entail domain- or process-specificity of particular frequency bands. The human brain's oscillatory frequency band inventory is well too limited to ascribe

Correspondence: Lars Meyer, as above.
E-mail: lmeyer@cbs.mpg.de

Received 11 July 2017, revised 14 September 2017, accepted 9 October 2017

Edited by Ali Mazaheri

Reviewed by Nicola Molinaro, Basque center on Cognition, Brain and Language, Spain; and Elana Zion-Golumbic, Bar-Ilan University, Israel

The associated peer review process communications can be found in the online version of this article.

specific oscillatory frequency bands to specific linguistic processing domains alone (for review, see Penttonen & Buzsáki, 2003; Buzsaki, 2006). Instead, I share the more parsimonious view that linguistic processing domains instantiate a small general set of basic electrophysiological mechanisms that resurface as oscillatory patterns across cognitive domains (e.g. Friederici & Singer, 2015; Lewis *et al.*, 2015). By implication, the use of neural oscillations as a dependent measure for experimental research may allow for a cognitive neuropsychology of speech and language that is not only faithful to textbook conceptualisations of linguistics, psycholinguistics and neurolinguistics, but also to the systems neuroscientific view that any cognitive function instantiates a circumscribed set of fundamental mechanisms of cortical information processing.

Neural oscillations in speech processing

Speech processing is the set of neural processes enabling the segmentation and identification of more or less discrete phonological units in the acoustic spectrum, which encode to-be-communicated meaning. Speech is built hierarchically from units decreasing in temporal granularity: a series of phonemes constitutes a syllable; a series of syllables constitutes an intonation phrase. To a certain extent, the linguistic units of phonemes, syllables and intonation phrases can be recognised from physical counterparts in speech acoustics—temporal counterparts include repetitive trains of air pressure maxima that mark the on- and offsets of many discrete phonetic units; spectral counterparts distinguish between individual phonetic units. Across a range of operating frequencies, neural oscillations have been proposed to aid speech processing by aligning their phase to acoustic amplitude extrema that occur during the on- or offsets of phonemes, syllables and intonation phrases. In the first part of this review, I summarise these proposals (see Fig. 1 for an overview).

Synchronisation with speech: phonemes, syllables and intonation phrases

During speech processing, neural oscillations have been proposed to track linguistically meaningful acoustic properties of speech across three characteristic frequency bands (for review, see Giraud & Poeppel, 2012; Kösem & Van Wassenhove, 2017). During rest already, oscillatory frequency bands resembling the paces of phonemic units are prevalent in the power spectrum of the auditory cortex (Giraud *et al.*, 2007). During speech processing, these oscillations have been found to synchronise with the pace of both acoustic amplitude modulations and phonemes (Di Liberto *et al.*, 2015), syllables (Peelle *et al.*, 2013) and intonation phrases (Bourguignon *et al.*, 2013). Oscillatory synchronisation facilitates speech processing, given that the magnitude of synchronicity predicts speech intelligibility (Ahissar *et al.*, 2001; Luo & Poeppel, 2007; Nourski *et al.*, 2009; Peelle *et al.*, 2013; Doelling *et al.*, 2014). It is unlikely that the proposed isomorphism between the frequencies of neural oscillations and speech segments of different granularities entails a functional specificity of neural oscillations for speech processing—for instance, phoneme-rate sampling is attested even in rodent models, and deficient sampling in rodent models may associate with decreased rhythmicity of spiking (Che *et al.*, 2014; Centanni *et al.*, 2016). In addition, perceptual sampling is rhythmic not only in the auditory, but also in the visual and tactile modalities (Holcombe, 2009; Giraud & Poeppel, 2012; VanRullen, 2016). It thus appears more plausible to conceive of the rhythmicity of speech at its major temporal granularities as having been shaped evolutionarily by the brain's pre-existing oscillatory frequencies.

Synchronisation of neural oscillations to speech is thought to occur via so-called entrainment, the electrophysiological underpinnings of which are not clearly defined: on the one hand, the phase of the speech amplitude envelope could attract neural phase to synchronise (i.e. phase synchronisation); on the other hand, speech

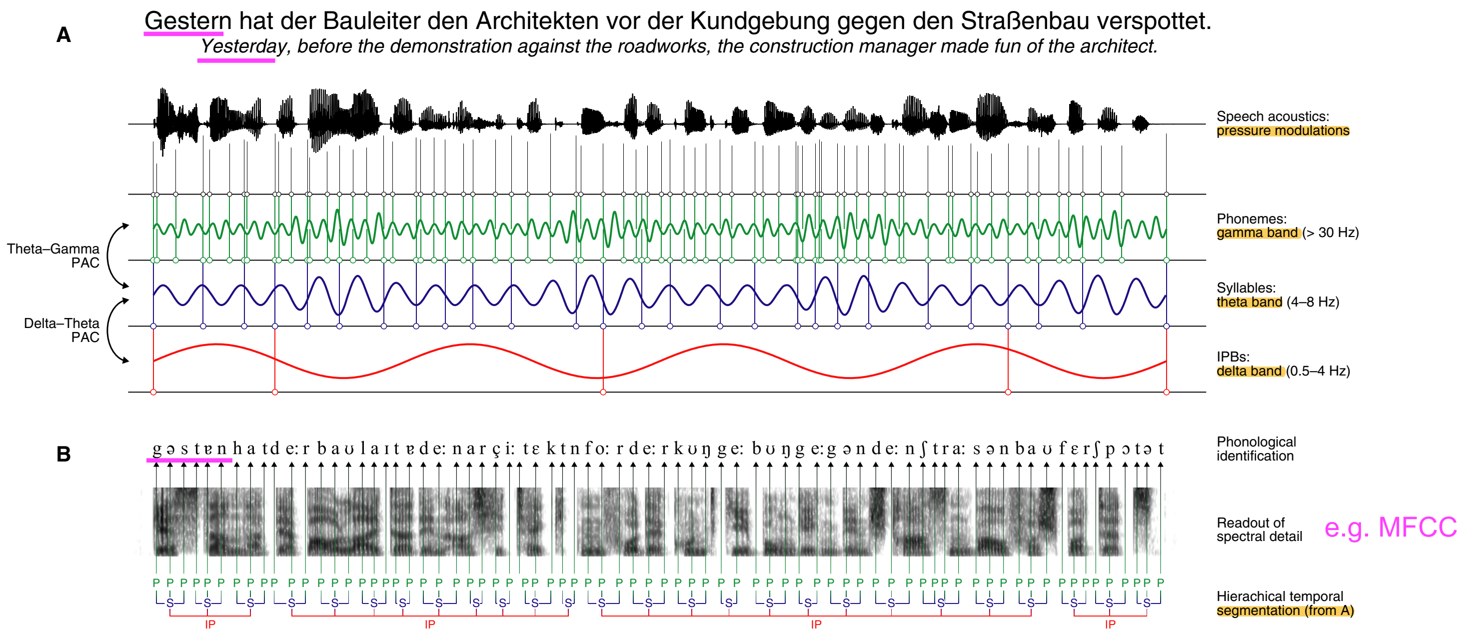

FIG. 1. Neural oscillations in speech processing; (A) gamma- (green), theta- (blue) and delta-band oscillations (red) synchronise with air pressure maxima in the acoustic speech signal that delineates phonemes (green), syllables (blue) and intonational phrase boundaries (IPB; red), respectively; delta–theta and theta–gamma phase–amplitude coupling (PAC) could serve both to (B) bind lower-level phonological units into coherent high-level percepts (i.e. phonemes into syllables, syllables into intonation phrases) and to align neural excitability to linguistically critical spectral features, thereby facilitating phonological identification in spite of sparse amplitude modulations that prohibit stimulus-driven synchronisation.

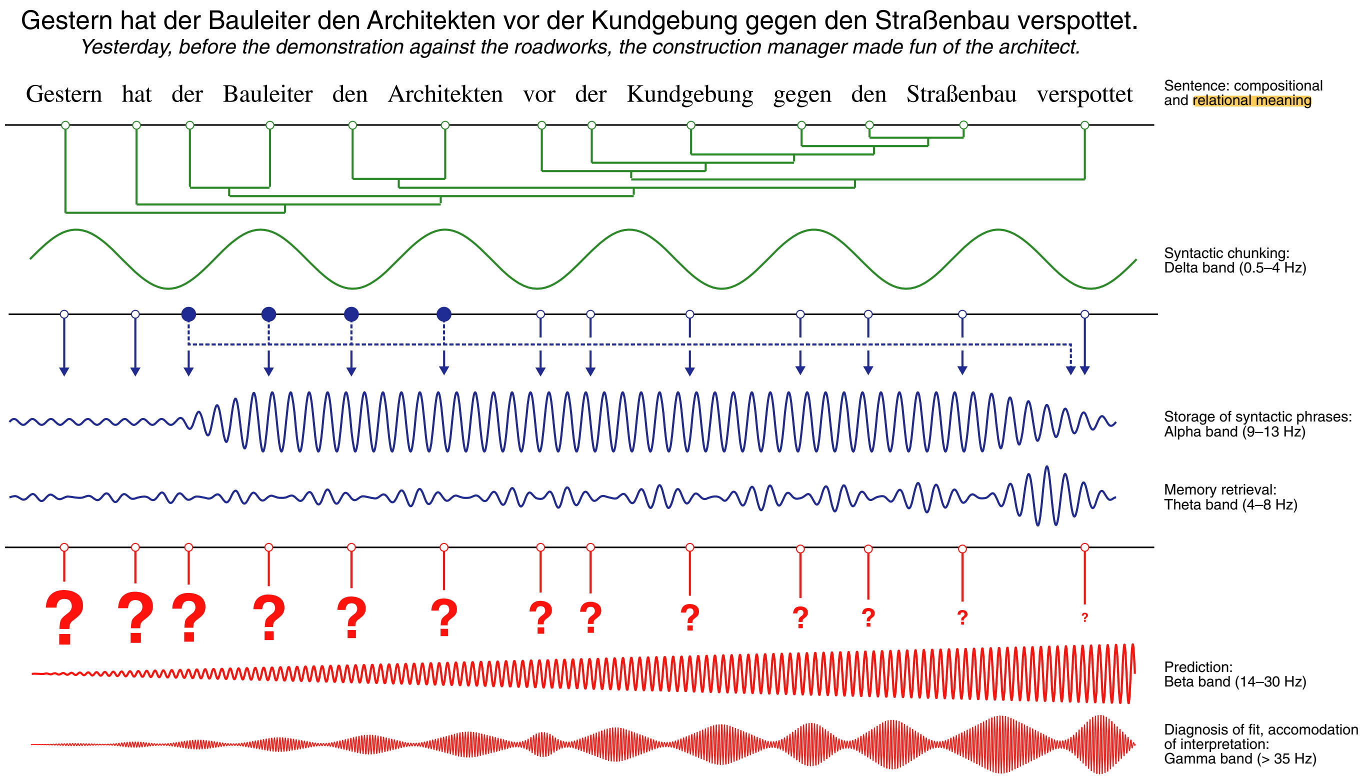

FIG. 2. Neural oscillations in language comprehension; green: chunking of words into syntactic phrases is subserved by delta-band cycles; blue: phrases are stored in working memory (solid circles) with the help of alpha-band oscillations; alpha-band amplitude increases during the storage of the subject (i.e. *der Bauleiter/the construction manager*) and object (i.e. *den Architekten/the architect*), until these can be linked to the main verb (i.e. *verspottet/made fun of*); stored information is retrieved from working memory and long-term memory for integration into the ongoing sentence-level working memory representation, with the support of theta-band oscillations; theta-band power transiently increases during the retrieval of individual words from long-term memory; burst amplitude increases along the sentence, potentially indicating facilitated integration of retrieved words into the sentence-level working memory representation is strongest during working memory retrieval of the subject and object at the sentence-final verb; red: aided by beta-band oscillations, and upcoming words are predicted from the prior cumulative semantic interpretation, based on probabilistic lexical-semantic knowledge; the stronger the predictions (end of the sentence), the higher the beta power on prediction fulfilment; gamma-band oscillations assist diagnosis of lexical-semantic contextual fit of incoming words and the accommodation of the cumulative interpretation; the better the fit of incoming words, the higher the gamma power on prediction fulfilment.

noteworthy that slow cortical potentials (i.e. < 0.5 Hz; Northoff, 2016) have previously been ascribed a domain-general function in the perception of simultaneity (cf. Stefanics *et al.*, 2010; Northoff, 2016). This allows to speculate that the delta band's pre-existing general role in cortical information chunking offered itself as a mechanism for syntactic phrase formation and, potentially, informational fusion of complex meaning by simulating perceptual simultaneity of a phrase's words (Meyer *et al.*, 2016). Hence, rather than thinking that delta-band oscillations specifically fit the necessities of language comprehension, syntactic phrases as a linguistic phenomenon should rather be viewed as an ideal means of exploiting the brain's pre-existing chunking abilities. In line with this speculation, event-related brain potentials indexing the chunking of words into syntactic phrases occur with a regular period (2.6–2.7 s; Roll *et al.*, 2012; Schremm *et al.*, 2015) that matches the canonical duration of syntactic phrases observed in linguistic corpus analyses (Vollrath *et al.*, 1992) and falls into the frequency range of either the lower delta band or slow cortical potentials. Such slow frequencies might just suffice electrophysiologically to integrate information on the time and network scales needed for the composition of complex information—the brain employs long cycles when integrating complex information across large networks, given the variance in conduction delays of large neuronal networks (Buzsaki, 2006). In line with this, slow oscillatory frequencies relate to long-lasting excitatory postsynaptic potentials in supragranular and granular cortical layers, which exhibit a high degree of cortico–cortical connectivity (Steriade *et al.*, 1993; Lakatos *et al.*, 2008; Northoff, 2016).

The alpha band: storage of syntactic phrases in verbal working memory

The storage of syntactic phrases in verbal working memory for the downstream establishment of dependencies with other phrases associates with increased alpha-band activity (Haarmann & Cameron, 2005; Weiss *et al.*, 2005; Meyer *et al.*, 2013; Bonhage *et al.*, 2017). As an example, an experiment by Meyer *et al.* (2013) manipulated the storage interval for a sentence-initial subject noun phrase that needed to be linked to a sentence-final verb (e.g. Fodor, 1978; Wanner & Maratsos, 1979; Frazier *et al.*, 1983; Kluender & Kutas, 1993; Matzke *et al.*, 2002). First, alpha-band power was observed to increase with storage demands; second, the effect was source-localised to inferior parietal cortex, which is classically related to verbal working memory (for review, see Wager & Smith, 2003; Owen *et al.*, 2005; Leff *et al.*, 2009); and third, source power correlated with verbal working memory capacity.

In general, these results extend prior evidence for a role of alpha-band oscillations in verbal working memory to the sentence comprehension domain (Krause *et al.*, 1996; Maltseva *et al.*, 2000; Jensen *et al.*, 2002; Schack *et al.*, 2005; Leiberg *et al.*, 2006; Van Dijk