"Entraining" to speech, generating language?

Lars Meyer , Yue Sun & Andrea E. Martin

Published online: 04 Oct 2020.

Submit your article to this journal

View related articles

View Crossmark data

Routledge
Taylor & Francis Group

AUTHORS' RESPONSE

"Entraining" to speech, generating language?

Lars Meyer[a,b], Yue Sun[c] and Andrea E. Martin[d,e]

[a]Research Group Language Cycles, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany; [b]Department of Phoniatrics and Pedaudiology, University Hospital Münster, University of Münster, Münster, Germany; [c]Department of Neuroscience, Max Planck Institute for Empirical Aesthetics, Frankfurt am Main, Germany; [d]Language and Computation in Neural Systems Group, Max Planck Institute for Psycholinguistics, Nijmegen, Netherlands; [e]Donders Centre for Cognitive Neuroimaging, Radboud University, Nijmegen, Netherlands

ABSTRACT
Could meaning be read from acoustics, or from the refraction rate of pyramidal cells innervated by the cochlea, everyone would be an omniglot. Speech does not contain sufficient acoustic cues to identify linguistic units such as morphemes, words, and phrases without prior knowledge. Our target article (Meyer, L., Sun, Y., & Martin, A. E. (2019). Synchronous, but not entrained: Exogenous and endogenous cortical rhythms of speech and language processing. *Language, Cognition and Neuroscience*, 1–11. https://doi.org/10.1080/23273798.2019.1693050) thus questioned the concept of "entrainment" of neural oscillations to such units. We suggested that synchronicity with these points to the existence of endogenous functional "oscillators"—or population rhythmic activity in Giraud's (2020) terms—that underlie the inference, generation, and prediction of linguistic units. Here, we address a series of inspirational commentaries by our colleagues. As apparent from these, some issues raised by our target article have already been raised in the literature. Psycho– and neurolinguists might still benefit from our reply, as "oscillations are an old concept in vision and motor functions, but a new one in linguistics" (Giraud, A.-L. 2020. Oscillations for all A commentary on Meyer, Sun & Martin (2020). *Language, Cognition and Neuroscience*, 1–8).

Introduction

Speech does not mark the boundaries of every linguistic segment, be it morpheme, word, or phrase, in an injective fashion with what we as comprehenders perceive. In our target article (Meyer, Sun, & Martin, 2019), we argued that the alleged entrainment of neural oscillations with such symbolic units is thus implausible. We suggested that synchronicity between oscillations generated in sensory cortices by the physical nature of the sensory stimulus is combined with signals from cortical areas that encode symbolic units, such as morphemes, words, and phrases. In short, and though not fully formalised here (see Martin, 2020), such a claim implies a system architecture where the resulting combined *population-level rhythmic activity* (Giraud, 2020) fluctuates in step with higher-level linguistic structure. Such intrinsic synchronicity, then, can be said to reflect the inference, generation, and prediction of symbolic units like morphemes, words, and phrases. We are honoured and thankful that our article yielded a series of commentaries from auditory neuroscience to neurolinguistics (Ghitza, 2020; Giraud, 2020; Gwilliams, 2020; Haegens, 2020;

Kandylaki & Kotz, 2020; Klimovich-Gray & Molinaro, 2020; Lewis, 2020) ranging from helpfully skeptical to further expansions of the premises put forth in our target article. Below, we discuss the following leitmotifs from these commentaries:

(1) A narrow definition should be met when invoking "entrainment"
(2) If it's not entrainment, call it tracking?
(3) Is intrinsic synchronicity the same as top-down modulation of entrainment?
(4) How do entrainment and intrinsic synchronicity interact?
(5) Periodic linguistic processing without entrainment?
(6) Synchronicity with symbolic units: periodic ERPs?

We did not intend to underemphasise the importance of recent opinion and review articles that focus on the role of endogenous neural oscillations in the top-down modulation of speech entrainment (Haegens & Zion-Golumbic, 2018; cited by Haegens, 2020; Lakatos et al., 2019; Obleser & Kayser, 2019; Rimmele et al., 2018; Zoefel et al., 2018; cited by Haegens, 2020). Publication

CONTACT Lars Meyer ✉ lmeyer@cbs.mpg.de

of these overlapped in time with revisions of our article and we still believe to present timely and interesting perspectives on speech processing that are complementary to our account of language processing. In contrast to the "bottom-up" processing of the speech signal alone, we focused on the generation of linguistic inferences and predictions rather than their possible function in the top-down physiological modulation of "entrainment" proper separate from linguistic representation.

1. A narrow definition should be met when invoking "entrainment"

Haegens' (2020) commentary underlines the necessity to reflect on the loose use of the entrainment term before coining new terms in uncertainty or negation. Because our psycho– and neurolinguistic readership may not be familiar with the need for a narrower definition of entrainment, we made the *Entrainment Checklist* below, copying and explaining (a) to (c) from Obleser & Kayser (2019) and Haegens (2020). For psycho– and neurolinguists who want to study symbolic computations instead of acoustics, we added (d) and (e).

a. *Oscillatory activity in the absence of rhythmic stimulation*

Oscillatory activity in the absence of rhythmic stimulation is a prerequisite for entrainment (Haegens, 2020; Obleser & Kayser, 2019). Resting-state data may provide a data-driven test of this criterion. For example, a magnetoencephalography study by Keitel and Gross (2016) has shown oscillatory activity across the cerebral cortex in the absence of rhythmic stimulation. Their report of spectral differences between cortical areas entails the use of such data for sanity-checking whether speech–brain synchronicity during rhythmic stimulation modulates the activity of a preexisting oscillator or whether it spuriously occurs in an area that does not exhibit an oscillator that operates at the stimulation frequency to start with. Researchers could restrict analyses of data recorded during stimulation to spectro–temporal regions of interest that display preexisting resting-state activity at or near the frequency of rhythmic acoustic stimulation. We note however, that while endogenous eigenfrequencies should play a role—if only to reflect the fact that the brain is a living organ whose metabolic and computational infrastructure is composed of periods of excitation, refraction, and inhibition (Lakatos et al., 2019)—their presence is not inconsistent with electrophysiological responses to speech and language processing being composed of evoked responses, even if the events evoking those responses simply do not exist in speech, but are symbolic.

b. *Frequency-selective phase alignment with a rhythmic external physical stimulus*

An oscillator that comes with its genuine eigenfrequency should not entrain to arbitrary stimulation rhythms. Instead, the magnitude of entrainment should depend on the proximity between the stimulation frequency and the oscillator's eigenfrequency (i.e. the so-called *Arnold Tongue*; Hahn et al., 2019; Hyafil et al., 2015; Lakatos et al., 2019; Obleser & Kayser, 2019; Pikovsky et al., 2002). Arnold Tongues are emerging in the domain of speech entrainment, where oscillatory activity at phoneme– and syllable rates prevails at rest in auditory brain regions (Daube et al., 2019; Giraud et al., 2007; Peelle et al., 2013). Moreover, reduced frontal delta-band activity at rest has been linked to reduced delta-band entrainment (Arns et al., 2007; Hämäläinen et al., 2012; Molinaro et al., 2016; Pagnotta et al., 2015). Of potential interest to linguists, evidence for Arnold Tongues during speech processing could be used to study biological links between speech as an object of cortical information processing and language as a cognitive and cultural system.

c. *Rhythmic activity transiently continues at the stimulus rate after stimulus offset*

The workhorse of entrainment in speech and language processing research is the recording of electrophysiological signals while presenting continuous speech. As most of our readers know, frequency-domain speech–brain synchronicity in such experiments could as easily be mimicked by sequences of evoked responses as it could be driven by "true" oscillations (e.g. Ding & Simon, 2014; Haegens, 2020; Klimesch et al., 2007; Obleser & Kayser, 2019). Critical for psycho– and neurolinguists, this holds not only for acoustic speech rhythms, but for internal computations as well (e.g. N400, P600, and CPS; Frank et al., 2015; Kuperberg et al., 2019; Steinhauer & Friederici, 2001). Alternative experimental designs employ two-phase trial structures involving initial playback of an acoustic rhythm and subsequent presentation of a target stimulus; the rationale is to affect target processing through prior entrainment (e.g. Bosker, 2017; Hickok et al., 2015; Kösem et al., 2018). While, as noted by Haegens (2020), spectral smearing can still jeopardise analysis in the target phase, behavioural effects in the target phase depending on manipulations of the first-phase carrier can nevertheless be interpreted to indicate entrainment. Such designs may thus be preferable.

d. *The stimulus feature under study is physical, not symbolic*

As stressed in our target article, entrainment is a phenomenon that is best understood in relation to physical stimuli, but it is unclear how symbols would drive entrainment proper[1] (e.g. Lewis, 2020). If the phase of an electrophysiological oscillator is supposed to inherit

the phase of a stimulus, the stimulus needs a phase to start with. To support an experimenter's assumption that stimuli exhibit acoustic rhythms that could possibly drive entrainment, experimenters should append the result of spectral analysis of their stimuli. Examples of symbols that cannot cause entrainment in ecologically valid settings are morphemes, words, phonetic features, parts of speech, and information-theoretic complexity metrics, mostly because their onsets and offsets need not be predictable in order to be perceived. Causally speaking, proficient listeners generate associated linguistic symbols on encountering familiar patterns in the electrophysiological imprint served by their auditory periphery. We refer the interested reader to the commentary by Giraud (2020) for an inspiring discussion of the possible neurobiological underpinnings.

e. Rhythmic stimulation is ecologically valid

Experimenters sometimes operationalise speech through the temporal straightjacket of isochrony (for discussion, see Cummins, 2012; Goswami & Leong, 2013; for examples of affected linguistic levels, see also Gwilliams, 2020). We trivially note that when introducing acoustic rhythmicity into an experiment, the odds for observing electrophysiological rhythmicity are good. Before labelling such results entrainment and concluding that the associated processing mechanism is oscillatory, experimenters should motivate their choice of introducing rhythmicity via statistical assessments of natural speech as such (for a good example, see Ding et al., 2017). Critically, the use of naturalistic stimuli (cf. Kandylaki & Kotz, 2020) does not lift this analytical burden: Methodologically speaking, speech–brain synchronicity with natural speech can be significant in spite of lacking stimulus rhythmicity (see Kaufeld et al., 2020).[2] As an example, consider the assumption that speech prosody is rhythmic enough for entraining an electrophysiological oscillator (e.g. Ghitza, 2020), as well as reports of entrainment to natural prosody (Bourguignon et al., 2013; Kaufeld et al., 2020; Mai et al., 2016; Meyer & Gumbert, 2018; Meyer et al., 2016). While there are physiological, environmental, and possibly electrophysiological constraints that may lead to a non-uniform distribution of prosodic events in speech (cf. Kreiner & Eviatar, 2014; Rochet-Capellan & Fuchs, 2014), corpus analyses still suggest substantial variance in the duration of prosodic units (Vollrath et al., 1992). It would be very important to assess whether this variance fits the bandwidth of an electrophysiological oscillator assumed to be devoted to prosody processing.

2. If it's not entrainment, call it tracking?

In line with Obleser & Kayser (2019), Haegens (2020) suggests that phenomena that do not meet a narrow definition of *entrainment*—including cases of synchronicity with symbolic units—should rather be labelled *tracking* or *entrainment in the broad sense*. While *tracking* and *entrainment* are sometimes used interchangeably to label synchronicity with both acoustic and symbolic units (Brennan & Martin, 2020; Cogan & Poeppel, 2011; Daube et al., 2019; Gross et al., 2013; Hämäläinen et al., 2012; Jochaut et al., 2015; Kaufeld et al., 2020; Kayser et al., 2015; Luo & Poeppel, 2007; Luo et al., 2010; Mai et al., 2016; Meyer & Gumbert, 2018; Molinaro et al., 2016; Park et al., 2018; Weissbart, Kandylaki, Reichenbach, et al., 2019; Zoefel & VanRullen, 2016), some authors use *tracking* to label synchronicity with symbolic units alone (Bourguignon, Molinaro, et al., 2020; Brennan & Martin, 2020; Ding et al., 2016; Kaufeld et al., 2020; Zhang & Ding, 2016).

We still worry that *tracking* of symbolic units is not fully consistent: As noted by Lewis (2020), symbolic units cannot be tracked in the literal sense, because the physicality of speech does not feature a one-to-one mapping to symbolic units for an electrophysiological observer to track (Martin, 2016, 2020). For a concrete example, the seminal study by Ding et al. (2016) presented isochronous word sequences in a language that was either native to their participants or not. Each pair of words denoted a syntactic phrase and each quadruplet of words denoted a sentence—which non-native listeners could not recognise. Frequency components in the magnetoencephalogram mirrored the paces of phrases and sentences in native listeners only. Hence, while not physically present in acoustics, these symbolic units were present in listeners' electrophysiology, possibly in a cyclic fashion. Our worry with applying the *tracking* label here: If electrophysiology *tracked* syntactic structures, one would logically entail that electrophysiology tracked *itself*.

Giraud (2020) and Gwilliams (2020) come to our terminological rescue, discussing that synchronicity with symbolic units might not reflect their *tracking*, but their *generation* proper (see also Martin, 2016, 2020; Martin & Doumas, 2017). This suggestion draws a clear line between electrophysiological functions that perceive and process acoustic stimuli and functions that invoke or infer the symbolic units. See below for the implications that this terminological dissociation has for capturing the exogenous–endogenous interplay.

3. Is intrinsic synchronicity redundant with top-down modulation of entrainment?

Klimovich-Gray and Molinaro (2020) worry that the separation of endogenous oscillatory activity from the top-down modulation of entrainment may be overly analytic: Without being at pace with perceptual sampling, the inference, generation, and prediction of linguistic units would be of little use for comprehension. We certainly agree; our target article did not intend to question

interactions between entrainment and intrinsic synchronicity. We acknowledge the role of oscillatory coupling in the functional connectivity between auditory regions and frontal cortices (e.g. Molinaro et al., 2016; Park et al., 2015) as substrate of top-down amplification (e.g. Schroeder et al., 2008; for review, see Vanrullen et al., 2011) and temporal binding (e.g. Giraud & Poeppel, 2012; Morillon et al., 2012). Also, as discussed by Giraud (2020), exogenous and endogenous oscillators may not always be neuroanatomically distinct; instead, a single network might show entrainment while still acting as a pacemaker. Yet, Giraud (2020) also proposes that oscillators that fulfil abstract purposes might lean towards serving as pacemakers, whereas less abstract processes might be dominated by entrainment. Along these lines, both Giraud (2020) and Lewis (2020) suggest that the separation between entrainment and pacemaking might conceptually depend on abstraction and neuroanatomically entail an increasing network size.

Our proposal of intrinsic synchronicity aims to conceptualise these pacemakers as such. Here, we suggest a clinical approach to dissociate entrainment and pacemaking while acknowledging their mutual interactions. As a first example, Broca's aphasia could be a case of intact speech entrainment in spite of abnormal periodic chunking—potentially allowing for the dissociation of exogenous entrainment proper and endogenous linguistic processing. Patients suffering from Broca's aphasia after left-hemispheric precentral lesions exhibit altered auditory chunking time windows in spite of structurally intact auditory regions (Szelag et al., 1997). While we are not aware of reports of intact prosody entrainment in Broca's aphasia, we note that such patients are certainly able to shadow speech (Fridriksson et al., 2012; Fridriksson et al., 2015). In healthy subjects, prosody entrainment associates with auditory, but not precentral activity (Bourguignon et al., 2013). In line with this picture, we found repetitive transcranial magnetic stimulation of left inferior frontal cortex in healthy subjects to affect linguistic chunking while leaving prosody perception intact (Meyer et al., 2018). In principle, lesion data could thus dissociate prosody entrainment from intrinsic synchronicity with endogenously generated linguistic chunks, helping to address the question of whether multi-word chunking relies on prosody entrainment, an internal oscillatory pacemaker, or both (Ghitza, 2020).

As a second example, linguistic dysfunction in schizophrenia (Kircher et al., 2018; Sterzer et al., 2018) has been argued to reflect an imbalance between speech perception and the internal generation of linguistic predictions (Brown & Kuperberg, 2015). In schizophrenia patients suffering from auditory hallucinations, overly strong predictions can trigger the hallucination of words that do not exist outside of the patient's brain (Alderson-Day et al., 2017). Auditory stimulation in schizophrenia patients associates with abnormal beta– and delta-band oscillations (Lakatos et al., 2013), previously proposed to subserve the prediction of content and timing, respectively (Lewis & Bastiaansen, 2015; Schroeder & Lakatos, 2009; Stefanics et al., 2010). In the case of auditory hallucinations, there is no stimulus to entrain to, and thus no entrainment proper to be modulated—could endogenous oscillatory activity underlie the hallucination as such? This hypothesis is supported by the observation of auditory activity to visual-only stimulation with lip movements (Bourguignon et al., 2020; cited by Klimovich-Gray & Molinaro, 2020).

4. How do entrainment and intrinsic synchronicity interact?

Lewis (2020) points out that our target article lacks a proposal as to how entrainment and intrinsic synchronicity might interact. We are thankful for Giraud's (2020) insightful suggestion that the generation of symbolic units may involve a matching between quasi-periodic speech segments and internally stored or generated symbols—such that generative rhythms impose their preferred pace onto perceptual systems to enforce an according segmentation of speech. This is entirely consistent with the analysis-by-synthesis type of proposal laid out in recent models of language processing (Martin, 2016, 2020) which have also been realised in an abstracted computational instantiation (Martin & Doumas, 2017). Such a view is based on claims about how sensory systems make contact with action systems writ large (Buzsáki, 2019; Ernst & Bülthoff, 2004; Olshausen, 2013), but also on classic and core ideas in psycholinguistics (Halle & Stevens, 1962; Marslen-Wilson & Welsh, 1978). Lewis' (2020) concerns resonate with the proposal of a now-or-never bottleneck, according to which abstraction must occur before speech segments corresponding to phonemes, syllables, words, and syntactic structures are forgotten (e.g. Christiansen & Chater, 2015; but note that perceptual memory and memory in language processing may not only be a function of time; see Brown, 1958; McElree, 2006; Peterson & Peterson, 1959; Sperling, 1983). This may also partially answer Lewis' (2020) question of the exact weighting between entrainment and intrinsic synchronicity: Ultimately, the weighting should depend on temporal compatibility between stimulus and inference—balance when the stimulus pace matches the preferred linguistic pace, imbalance when there is a mismatch. Gain and inhibition, and as a result, the phase of the signals that emerge when neural assemblies form dynamically across the language network, likely play crucial roles in

titrating how sensory and abstract neural signals are combined (Martin, 2020).

Giraud (2020) discusses that gamma-band oscillations have been associated both with acoustic processing at the sub-syllabic rate (Daube et al., 2019; Gross et al., 2013) and with the invocation of phonemic categories (Di Liberto et al., 2015; Lehongre et al., 2011; Mesgarani et al., 2014; Nourski et al., 2015). We thank Giraud (2020) for pointing us to literature on endogenous timing constraints on phoneme perception, suggesting that sounds can only be dissociated in time when paced within the period range of lower-gamma-band cycles (Joliot et al., 1994). This observation is consistent with lower-gamma-band cycles acting as pacemakers to sub-syllabic sampling, imposing a preferred pace of invocation of phonemic categories onto auditory sampling. In other words: The endogenous generation of phonemic units sets a pace for the entrainment by the sub-syllabic acoustic rhythm. This should be investigated further.

As a second example from the slow end of the frequency axis, multi-word chunks cannot exceed a duration of about 2–3 s, after which auditory short-term memory fades (Baddeley et al., 1975); in parallel, chunking-related event-related potentials occur without having been triggered by prosodic cues (Schremm et al., 2015). Functional neuroimaging results are consistent with endogenous operation time windows in the order of seconds that are devoted to the processing of multi-word sequences (Hasson et al., 2008; Lerner et al., 2011). Our readers know that it is debated whether delta-band oscillations are exogenously entrained by speech prosody alone to support linguistic chunking (Bourguignon et al., 2013; Gross et al., 2013; Mai et al., 2016) or whether they also underlie the endogenous generation of multi-word chunks (Boucher et al., 2018; Ding et al., 2016; Meyer et al., 2016) or temporal predictions on the time scale of seconds (Arnal et al., 2015; Breska & Deouell, 2017; cited by Haegens, 2020; Donhauser & Baillet, 2019; Lakatos et al., 2008; Lakatos et al., 2013; Meyer & Gumbert, 2018; Stefanics et al., 2010; Weissbart et al., 2019). We concur with Ghitza's (2020) suggestion that all of these can be true, such that speech prosody could exogenously entrain delta-band oscillations, the cycles of which then act as endogenous temporal limiters of chunk duration. This could help explaining why multi-word chunks can be generated without the presence of prosodic cues (i.e. «every prosodic unit is a syntactic unit» does not entail that «every syntactic unit is a prosodic unit»; e.g. Drury et al., 2016; Steinhauer & Friederici, 2001).

5. Periodic linguistic processing without entrainment?

As mentioned in (d) of the Entrainment Checklist, rhythmic acoustic cues may be necessary for entrainment. Yet, as discussed in our target article, there could still be periodic linguistic processing without acoustic cues. In general, periodicity may be the standard, rather than the exception of electrophysiology (Buzsáki, 2006, 2019; Lakatos et al., 2019; Palva et al., 2005; VanRullen, 2016; see also (a) of the Entrainment Checklist). As discussed above, there are endogenous timing constraints on phoneme perception and multi-word chunking (Joliot et al., 1994; Schremm et al., 2015). Auditory neuroscience may inspire psycholinguists to reconsider such evidence in terms of neural oscillations (e.g. Martin, 2020; Tilsen, 2018).

We thank Kandylaki and Kotz (2020) for sketching further directions for such research. In particular, they raise the fascinating possibility that the formation of sentences' verb–argument structure (i.e. those syntactic phrases denoting the who and whom involved in an action, plus the verb denoting the action; Chomsky, 1965) could be a periodic process, pointing to an underlying endogenous oscillatory generator. Indeed, a number of correlations between the phase of delta-band oscillations and syntactic structures (Brennan & Martin, 2020; Ding et al., 2016; Kaufeld et al., 2020; Meyer et al., 2016) or information-theoretic metrics of sequential and syntactic complexity (Meyer & Gumbert, 2018; Weissbart, Kandylaki, Reichenbach, et al., 2019) have been published. In principle, it is still unknown whether these findings reflect verb–argument structure, implicit prosodic phrases (Drury et al., 2016; Frazier et al., 2006; Kreiner & Eviatar, 2014; Schremm et al., 2015), or even information structure (i.e. given–new, theme–rheme, topic–comment alternations over time; for an introduction, see Krifka, 2008; Hagoort, personal communication). Referring to (e) of the Entrainment Checklist, we encourage corpus linguists to haunt for periodicity across these diverse levels of theoretical linguistic description.

6. Synchronicity with symbolic units: periodic ERPs?

We thank Gwilliams (2020) for highlighting one of the biggest challenges for testing whether linguistic processing involves endogenous oscillators: the ambiguity of evoked responses and oscillations (e.g. (c) of the Entrainment Checklist). Linguistic processes with a deterministic time lag can phase-lock endogenous oscillations across trials, masquerading as an evoked response in the average (e.g. Klimesch et al., 2007). In turn, linguistic processes occurring at every given linguistic segment (e.g. syllable, word, or phrase) of a sentence or narrative will give rise to speech–brain synchronicity at the segment frequency. The plot thickens, because, in the limit, it

may be the case that systems of oscillators and series of evoked responses can approximate each other.[3] Here, we discuss exemplary evoked responses for which this might be the case. Afterwards, we provide strategies for assessing whether these have endogenous oscillatory substrates.

In the time domain, linguistic prediction relates to the N400 (Bornkessel-Schlesewsky & Schlesewsky, 2019; Cowles et al., 2007; Fitz & Chang, 2018; Frank et al., 2015; Kuperberg et al., 2019; Rabovsky et al., 2018). In the frequency domain, delta-band oscillations have been linked to computational metrics of linguistic prediction (Meyer & Gumbert, 2018; Weissbart et al., 2019). Frequency decomposition of the N400 shows a dominant delta-band component (Roehm et al., 2009). This pattern leaves it unclear whether linguistic prediction has an evoked or oscillatory substrate.

A second example, the P600, is thought to reflect a revision of the current syntactic structure or overall interpretation (Bornkessel-Schlesewsky & Schlesewsky, 2008, 2019; Kaan & Swaab, 2003; Kuperberg et al., 2019). Noteworthy, the P600 may be elicited by every single word in continuous narratives, depending on the amount of revision or reinterpretation that is required (Hale et al., 2018). While we are not aware of a published frequency decomposition of the P600, a sequence of single-word P600s would likely surface as oscillatory synchronicity between revision or integration demands and the EEG.

Our third example is chunking. The boundaries of multi-word chunks are accompanied by the CPS (Steinhauer et al., 1999). While the CPS can be triggered by prosody (Gilbert et al., 2015; Holzgrefe et al., 2013; Steinhauer, 2003), it can also be triggered by visual cues during reading (Drury et al., 2016; Steinhauer, 2003). Strikingly, the CPS appears with an endogenous period of 2–3 s even in the absence of prosody (Roll et al., 2012; Schremm et al., 2015). Frequency-domain analyses show delta-band phase in the CPS window to predict chunking decisions (Meyer et al., 2016). In principle, the CPS could thus reflect a phase reset of endogenous delta-band oscillations that are devoted to chunking (Boucher et al., 2018; Ding et al., 2016).

To understand the relationship between these evoked responses and endogenous oscillatory activity, single-trial phase should be assessed concurrently to averaging across trials. Averaging is often advocated as prerequisite of sufficient signal-to-noise ratio (Luck, 2014), but multivariate approaches question this assumption (e.g. Sassenhagen & Fiebach, 2019; Sassenhagen et al., 2014). Averaging is also still thought to yield temporally invariant electrophysiological counterparts of discrete processing steps—the boxes and arrows of cognitive (neuro)science (e.g. Luck, 2014). Still, it is unclear whether evoked responses reflect singular stimulus-driven amplitude events rather than the average modulation of ongoing oscillatory activity (Klimesch et al., 2007). In contrast, phase is an indicator of neuronal excitability that directly predicts behavioural responses, thus providing a parsimonious substrate of cognitive processing (Henry & Obleser, 2012; Meyer & Gumbert, 2018; Schroeder et al., 2008; Stefanics et al., 2010). While phase-locking must surface as an evoked response in the average, an association between single-trial phase and a downstream behavioural task can still point to an oscillatory substrate. As related recommendation raised in (e) of the Entrainment Checklist, artificial rhythmicity of endogenous linguistic processing should not be induced by the experimental operationalisation. To close, Giraud (2020) posits that symbolic representations are not extended in time. While we agree that one of the benefits of symbolising a representation is that it need no longer be tied to the vagaries or particulars of the stimulus, environment, or instance, there is no reason that the brain could not generate expectations or gather statistics about when and in what contexts symbolic representations are likely to occur or should be inferred. Thus, decoupling symbols from time is not a necessary condition. Martin and Doumas (2017) showed that a symbolic-connectionist model that uses time to encode functionally symbolic representations in the state dynamics of a neural network could approximate the pattern of neural oscillations found by Ding et al. (2016). The model relies on symbols being extended in time; the rhythmic activity in the neural network, and the ability to separate patterns in the network by when they occur, is what leads to functionally symbolic representations.

Conclusion

Entrainment is a useful concept to describe auditory speech processing. Neural oscillations might also have a role in the inference, generation, and prediction of linguistic units, but this should not be termed entrainment. In turn, the assessment of electrophysiological periodicity of linguistic processing requires dedicated experimental paradigms. In addition, linguists should test for periodicity in speech and text corpora to assess whether such a hypothesis is ecologically valid. Once periodicity has been established, the electrophysiological basis and according limitations of exogenous–endogenous interactions can be pursued. Speech acoustics entrain neural oscillations, but neural oscillations—or population rhythmic activity—likely generate language.

Notes

1. One can imagine an experiment where a participant is asked to monitor a masked visual stream of images for an object, and that if that object were to appear periodically and predictably, perhaps entrainment over and above entrainment to the presentation rate could be observed.
2. To complicate things further, it is not well understood how much isochrony is required for entrainment to occur—in principle, oscillators can tolerate some temporal variability of their entraining stimuli; that is, anisochrony does not necessarily rule out entrainment (Lakatos et al., 2019).
3. Trains of evoked responses and oscillations might not even be mutually exclusive—consider the frequent observation that specific phase intervals raise the probability of evoked responses to occur (e.g., Henry & Obleser, 2012; Lakatos et al., 2019; Stefanics et al., 2010).

Acknowledgments

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

Alderson-Day, B., Lima, C. F., Evans, S., Krishnan, S., Shanmugalingam, P., Fernyhough, C., & Scott, S. K. (2017). Distinct processing of ambiguous speech in people with non-clinical auditory verbal hallucinations. *Brain*, *140*(9), 2475–2489. https://doi.org/10.1093/brain/awx206

Arnal, L. H., Doelling, K. B., & Poeppel, D. (2015). Delta-beta coupled oscillations underlie temporal prediction accuracy. *Cerebral Cortex*, *25*(9), 3077–3085. https://doi.org/10.1093/cercor/bhu103

Arns, M. M., Peters, S., Breteler, R., & Verhoeven, L. (2007). Different brain activation patterns in dyslexic children: Evidence from EEG power and coherence patterns for the double-deficit theory of dyslexia. *Journal of Integrative Neuroscience*, *6*(01), 175–190. https://doi.org/10.1142/S0219635207001404

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior*, *14*(6), 575–589. https://doi.org/10.1016/S0022-5371(75)80045-4

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2008). An alternative perspective on "semantic P600" effects in language comprehension. *Brain Research Reviews*, *59*(1), 55–73. https://doi.org/10.1016/j.brainresrev.2008.05.003

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2019). Toward a neurobiologically plausible model of language-related, negative event-related potentials. *Frontiers in Psychology*, *10*(FEB), 1–17. https://doi.org/10.3389/fpsyg.2019.00298

Bosker, H. R. (2017). Accounting for rate-dependent category boundary shifts in speech perception. *Attention, Perception, & Psychophysics*, *79*(1), 333–343. https://doi.org/10.3758/s13414-016-1206-4

Boucher, V. J., Gilbert, A. C., & Jemel, B. (2018). The role of low-frequency neural oscillations in speech processing: Revisiting delta entrainment. *Journal of Cognitive Neuroscience*, *31*(8), 1205–1215. https://doi.org/10.1162/jocn_a_01410

Bourguignon, M., Baart, M., Kapnoula, E. C., & Molinaro, N. (2020). Lip-reading enables the brain to synthesize auditory features of unknown silent speech. *Journal of Neuroscience*, *40*(5), 1053–1065. https://doi.org/10.1523/JNEUROSCI.1101-19.2019

Bourguignon, M., De Tiege, X., de Beeck, M. O., Ligot, N., Paquier, P., Van Bogaert, P., Goldman, S., Hari, R., & Jousmäki, V. (2013). The pace of prosodic phrasing couples the listener's cortex to the reader's voice. *Human Brain Mapping*, *34*(2), 314–326. https://doi.org/10.1002/hbm.21442

Bourguignon, M., Molinaro, N., Lizarazu, M., Taulu, S., Jousmäki, V., Lallier, M., Carreiras, M., & De Tiège, X. (2020). Neocortical activity tracks the hierarchical linguistic structures of self-produced speech during reading aloud. *NeuroImage*, *216* (March), https://doi.org/10.1016/j.neuroimage.2020.116788

Brennan, J. R., & Martin, A. E. (2020). Phase synchronization varies systematically with linguistic structure composition. *Philosophical Transactions of the Royal Society B: Biological Sciences*. https://doi.org/10.1098/rstb.2019.0305

Breska, A., & Deouell, L. Y. (2017). Neural mechanisms of rhythm-based temporal prediction: Delta phase-locking reflects temporal predictability but not rhythmic entrainment. *PLoS Biology*, *15*(2), 1–30. https://doi.org/10.1371/journal.pbio.2001665

Brown, J. (1958). Some tests of the decay theory of immediate memory. *Quarterly Journal of Experimental Psychology*, *10*(1), 12–21. https://doi.org/10.1080/17470215808416249

Brown, M., & Kuperberg, G. R. (2015). A hierarchical generative framework of language processing: Linking language perception, interpretation, and production abnormalities in schizophrenia. *Frontiers in Human Neuroscience*. https://doi.org/10.3389/fnhum.2015.00643

Buzsáki, G. (2006). *Rhythms of the brain*. Oxford University Press.

Buzsáki, G. (2019). The Brain from inside out. *The Brain from Inside Out*. https://doi.org/10.1093/oso/9780190905385.001.0001

Chomsky, N. (1965). *Aspects of the theory of syntax*. MIT Press.

Christiansen, M. H., & Chater, N. (2015). The now-or-never bottleneck: A fundamental constraint on language. *Behavioral and Brain Sciences*, *39*, 1–72. https://doi.org/10.1017/S0140525X1500031X

Cogan, G. B., & Poeppel, D. (2011). A mutual information analysis of neural coding of speech by low-frequency MEG phase information. *Journal of Neurophysiology*, *106*(2), 554–563. https://doi.org/10.1152/jn.00075.2011

Cowles, H. W., Kluender, R., Kutas, M., & Polinsky, M. (2007). Violations of information structure: An electrophysiological study of answers to wh-questions. *Brain and Language*, *102*(3), 228–242. https://doi.org/10.1016/j.bandl.2007.04.004

Cummins, F. (2012). Oscillators and syllables: A cautionary note. *Frontiers in Psychology*, *3*(OCT), 1–2. https://doi.org/10.3389/fpsyg.2012.00364

Daube, C., Ince, R. A. A., & Gross, J. (2019). Simple acoustic features can explain phoneme-based predictions of cortical

responses to speech. *Current Biology*, *29*(12), 1924–1937.e9. https://doi.org/10.1016/j.cub.2019.04.067

Di Liberto, G. M., O'Sullivan, J. A., & Lalor, E. C. (2015). Low-frequency cortical entrainment to speech reflects phoneme-level processing. *Current Biology*, *25*(19), 2457–2465. https://doi.org/10.1016/j.cub.2015.08.030

Ding, N., Melloni, L., Zhang, H., Tian, X., & Poeppel, D. (2016). Cortical tracking of hierarchical linguistic structures in connected speech. *Nature Neuroscience*, *19*(1), 158–164. https://doi.org/10.1038/nn.4186

Ding, N., Patel, A. D., Chen, L., Butler, H., Luo, C., & Poeppel, D. (2017). Temporal modulations in speech and music. *Neuroscience and Biobehavioral Reviews*. https://doi.org/10.1016/j.neubiorev.2017.02.011

Ding, N., & Simon, J. Z. (2014). Cortical entrainment to continuous speech: Functional roles and interpretations. *Frontiers in Human Neuroscience*, *8*, 1–7. https://doi.org/10.3389/fnhum.2014.00311

Donhauser, P. W., & Baillet, S. (2019). Two distinct neural time-scales for predictive speech processing. *Neuron*, 1–9. https://doi.org/10.1016/j.neuron.2019.10.019

Drury, J. E., Baum, S. R., Valeriote, H., & Steinhauer, K. (2016). Punctuation and implicit prosody in silent reading: An ERP study investigating English garden-path sentences. *Frontiers in Psychology*, *7*, SEP. https://doi.org/10.3389/fpsyg.2016.01375

Ernst, M. O., & Bülthoff, H. H. (2004). Merging the senses into a robust percept. *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2004.02.002

Fitz, H., & Chang, F. (2018). Sentence-level ERP effects as error propagation: A neurocomputational model. *Cognitive Psychology*, 1–46. https://doi.org/10.31234/OSF.IO/FRX2W

Frank, S. L., Otten, L. J., Galli, G., & Vigliocco, G. (2015). The ERP response to the amount of information conveyed by words in sentences. *Brain and Language*, *140*, 1–11. https://doi.org/10.1016/j.bandl.2014.10.006

Frazier, L., Carlson, K., & Clifton, C. (2006). Prosodic phrasing is central to language comprehension. *Trends in Cognitive Sciences*, *10*(6), 244–249. https://doi.org/10.1016/j.tics.2006.04.002

Fridriksson, J., Basilakos, A., Hickok, G., Bonilha, L., & Rorden, C. (2015). Speech entrainment compensates for Broca's Area Damage. *Cortex*, *69*(5), 68–75. https://doi.org/10.1016/j.cortex.2015.04.013

Fridriksson, J., Hubbard, H. I., Hudspeth, S. G., Holland, A. L., Bonilha, L., Fromm, D., & Rorden, C. (2012). Speech entrainment enables patients with Broca's aphasia to produce fluent speech. *Brain*, *135*(12), 3815–3829. https://doi.org/10.1093/brain/aws301

Ghitza, O. (2020). "Acoustic-driven oscillators as cortical pacemaker": a commentary on Meyer, Sun & Martin (2019). *Language, Cognition and Neuroscience*, *0*(0), 1–6. https://doi.org/10.1080/23273798.2020.1737720

Gilbert, A. C., Boucher, V. J., & Jemel, B. (2015). The perceptual chunking of speech: A demonstration using ERPs. *Brain Research*, *1603*, 101–113. https://doi.org/10.1016/j.brainres.2015.01.032

Giraud, A.-L. (2020). Oscillations for all A commentary on Meyer, Sun & Martin (2020). *Language, Cognition and Neuroscience*, 1–8. https://doi.org/10.1080/23273798.2020.1764990

Giraud, A., Kleinschmidt, A., Poeppel, D., Lund, T. E., Frackowiak, R. S. J., & Laufs, H. (2007). Endogenous cortical rhythms determine cerebral specialization for speech perception and production. *Neuron*, *56*(6), 1127–1134. https://doi.org/10.1016/j.neuron.2007.09.038

Giraud, A.-L., & Poeppel, D. (2012). Cortical oscillations and speech processing: Emerging computational principles and operations. *Nature Neuroscience*, *15*(4), 511–517. https://doi.org/10.1038/nn.3063

Goswami, U., & Leong, V. (2013). Speech rhythm and temporal structure: Converging perspectives? *Laboratory Phonology*, *4*(1), 67–92. https://doi.org/10.1515/lp-2013-0004

Gross, J., Hoogenboom, N., Thut, G., Schyns, P., Panzeri, S., Belin, P., & Garrod, S. (2013). Speech rhythms and multiplexed oscillatory sensory coding in the human brain. *PLoS Biology*, *11*(12), e1001752. https://doi.org/10.1371/journal.pbio.1001752

Gwilliams, L. (2020). Hierarchical oscillators in speech comprehension: A commentary on Meyer, Sun, and Martin (2019). *Language, Cognition and Neuroscience*. https://doi.org/10.1080/23273798.2020.1740749

Haegens, S. (2020). Entrainment revisited: A commentary on Meyer, Sun, and Martin (2020). *Language, Cognition and Neuroscience*, 1–5. https://doi.org/10.1080/23273798.2020.1758335

Haegens, S., & Zion-Golumbic, E. M. (2018). Rhythmic facilitation of sensory processing: A critical review. *Neuroscience and Biobehavioral Reviews*, *86*, 150–165. https://doi.org/10.1016/j.neubiorev.2017.12.002

Hahn, G., Ponce-Alvarez, A., Deco, G., Aertsen, A., & Kumar, A. (2019). Portraits of communication in neuronal networks. *Nature Reviews Neuroscience*, *20*(2), 117–127. https://doi.org/10.1038/s41583-018-0094-0

Hale, J. T., Dyer, C., Kuncoro, A., & Brennan, J. R. (2018). Finding Syntax in Human Encephalography with Beam Search. *ACL*, *1*, 1–9. https://doi.org/10.18653/v1/p18-1254

Halle, M., & Stevens, K. (1962). Speech recognition: A model and a program for research. *IEEE Transactions on Information Theory*, *8*(2), 155–159. https://doi.org/10.1109/TIT.1962.1057686

Hämäläinen, J. A., Rupp, A., Soltész, F., Szücs, D., & Goswami, U. (2012). Reduced phase locking to slow amplitude modulation in adults with dyslexia: An MEG study. *NeuroImage*, *59*(3), 2952–2961. https://doi.org/10.1016/j.neuroimage.2011.09.075

Hasson, U., Yang, E., Vallines, I., Heeger, D. J., & Rubin, N. (2008). A hierarchy of temporal receptive windows in human cortex. *Journal of Neuroscience*, *28*(10), 2539–2550. https://doi.org/10.1523/JNEUROSCI.5487-07.2008

Henry, M. J., & Obleser, J. (2012). Frequency modulation entrains slow neural oscillations and optimizes human listening behavior. *Proceedings of the National Academy of Sciences*, *109*(49), 20095–20100. https://doi.org/10.1073/pnas.1213390109

Hickok, G., Farahbod, H., & Saberi, K. (2015). The rhythm of perception entrainment to acoustic rhythms induces subsequent perceptual oscillation. *Psychological Science*, *26*. https://doi.org/10.1177/0956797615576533

Holzgrefe, J., Wellmann, C., Petrone, C., Truckenbrodt, H., Höhle, B., & Wartenburger, I. (2013). Brain response to prosodic boundary cues depends on boundary position. *Frontiers in Psychology*, *4*(JUL), 1–14. https://doi.org/10.3389/fpsyg.2013.00421

Hyafil, A., Giraud, A.-L., Fontolan, L., Gutkin, B., & Hya, A. (2015). Neural cross-frequency coupling: connecting architectures,

mechanisms, and functions. *Trends in Neurosciences*, *38*(11), 725–740. https://doi.org/10.1016/j.tins.2015.09.001

Jochaut, D., Lehongre, K., Saitovitch, A., Devauchelle, A. D., Olasagasti, I., Chabane, N., Zilbovicius, M., & Giraud, A. L. (2015). Atypical coordination of cortical oscillations in response to speech in autism. *Frontiers in Human Neuroscience*, *9*(March), 171. https://doi.org/10.3389/fnhum.2015.00171

Joliot, M., Ribary, U., & Llinás, R. (1994). Human oscillatory brain activity near 40 Hz coexists with cognitive temporal binding. *Proceedings of the National Academy of Sciences of the United States of America*. https://doi.org/10.1073/pnas.91.24.11748

Kaan, E., & Swaab, T. Y. (2003). Repair, revision, and complexity in syntactic analysis: An electrophysiological differentiation. *Journal of Cognitive Neuroscience*, *15*(1), 98–110. https://doi.org/10.1162/089892903321107855

Kandylaki, K. D., & Kotz, S. A. (2020). Distinct cortical rhythms in speech and language processing and some more: A commentary on Meyer, Sun, & Martin (2019). *Language, Cognition and Neuroscience*, 1–5. https://doi.org/10.1080/23273798.2020.1757729

Kaufeld, G., Bosker, H. R., Alday, P. M., Meyer, A. S., & Martin, A. E. (2020). Linguistic structure and meaning organize neural oscillations into a content-specific hierarchy. *BioRxiv*.

Kayser, S. J., Ince, R. A. A., Gross, J., & Kayser, C. (2015). Irregular speech rate dissociates auditory cortical entrainment, evoked responses, and frontal alpha. *Journal of Neuroscience*, *35*(44), 14691–14701. https://doi.org/10.1523/JNEUROSCI.2243-15.2015

Keitel, A., & Gross, J. (2016). Individual human brain areas can be identified from their characteristic spectral activation fingerprints. *PLoS Biology*, *14*(6), 1–22. https://doi.org/10.1371/journal.pbio.1002498

Kircher, T., Bröhl, H., Meier, F., & Engelen, J. (2018). Formal thought disorders: From phenomenology to neurobiology. *The Lancet Psychiatry*, https://doi.org/10.1016/S2215-0366(18)30059-2

Klimesch, W., Sauseng, P., Hanslmayr, S., Gruber, W. R., & Freunberger, R. (2007). Event-related phase reorganization may explain evoked neural dynamics. *Neuroscience and Biobehavioral Reviews*, *31*(7), 1003–1016. https://doi.org/10.1016/j.neubiorev.2007.03.005

Klimovich-Gray, A., & Molinaro, N. (2020). Synchronising internal and external information: A commentary on Meyer, Sun & Martin (2020). *Language, Cognition and Neuroscience*, 1–4. https://doi.org/10.1080/23273798.2020.1743875

Kösem, A., Bosker, H. R., Takashima, A., Meyer, A. S., Jensen, O., & Hagoort, P. (2018). Neural entrainment determines the words we hear. *Current Biology*, *28*(18), 2867–2875.e3. https://doi.org/10.1016/j.cub.2018.07.023

Kreiner, H., & Eviatar, Z. (2014). The missing link in the embodiment of syntax: Prosody. *Brain and Language*, *137*, 91–102. https://doi.org/10.1016/j.bandl.2014.08.004

Krifka, M. (2008). Basic notions of information structure. *Acta Linguistica Hungarica*, *55*(3–4), 243–276. https://doi.org/10.1556/ALing.55.2008.3-4.2

Kuperberg, G. R., Brothers, T., & Wlotko, E. W. (2019). A tale of two positivities and the N400: Distinct neural signatures Are evoked by confirmed and violated predictions at different levels of representation. *Journal of Cognitive Neuroscience*, *32*(1), 12–35. https://doi.org/10.1162/jocn_a_01465

Lakatos, P., Gross, J., & Thut, G. (2019). A new unifying account of the roles of neuronal entrainment. *Current Biology*. https://doi.org/10.1016/j.cub.2019.07.075

Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I., & Schroeder, C. E. (2008). Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science*, *320*(5872), 110–113. https://doi.org/10.1126/science.1154735

Lakatos, P., Schroeder, C. E., Leitman, D. I., & Javitt, D. C. (2013). Predictive suppression of cortical excitability and its deficit in schizophrenia. *Journal of Neuroscience*, *33*(28), 11692–11702. https://doi.org/10.1523/JNEUROSCI.0010-13.2013

Lehongre, K., Ramus, F., Villiermet, N., Schwartz, D., & Giraud, A.-L. (2011). Altered low-gamma sampling in auditory cortex accounts for the three main facets of dyslexia. *Neuron*, *72*(6), 1080–1090. https://doi.org/10.1016/j.neuron.2011.11.002

Lerner, Y., Honey, C. J., Silbert, L. J., & Hasson, U. (2011). Topographic mapping of a hierarchy of temporal receptive windows using a narrated story. *Journal of Neuroscience*, *31*(8), 2906–2915. https://doi.org/10.1523/JNEUROSCI.3684-10.2011

Lewis, A. G. (2020). Balancing exogenous and endogenous cortical rhythms for speech and language requires a lot of entraining: a commentary on Meyer, Sun & Martin (2020). *Language, Cognition and Neuroscience*, 1–5. http://doi.org/10.1080/23273798.2020.1734640

Lewis, A. G., & Bastiaansen, M. C. M. (2015). A predictive coding framework for rapid neural dynamics during sentence-level language comprehension. *Cortex*, *68*, 155–168. https://doi.org/10.1016/j.cortex.2015.02.014

Luck, S. J. (2014). *An introduction to the event-related potential technique* (2nd ed.). The MIT Press. https://doi.org/10.1118/1.4736938.

Luo, H., Liu, Z., & Poeppel, D. (2010). Auditory cortex tracks both auditory and visual stimulus dynamics using low-frequency neuronal phase modulation. *PLoS Biology*, *8*(8), 25–26. https://doi.org/10.1371/journal.pbio.1000445

Luo, H., & Poeppel, D. (2007). Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. *Neuron*, *54*(6), 1001–1010. https://doi.org/10.1016/j.neuron.2007.06.004

Mai, G., Minett, J. W., & Wang, W. S. Y. (2016). Delta, theta, beta, and gamma brain oscillations index levels of auditory sentence processing. *NeuroImage*, *133*, 516–528. https://doi.org/10.1016/j.neuroimage.2016.02.064

Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology*, *10*(1), 29–63. https://doi.org/10.1016/0010-0285(78)90018-X

Martin, A. E. (2016). Language processing as Cue integration: Grounding the psychology of language in perception and neurophysiology. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.00120

Martin, A. E. (2020). A compositional neural architecture for language. *Journal of Cognitive Neuroscience*, 1–20. https://doi.org/10.1162/jocn_a_01552

Martin, A. E., & Doumas, L. A. A. (2017). A mechanism for the cortical computation of hierarchical linguistic structure. *PLoS Biology*, *15*(3), e2000663. https://doi.org/10.1371/journal.pbio.2000663

McElree, B. (2006). Accessing recent events. In B. H. Ross (Ed.), *The psychology of learning and motivation: advances in research and theory* (Vol. 46, pp. 155–200). https://doi.org/10.1016/S0079-7421(06)46005-9.

Mesgarani, N., Cheung, C., Johnson, K., & Chang, E. F. (2014). Phonetic feature encoding in human superior temporal gyrus. *Science*, *343*(6174), 1006–1010. https://doi.org/10.1126/science.1245994

Meyer, L., Elsner, A., Turker, S., Kuhnke, P., & Hartwigsen, G. (2018). Perturbation of left posterior prefrontal cortex modulates top-down processing in sentence comprehension. *Neuroimage*, *181*, 598–604. https://doi.org/10.1016/j.neuroimage.2018.07.059

Meyer, L., & Gumbert, M. (2018). Synchronization of delta-band oscillations to syntax benefits linguistic information processing. *Journal of Cognitive Neuroscience*, *30*, 1066–1074. https://doi.org/10.1162/jocn_a_01236

Meyer, L., Henry, M. J., Gaston, P., Schmuck, N., & Friederici, A. D. (2016). Linguistic bias modulates interpretation of speech via neural delta-band oscillations. *Cerebral Cortex*, *27*(9), 4293–4302. https://doi.org/10.1093/cercor/bhw228

Meyer, L., Sun, Y., & Martin, A. E. (2019). Synchronous, but not entrained: Exogenous and endogenous cortical rhythms of speech and language processing. *Language, Cognition and Neuroscience*, 1–11. http://doi.org/10.1080/23273798.2019.1693050

Molinaro, N., Lizarazu, M., Lallier, M., Bourguignon, M., & Carreiras, M. (2016). Out-of-synchrony speech entrainment in developmental dyslexia. *Human Brain Mapping*, *37*(8), 2767–2783. https://doi.org/10.1002/hbm.23206

Morillon, B., Liégeois-Chauvel, C., Arnal, L. H., Benar, C. G., Giraud, A.-L., & Bénar, C. (2012). Asymmetric function of theta and gamma activity in syllable processing: An intra-cortical study. *Frontiers in Psychology*, *3*(July), 248. https://doi.org/10.3389/fpsyg.2012.00248

Nourski, K. V., Steinschneider, M., Rhone, A. E., Oya, H., Kawasaki, H., Howard, M. A., & McMurray, B. (2015). Sound identification in human auditory cortex: Differential contribution of local field potentials and high gamma power as revealed by direct intracranial recordings. *Brain and Language*, *148*, 37–50. https://doi.org/10.1016/j.bandl.2015.03.003

Obleser, J., & Kayser, C. (2019). Neural entrainment and attentional selection in the listening brain. *Trends in Cognitive Sciences*, 1–14. https://doi.org/10.1016/j.tics.2019.08.004

Olshausen, B. (2013). Perception as an inference problem. In M. S. Gazzaniga & G. R. Mangun (Eds.), *The cognitive Neurosciences* (pp. 295–304). MIT Press.

Pagnotta, M. F., Zouridakis, G., Li, L., Lizarazu, M., Lallier, M., Molinaro, N., & Carreiras, M. (2015). Low frequency overactivation in dyslexia: Evidence from resting state Magnetoencephalography. 2015 *37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 6959–6962. IEEE.

Palva, J. M., Palva, S., & Kaila, K. (2005). Phase synchrony among neuronal oscillations in the human cortex. *Journal of Neuroscience*, *25*(15), 3962–3972. https://doi.org/10.1523/JNEUROSCI.4250-04.2005

Park, H., Ince, R. A. A., Schyns, P. G., Thut, G., & Gross, J. (2015). Frontal top-down signals increase coupling of auditory low-frequency oscillations to continuous speech in human listeners. *Current Biology*, *25*(12), 1649–1653. https://doi.org/10.1016/j.cub.2015.04.049

Park, H., Thut, G., & Gross, J. (2018). Predictive entrainment of natural speech through two fronto-motor top-down channels. *Language, Cognition and Neuroscience*. https://doi.org/10.1080/23273798.2018.1506589

Peelle, J. E., Gross, J., & Davis, M. H. (2013). Phase-locked responses to speech in human auditory cortex are enhanced during comprehension. *Cerebral Cortex*, *23*(6), 1378–1387. https://doi.org/10.1093/cercor/bhs118

Peterson, L., & Peterson, M. J. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology*, *58*(3), 193–198. https://doi.org/10.1037/h0049234

Pikovsky, A., Rosenblum, M., Kurths, J., & Hilborn, R. C. (2002). Synchronization: A universal concept in nonlinear science. *American Journal of Physics*, *70*(6), 655–655. https://doi.org/10.1119/1.1475332

Rabovsky, M., Hansen, S. S., & McClelland, J. L. (2018). Modelling the N400 brain potential as change in a probabilistic representation of meaning. *Nature Human Behaviour*, *2*(9), 693–705. https://doi.org/10.1038/s41562-018-0406-4

Rimmele, J. M., Morillon, B., Poeppel, D., & Arnal, L. H. (2018). Proactive sensing of periodic and aperiodic auditory patterns. *Trends in Cognitive Sciences*, *22*(10), 870–882. https://doi.org/10.1016/j.tics.2018.08.003

Rochet-Capellan, A., & Fuchs, S. (2014). Take a breath and take the turn: How breathing meets turns in spontaneous dialogue. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1658), 20130399. https://doi.org/10.1098/rstb.2013.0399

Roehm, D., Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2009). The internal structure of the N400: Frequency characteristics of a language related ERP component. In *Chaos and Complexity: New Research*.

Roll, M., Lindgren, M., Alter, K., & Horne, M. (2012). Time-driven effects on parsing during reading. *Brain and Language*, *121*(3), 267–272. https://doi.org/10.1016/j.bandl.2012.03.002

Sassenhagen, J., & Fiebach, C. J. (2019). Finding the P3 in the P600: Decoding shared neural mechanisms of responses to syntactic violations and oddball targets. *NeuroImage*, *200*, 425–436. https://doi.org/10.1016/j.neuroimage.2019.06.048

Sassenhagen, J., Schlesewsky, M., & Bornkessel-Schlesewsky, I. (2014). The P600-as-P3 hypothesis revisited: Single-trial analyses reveal that the late EEG positivity following linguistically deviant material is reaction time aligned. *Brain and Language*, *137*, 29–39. https://doi.org/10.1016/j.bandl.2014.07.010

Schremm, A., Horne, M., & Roll, M. (2015). Brain responses to syntax constrained by time-driven implicit prosodic phrases. *Journal of Neurolinguistics*, *35*, 68–84. https://doi.org/10.1016/j.jneuroling.2015.03.002

Schroeder, C. E., & Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. *Trends in Neurosciences*, *32*(1), 9–18. https://doi.org/10.1016/j.tins.2008.09.012

Schroeder, C. E., Lakatos, P., Kajikawa, Y., Partan, S., & Puce, A. (2008). Neuronal oscillations and visual amplification of speech. *Trends in Cognitive Sciences*, *12*(3), 106–113. https://doi.org/10.1016/j.tics.2008.01.002

Sperling, G. (1983). Why we need iconic memory. *Behavioral and Brain Sciences*. https://doi.org/10.1017/S0140525X00014564

Stefanics, G., Hangya, B., Hernádi, I., Winkler, I., Lakatos, P., & Ulbert, I. (2010). Phase entrainment of human delta oscillations can mediate the effects of expectation on reaction speed. *Journal of Neuroscience*, *30*(41), 13578–13585. https://doi.org/10.1523/JNEUROSCI.0703-10.2010

Steinhauer, K. (2003). Electrophysiological correlates of prosody and punctuation. *Brain and Language*, *86*(1), 142–164. https://doi.org/10.1016/S0093-934X(02)00542-4

Steinhauer, K., Alter, K., & Friederici, A. D. (1999). Brain potentials indicate immediate use of prosodic cues in natural speech processing. *Nature Neuroscience*, *2*(2), 191–196. https://doi.org/10.1038/5757

Steinhauer, K., & Friederici, A. D. (2001). Prosodic boundaries, comma rules, and brain responses: The closure positive shift in ERPs as a universal marker for prosodic phrasing in listeners and readers. *Journal of Psycholinguistic Research*, *30*(3), 267–295. https://doi.org/10.1023/A:1010443001646

Sterzer, P., Adams, R. A., Fletcher, P., Frith, C., Lawrie, S. M., Muckli, L., Petrovic, P., Uhlhaas, P., Voss, M., & Corlett, P. R. (2018). The Predictive Coding account of Psychosis. *Biological Psychiatry*, *84*(9), 634–643. https://doi.org/10.1016/j.biopsych.2018.05.015

Szelag, E., Von Steinbüchel, N., & Pöppel, E. (1997). Temporal processing disorders in patients with Broca's aphasia. *Neuroscience Letters*, *235*(1), 33–36. https://doi.org/10.1016/S0304-3940(97)00703-9

Tilsen, S. (2018). *Syntax with oscillators and energy levels*.

VanRullen, R. (2016). Perceptual cycles. *Trends in Cognitive Sciences*, *20*(10), 723–735. https://doi.org/10.1016/j.tics.2016.07.006

Vanrullen, R., Busch, N., Drewes, J., & Dubois, J. (2011). Ongoing EEG phase as a trial-by-trial predictor of perceptual and attentional variability. *Frontiers in Psychology*, *2*(April), 60. https://doi.org/10.3389/fpsyg.2011.00060

Vollrath, M., Kazenwadel, J., & Krüger, H.-P. (1992). A universal constant in temporal segmentation of human speech. *Naturwissenschaften*, *79*(10), 479–480. https://doi.org/10.1007/BF01139205

Weissbart, H., Kandylaki, K. D., & Reichenbach, T. (2019). Cortical tracking of surprisal during continuous speech comprehension. *Journal of Cognitive Neuroscience*, 1–12. https://doi.org/10.1097/PSY.0000000000000183

Weissbart, H., Kandylaki, K. D., Reichenbach, T., Weißbart, H., Kandylaki, K. D., & Reichenbach, T. (2019). Cortical tracking of surprisal during continuous speech comprehension. *Journal of Cognitive Neuroscience*, 1–12. https://doi.org/10.1097/PSY.0000000000000183

Zhang, W., & Ding, N. (2016). Time-domain analysis of neural tracking of hierarchical linguistic structures. *NeuroImage*, *146*(July 2016), 333–340. https://doi.org/10.1016/j.neuroimage.2016.11.016

Zoefel, B., ten Oever, S., & Sack, A. T. (2018). The involvement of endogenous neural oscillations in the processing of rhythmic input: More than a regular repetition of evoked neural responses. *Frontiers in Neuroscience*, *12*. https://doi.org/10.3389/fnins.2018.00095

Zoefel, B., & VanRullen, R. (2016). EEG oscillations entrain their phase to high-level features of speech sound. *NeuroImage*, *124*, 16–23. https://doi.org/10.1016/j.neuroimage.2015.08.054