

Synchronous, but not entrained: exogenous and endogenous cortical rhythms of speech and language processing

Lars Meyer , Yue Sun & Andrea E. Martin

To cite this article: Lars Meyer , Yue Sun & Andrea E. Martin (2020) Synchronous, but not entrained: exogenous and endogenous cortical rhythms of speech and language processing, Language, Cognition and Neuroscience, 35:9, 1089-1099, DOI: [10.1080/23273798.2019.1693050](https://doi.org/10.1080/23273798.2019.1693050)

To link to this article: <https://doi.org/10.1080/23273798.2019.1693050>



© 2019 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 11 Dec 2019.



Submit your article to this journal [↗](#)



Article views: 2422



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 8 View citing articles [↗](#)

Synchronous, but not entrained: exogenous and endogenous cortical rhythms of speech and language processing

Lars Meyer^a, Yue Sun^b and Andrea E. Martin^{c,d,e}

^aResearch Group Language Cycles, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany; ^bDepartment of Neuroscience, Max Planck Institute for Empirical Aesthetics, Frankfurt am Main, Germany; ^cPsychology of Language Department, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands; ^dLanguage and Computation in Neural Systems Group, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands; ^eDonders Centre for Cognitive Neuroimaging, Radboud University, Nijmegen, The Netherlands

ABSTRACT

Research on speech processing is often focused on a phenomenon termed “entrainment”, whereby the cortex shadows rhythmic acoustic information with oscillatory activity. Entrainment has been observed to a range of rhythms present in speech; in addition, synchronicity with abstract information (e.g. syntactic structures) has been observed. Entrainment accounts face two challenges: First, speech is not exactly rhythmic; second, synchronicity with representations that lack a clear acoustic counterpart has been described. We propose that apparent entrainment does not always result from acoustic information. Rather, internal rhythms may have functionalities in the generation of abstract representations and predictions. While acoustics may often provide punctate opportunities for entrainment, internal rhythms may also live a life of their own to infer and predict information, leading to intrinsic synchronicity – not to be counted as entrainment. This possibility may open up new research avenues in the psycho- and neurolinguistic study of language processing and language development.

ARTICLE HISTORY

Received 28 July 2019
Accepted 29 October 2019

KEYWORDS

Entrainment; neural oscillations; inference; prediction; intrinsic synchronicity

The assumed role of entrainment in speech processing

The functional interpretation of cortical rhythms remains an issue of great theoretical importance in research on speech perception and language comprehension (Friederici & Singer, 2015; Giraud & Poeppel, 2012; Lewis & Bastiaansen, 2015; Meyer, 2017). Given the fact that language comprehension is essentially the inference of meaning from vibrations of air, comprehension must require the synthesis of prior knowledge in the form of endogenous information (e.g. a network’s intrinsic activation state at stimulus onset) with sensory information in the form of exogenous acoustic input (Bever & Poeppel, 2010; Halle & Stevens, 1962; Martin, 2016; Poeppel & Monahan, 2011). Yet, in the past years, the auditory neuroscience and speech processing fields have focused mainly on the so-called “entrainment” of cortical rhythms during the processing of acoustics and pre-lexical representations such as phonemes, phonetic features, and syllables. In comparison, less focus has been put on potential cognitive and computational aspects of these signals for lexical, morphemic, syntactic, semantic, and discourse- and referential-level processing, all of which are crucial in understanding the meaning of speech, which is, of course, the goal of language comprehension.

Entrainment describes the phase-locking of a neural oscillation, presumed to emanate from a population of neurons that fire in synchrony, to the phase of an external physical stimulus, such as speech (Giraud & Poeppel, 2012; Lakatos, Karmos, Mehta, Ulbert, & Schroeder, 2008; Obleser & Kayser, 2019; Piskovsky, Rosenblum, & Kurths, 2003). In the narrow sense, which we refer to as entrainment proper, a given rhythmic sequence of acoustic cues drives the cycles of a given neural oscillation into a phase-aligned rhythmic sequence. Entrainment proper is contingent on the acoustic stimulus and requires acoustic cues. Entrainment of neural oscillations to amplitude rhythms in speech is attested at the gamma (Gross et al., 2013; Lehongre, Ramus, Villiermet, Schwartz, & Giraud, 2011), theta (Peelle, Gross, & Davis, 2013), and delta (Bourguignon et al., 2013) frequencies (Gross et al., 2013). It is thought that these electrophysiological rhythms can be entrained in the first place only because they match in frequency the rhythmic amplitude edges or peaks that accompany phonemes, syllables, and intonation phrases; the latter have a comparably weak rhythmicity. The general assumption that acoustic edges or peaks are consistently rhythmic and physically strong enough to entrain a neural oscillation is supported by rodent work, where

CONTACT Lars Meyer  lmeyer@cbs.mpg.de

© 2019 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

high-amplitude neuronal discharges to complex acoustic stimuli were observed to reset the phase of local field potentials in the auditory cortex (Szymanski, Rabinowitz, Magri, Panzeri, & Schnupp, 2011).

Challenge: synchronicity, but sparseness or absence of acoustic cues

It has often been proposed that the entrainment of neural oscillations plays a mechanistic role in speech and language processing. Evidence for this role is mostly confined to the entrainment of theta-band oscillations to speech amplitude modulations at the syllabic rate, which is emphasised in recent neurophysiological and computational models (Ghitza, 2011; Giraud & Poeppel, 2012). These models conceptualise theta-band entrainment as a mechanism to segment continuous speech signals into syllable-size acoustic chunks, guiding follow-up auditory decoding on shorter time scales, such as phonemes. As each syllable contains a phonetic unit that is acoustically salient, it is likely that exogenous acoustic landmarks are dominant, if not crucial for the elicitation of synchronicity between electrophysiological rhythms and speech. This view is further corroborated by a close correlation between amplitude modulations at theta-band frequency in speech and average syllable duration in speech corpora (Greenberg, 2001; Pellegrino, Coupé, & Marsico, 2011); further support comes from psychophysical studies (Ghitza & Greenberg, 2009).

In direct contrast to such evidence for entrainment proper in the syllabic range, there are clear-cut cases where speech does simply not exhibit any physical cues that could possibly entrain neural oscillations. Instead, synchronicity occurs between specific frequency bands of the electroencephalogram and linguistic representations that, in principle, only exist in the mind and brain through perceptual inference (Marslen-Wilson & Welsh, 1978; Martin, 2016): There is no isomorphic relationship between a sound wave and a word, its meaning, and its syntactic category. Instead, the relationship is symbolic: A physical sound wave is arbitrarily associated with a meaning that must be decoded in context (Ding & He, 2016; Ding, Melloni, Zhang, Tian, & Poeppel, 2016; Meyer, Henry, Schmuck, Gaston, & Friederici, 2016). First, words in running speech do not have acoustic boundaries; so already the segmentation of running speech into individual words is a case of inference (e.g. Kösem, Basirat, Azizi, & van Wassenhove, 2016; Lany & Saffran, 2010; Martin, 2016). Second, word meaning cannot be implicitly derived from a sound wave (Dingemans, Blasi, Lupyan, Christiansen, & Monaghan, 2015; Saussure, 1916); instead, the association between word meaning and sound

identity is mostly arbitrary cross-linguistically. In addition, a single segmented acoustic word is often associated with several meanings in the mental lexicon (Aitchison, 2012) the selection of which depends on context (Nieuwland & Van Berkum, 2006). Third, the syntactic categories of words in many languages are not marked acoustically – yet, sentence meaning derives from syntactic categories: Comprehending who-did-what-to-whom requires the establishment of relationships amongst words – the formation of a syntactic structure based on syntactic categories. Yet, often more than a single syntactic structure is compatible with a given utterance, leading to multiple mutually exclusive interpretations of the same utterance (Meyer et al., 2016). The assignment of syntactic categories, the establishment of relationships amongst words, and the comprehension of who-did-what-to-whom rely on an inferential link between sensory input and grammatical knowledge of language.

Hence, entrainment proper is an unlikely mechanism for the formation of higher-level linguistic representations. How could synchronicity then be triggered by something that does not have clear physicality in the external world? The proposed dissociation between stimulus-dominant entrainment proper and intrinsic synchronicity with abstract symbolic linguistic information is in line with a series of results: For example, language comprehension is largely intact when prosodic cues, which occur within the modulation frequency range of the delta band, are removed from the speech signal (Ding et al., 2016) – while removal of theta-range amplitude modulations results in a substantial loss of intelligibility (Doelling, Arnal, Ghitza, & Poeppel, 2014; Ghitza & Greenberg, 2009). In addition, phase-locking to linguistic structure occurs even when the speech stimulus contains distracting exogenous acoustic information (Meyer et al., 2016). It is even intensively debated and still requires future corpus analyses whether amplitude cues provided by pauses, duration information, and pitch modulations that occur in frequency ranges slower than theta are rhythmic enough to entrain oscillations in the first place; it is also unclear whether these cues are reliable enough to infer lexical or phrasal boundaries (Cummins, 2012; Fernald & McRoberts, 1996; Goswami & Leong, 2013; Jusczyk, 1997; Kelso, Saltzman, & Tuller, 1986; Martin, 2016).

Proposal: intrinsic synchronicity versus entrainment proper

We propose here that oscillatory synchronicity during speech processing and language comprehension is often not entrainment proper. Instead, the symbolic relationship between acoustic cues and the computation

of abstract structures and linguistic predictions implies that oscillatory rhythmicity could also be intrinsically synchronous with the pace of ongoing inferences – and thus strictly cognitive processing (Marslen-Wilson & Welsh, 1978; Martin, 2016; Martin & Doumas, 2017, 2019). Morphemes, lexical representations, syntactic and semantic structures, discourse and event structures, as well as pragmatic inferences cannot be considered sensory-driven consequences of speech alone; instead, they are generated in predicted internally.

Why should the computation of abstract linguistic structures be cyclic? There is evidence that such computation is bound by endogenous constraints on temporal regularity. In the domain of abstract syntactic processing, listeners are biased to group words into implicit phrases with a period that is highly regular across both time and participants (Fodor, 1998; Hwang & Steinhauer, 2011; Webman-Shafran & Fodor, 2016). This bias is strong enough to override exogenous prosodic cues that indicate phrase durations outside of the preferred phrasing period (Meyer, Elsner, Turker, Kuhnke, & Hartwigsen, 2018; Meyer et al., 2016). Furthermore, event-related brain potentials (ERPs) associated with the grouping of words into implicit phrases appear with a regular period that does not require the presence of periodic exogenous prosodic cues (Roll, Lindgren, Alter, & Horne, 2012; Schremm, Horne, & Roll, 2015; Steinhauer, Alter, & Friederici, 1999). Hidden in the frequency domain of such periodically occurring ERPs, there may be a slow-frequency oscillator that is synchronous with internally generated syntactic representations. This is consistent with the observation that the grouping of words into phrases associates with delta-band oscillatory activity, encompassing the range of periodicity of grouping-related ERPs (Bonhage, Meyer, Gruber, Friederici, & Mueller, 2017; Ding et al., 2016; Meyer et al., 2016; cf. Boucher, Gilbert, & Jemel, 2019). While grouping maximises information throughput beyond the capacity constraints of working memory (e.g. Baddeley, Hitch, & Allen, 2009; Bonhage et al., 2017), periodicity might ensure optimal use of electrophysiological constraints, such as the eigenfrequencies of cortical networks, and thus the time windows across which information can be integrated (e.g. Buzsaki, 2006, 2019; Keitel & Gross, 2016). The idea that endogenous oscillatory activity may be a reason for discretized information sampling has been discussed in detail elsewhere (e.g. Pöppel, 1997; VanRullen, 2016). Again, intrinsic synchronicity with internally generated word groups can have been “disguised” as entrainment in prior work, as it overlaps with the frequency band of apparent entrainment by speech prosody (i.e. <4 Hz; e.g. Bourguignon et al., 2013; Gross et al., 2013; Mai, Minett, & Wang, 2016). This hypothesis

would be supported if intrinsic synchronicity with word groups and prosodic entrainment were dissociated through their different cortical substrates. In fact, delta-band activity has not only been reported for speech entrainment in the vicinity of auditory cortices, but also for higher-level processes in frontal cortices (e.g. Molinaro, Lizarazu, Lallier, Bourguignon, & Carreiras, 2016; Park, Ince, Schyns, Thut, & Gross, 2015).

In addition to the internal generation of abstract linguistic structures, why should abstract linguistic predictions be assumed to employ ongoing oscillatory electrophysiological activity? Abstract linguistic representations live a life of their own, such that the lexical-semantic or syntactic context accumulating over time allows for the continuous derivation and refinement of linguistic predictions (Friston, 2005; Levy, 2008; Lewis & Bastiaansen, 2015; Martin, 2016; Nieuwland & Van Berkum, 2006), although it remains unclear in which granularity these are propagated and to what degree these are necessary to interpret language (Nieuwland et al., 2018). Predictions are made within their respective domain; for instance, linguistic predictions derived from preceding word meanings allow for predictions of the meaning of individual upcoming words; linguistic predictions derived from preceding syntactic categories allow for predictions of the syntactic category of upcoming groups of words (Frank, Otten, Galli, & Vigliocco, 2015; Hale, Dyer, Kuncoro, & Brennan, 2018; Meyer & Gumbert, 2018). Predictions may also capitalise on the fact that syntactic and semantic or syntactic and discourse phenomena are often correlated with one another, having analog forms that must correspond on each level of representation and allow for iterative resampling (Martin, 2016). In either case, oscillatory power in the beta band was repeatedly found to be modulated by linguistic predictability (e.g. Lewis, Schoffelen, Hoffmann, Bastiaansen, & Schriefers, 2017; Wang et al., 2012) and sensory predictability in the auditory and audio-visual modality (Arnal & Giraud, 2012; Arnal, Wyart, & Giraud, 2011; Kim & Chung, 2008; Weiss & Mueller, 2012). In fact, the beta band has been proposed to subserve the internal generation of predictions across linguistic domains (Lewis & Bastiaansen, 2015; Lewis, Schoffelen, Schriefers, & Bastiaansen, 2016), in line with the role of the beta band proposed in the literature on predictive coding (e.g. Chao, Takaura, Wang, Fujii, & Dehaene, 2018; Engel & Fries, 2010; Roopun et al., 2008). In line with our suspicion of the potential confoundedness of intrinsic synchronicity and entrainment, prediction-related beta-band activity (i.e. 13–30 Hz) is immediately adjacent to – or even overlaps with – lower gamma-band activity, where entrainment at phoneme rate has been claimed to occur (e.g. 25–35 Hz, Lehongre et al., 2011; 35–45 Hz, Gross et al., 2013; 30–45 Hz, Di Liberto, O’Sullivan, & Lalor, 2015; 30 Hz, Lizarazu et al., 2015). To dissociate phoneme-

rate entrainment and prediction-related intrinsic synchronicity, one could hypothesise to observe phoneme-rate entrainment in auditory cortex, but prediction-related intrinsic synchronicity in surrounding association cortex.

The strictly endogenous character of intrinsic synchronicity may dissociate it from the modulation of entrainment by domain-general or linguistic top-down processes (e.g. Obleser & Kayser, 2019; Rimmele, Morillon, Poeppel, & Arnal, 2018). For example, top-down activity can phase-shift entrained oscillations in sensory systems to increase the temporal alignment of neuronal excitability with external information (Arnal, Doelling, & Poeppel, 2015; Arnal & Giraud, 2012; Kayser, Wilson, Safaai, Sakata, & Panzeri, 2015; Lakatos et al., 2005, 2008; Park et al., 2015; Schroeder & Lakatos, 2009; Schroeder, Wilson, Radman, Scharfman, & Lakatos, 2010). In addition, the phase of neural oscillations prior to stimulus occurrence can determine how well a stimulus is perceived and processed in the visual, somatosensory, and auditory domains (e.g. Addante, Watrous, Yonelinas, Ekstrom, & Ranganath, 2011; Becker, Ritter, & Villringer, 2008; Hanslmayr et al., 2007; lemi et al., 2019; Malteva, Geissler, & Başar, 2000; Mathewson, Gratton, Fabiani, Beck, & Ro, 2009; Reinacher, Becker, Villringer, & Ritter, 2009; Rimmele et al., 2018; VanRullen, Busch, Drewes, & Dubois, 2011; Weisz et al., 2014; Wöstmann, Waschke, & Obleser, 2019). In contrast, we mean by intrinsic synchronicity that internal inferential, generative, and predictive processes as such may operate at their own rhythm. Conceptually speaking, the output of such intrinsically synchronous processes could in principle be “symbolic cues” or “hazard functions” (Rimmele et al., 2018) – that is, internal information that might affect entrainment by resetting the phase of entrained oscillators from the inside out. Still, the actual inference, generation, or prediction of a “symbolic cue” (Martin, 2016) or “hazard function” (Rimmele et al., 2018) is just that, but neither entrainment nor its top-down modulation.

As possible cases in principle of intrinsic synchronicity, entrainment by acoustic and temporal information that is not diagnostic of a unique linguistic unit can still result in the stable perception of a lexically ambiguous stimulus when task instructions are to detect either one or the other of the possible percepts (Kösem et al., 2016). A similar effect has been found during the processing of syntactically ambiguous sentences, where sentence interpretations that contradict acoustic cues are associated with a decreased entrainment by these cues, but a consistent phase shift towards the syntactic structure that is actually perceived (Meyer et al., 2016). While these types of effects strongly suggest that deployment of intrinsic signals in the form of cortical rhythms can shape stimulus comprehension, further studies are certainly required. A first setup to test our proposal could exploit the

phenomenon that auditory processing performance transiently keeps stimulation frequency even after stimulation offset (Hickok, Farahbod, & Saberi, 2015; Neuling, Rach, Wagner, Wolters, & Herrmann, 2012). Translating this to the language domain, behavioural speech perception research has found that prior speech rate can affect the perception of downstream words when speech rate is subsequently increased or reduced, such that incoming phonemes that do not match the prior speech rate are misperceived (Bosker, 2017; Dilley & Pitt, 2010). It has recently been shown that such effects are accompanied by endogenous oscillatory phase shifts (Kösem et al., 2018). To rule out that continued rhythmicity is not simply a reverberation of prior entrainment, continued rhythmicity would have to be experimentally elicited without acoustic cues. For instance, one could think of experimental paradigms that require the rhythmic internal generation of abstract linguistic representations, which then could be shown to affect abstract linguistic processing after stimulation offset.

A second setup to test our proposal could exploit the fact that many sentences allow for the formation of multiple possible syntactic structures. One example are ambiguous relative clauses, such as in *The doctor met the son of the colonel who died.*, where either *the son* or *the colonel* might have died, depending on the syntactic structure that the listener formed (Grillo, Costa, Fernandes, & Santi, 2015; Hemforth et al., 2015). A second example are prepositional phrases such as in *The client sued the murderer with the corrupt lawyer*, where *the corrupt lawyer* might either go with *the client* or with *the murderer*, depending on the syntactic structure that is generated (Meyer et al., 2016, 2018; Wales & Toner, 1979). In psycholinguistics, various endogenous sources of bias in the assignment of syntactic structure to ambiguous sentences have been identified, including verbs’ selection restrictions (Sedivy & Spivey-Knowlton, 1994), noun semantics (MacDonald, Pearlmuter, & Seidenberg, 1994), and working memory capacity limitations (Swets, Desmet, Hambrick, & Ferreira, 2007); all of these endogenous factors that are well understood in psycholinguistics should be experimentally tested for eliciting cases of apparent entrainment that are, however, cases of strictly intrinsic synchronicity.

A line of results that are problematic for the exogenous-dominant account of rhythmic neural oscillations in response to speech stimuli are compatible with our new proposal: First, rhythmicity of neural oscillations does still occur when the rhythmicity of external amplitude modulations is experimentally reduced (Calderone, Lakatos, Butler, & Castellanos, 2014; Mathewson et al., 2012; for review, see Ding & Simon, 2014; Zoefel & VanRullen, 2015), although we acknowledge that decreased temporal consistency of a stimulation rhythm has also been observed to reduce entrainment to some degree

(Mathewson et al., 2012). Second, when rhythmic amplitude cues are experimentally kept identical between a vocoded and a non-vocoded speech condition, phase-locking is still increased for the non-vocoded condition (Peelle et al., 2013). Third and more generally, our proposal may help to address the recurring concern that amplitude modulations in speech are too arrhythmic to allow for exogenous entrainment by acoustic cues, oscillatory activity may still be rhythmic (Cummins, 2012; Goswami & Leong, 2013; Kelso et al., 1986; Mathewson et al., 2012): An intrinsically synchronous oscillator could exhibit rhythmic behaviour disguised as entrainment, which, however, would in fact underlie the generation of abstract linguistic representations or predictions.

How and why entrainment and intrinsic synchronicity may relate

We next turn to the question of the relationship between entrainment proper and intrinsic synchronicity. The availability of acoustic cues for entrainment proper changes over time. Likewise, the richness, detail, and possible forecasting of internally generated abstract representations, mirrored by intrinsic synchronicity, may change over time. Thus, during speech processing and language comprehension, the relationship between exogenous acoustic and endogenous abstract information in cortical networks may be highly dynamic over time (Arnal & Giraud, 2012; Herrmann, Munk, & Engel, 2004; Marslen-Wilson & Welsh, 1978; Martin, 2016; Martin & Dumas, 2017, 2019; Seidl, 2007; Sherman, Kanai, Seth, & VanRullen, 2016). One could even hypothesise that behaviourally, speech perception and language comprehension might stay equally good under dynamically changing environmental conditions, due to dynamic fluctuations in the exogenous–endogenous weighting over time. When the speech signal is clear and environmental conditions are excellent, entrainment can dominate speech perception. In turn, any representation that has been generated internally allows for the generation of predictions across the various linguistic levels (Hale, 2001; Levy, 2008; Martin, 2016). This can result from the same inferential process that structure building would, without the invocation of an additional predictive mechanism (Marslen-Wilson & Welsh, 1978; Martin, 2016; Martin & Dumas, 2017, 2019). Endogenously generated linguistic information can thus keep neural oscillations in the auditory system in a rhythmic state, in effect stabilising temporal alignment of neuronal excitability with residual acoustic (e.g. spectral) cues in the speech stimulus (Arnal & Giraud, 2012; Arnal et al., 2015; Kayser et al., 2015; Lakatos et al., 2005; Park et al., 2015) even when

exogenous acoustic amplitude cues at these faster frequencies are arrhythmic or temporarily lacking (see Figure 1).

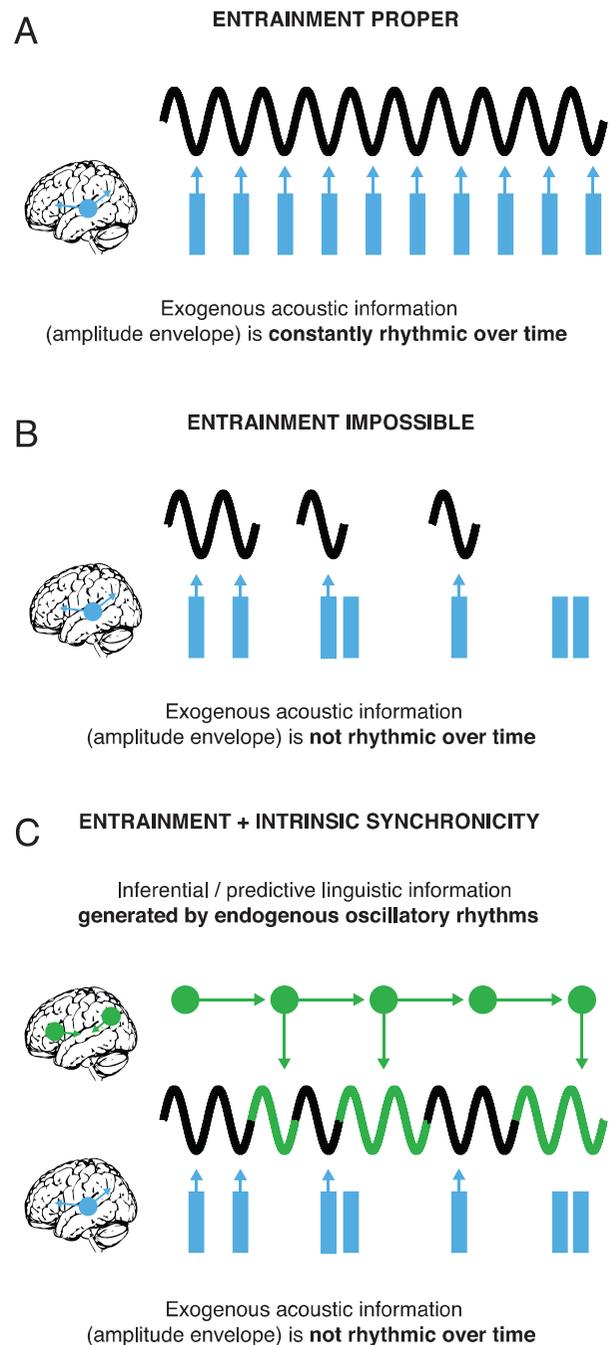


Figure 1. A: Entrainment of neural oscillations by a rhythmic exogenous acoustic stimulus (e.g. a regular tone sequence). Rhythmic edges or peaks in the amplitude envelope of the stimulus synchronise neural oscillation at stimulation frequency; B: Entrainment impossible in cases of non-rhythmic amplitude cues in speech stimulus; C: Intrinsic synchronicity of neural oscillations to a non-rhythmic acoustic stimulus (e.g. speech). Inferred and predictive linguistic knowledge (e.g. of abstract syntactic structure or predicted words) deployed by endogenous signals (i.e. local-field potentials or slower-frequency neural oscillations) establishes continued oscillatory rhythmicity disguised as entrainment, in spite of lacking acoustic cues.

In line with our proposal, the processing of clear speech, relative to the processing of acoustically degraded speech, is not only associated with reliable entrainment by sensory input, but also with an increase in the endogenous modulation of sensory entrainment (Herrmann et al., 2004; Park et al., 2015; Sherman et al., 2016) – endogenous neural oscillations that provide inferred linguistic information may thus provide human language with a powerful degree of redundancy, potentially compensating for the transient non-rhythmicity of acoustic information throughout speech. An analogy for the interplay of endogenous abstract and exogenous acoustic signals would be riding a bicycle, where you might achieve the same overall speed with strong pedalling of the left leg, but weaker pedalling of the right leg, or vice versa. When exogenous acoustic information is not strong enough to entrain to, endogenous information may jump in. Such ongoing imbalance would ensure an optimal use of the total amount of available information, be it acoustic or internal.

Functional neuroanatomy: abstraction from local to global?

The proposed ongoing entrained–intrinsic imbalance requires an underlying functional neuroanatomy that can handle increasing degrees of cognitive abstraction, possible relying on cortical networks of increasing size and complexity. At present, we may only speculate that the functional neuroanatomy of the dynamic interplay between exogenous acoustic (a.k.a., incoming, perceived, predicted) and endogenous abstract (a.k.a., internally generated, inferred, predictive) depends on the degree of linguistic abstraction. On lower abstraction levels (e.g. phoneme level), local networks might achieve the exogenous–endogenous interplay; on higher, abstract levels (e.g. syntax and semantics), networks may increase. In general, it has been proposed that representational abstraction goes hand in hand with an increase in the size of the involved oscillatory network (Buzsaki, 2006, 2019; Sarnthein, Petsche, Rappelsberger, Shaw, & von Stein, 1998; von Stein, Rappelsberger, Sarnthein, & Petsche, 1999). It has also been hypothesised that the associated increase in the variance in conduction delays results in an increase in the wavelength of those oscillations that underlie network-level processing (Buzsaki, 2006).

In support of this tentative proposal, gamma-band entrainment of circumscribed auditory regions has been observed both to phoneme-rate amplitude modulations (Di Liberto et al., 2015; Gross et al., 2013) and phonological categories (Lehongre et al., 2011; Nourski et al., 2015). Phonemic-categorical information as such is not present in the speech signal (i.e. exogenous), but can only be

inferred with the help of abstract linguistic knowledge (i.e. endogenously). Yet, the involved networks do not extend beyond auditory association cortex (e.g. superior posterior temporal cortex; Mesgarani, Cheung, Johnson, & Chang, 2014): Electroencephalographic data suggest that auditory association cortex is active when phonemic-categorical representations are inferred once the associated acoustic information is artificially removed from speech (Leonard, Baud, Sjerps, & Chang, 2016).

Some evidence is also compatible with the proposed involvement of large-scale, slow-frequency networks in the exogenous–endogenous interplay during abstract linguistic processing: Synchronicity of delta-band oscillations in frontal cortices increases for speech as compared to both amplitude-modulated white noise and spectrally rotated speech – that is, when linguistic information can be inferred in the first place (Molinaro & Lizarazu, 2017). Likewise, effective connectivity from frontal to posterior cortices in the delta band increases for clear speech compared to backward speech (Park et al., 2015). In general, literature on early evoked responses suggests that the diagnosis of anomalies that violate abstract expectations derived from internal linguistic knowledge is not achieved by sensory cortices alone (Dikker, Rabagliati, & Pykkänen, 2009; Herrmann, Maess, Hasting, & Friederici, 2009), but involves additional generators in the frontal cortices (Friederici, Wang, Herrmann, Maess, & Oertel, 2000). When testing this experimentally, it could be hypothesised that entrainment proper (i.e. driven by acoustic cues exclusively) is restricted to sensory cortices. Intrinsic synchronicity accompanying categorical abstraction (e.g. inference of phonemic features) should be observed in sensory association cortices. Increasingly abstract and generative processes (e.g. the generation of syntactic structure, linguistic predictions) would be hypothesised to associate with intrinsic synchronicity in larger frontal–posterior networks.

Consequences for language acquisition: from entrained to intrinsic?

In addition to providing a novel explanation for oscillatory phenomena in speech processing and language comprehension, our proposal also allows for new ways to study language acquisition – as a case in principle for a progression from exogenously-driven entrainment proper to endogenously-driven intrinsic synchronicity. Two examples offer support for the idea that the accumulation of linguistic knowledge leads to the deployment of endogenous signals that give rise to intrinsic synchronicity.

First, late in the womb and up to six months after birth, fetuses and neonates are able to distinguish both

native and non-native phonemes (Gervain, 2015; Mahmoudzadeh et al., 2013); six months after birth and possibly earlier (Moon, Lagercrantz, & Kuhl, 2013), this perceptual ability narrows selectively towards native phonemes (Kuhl, 2000; Kuhl, Ramirez, Bosseler, Lin, & Imada, 2014). Such perceptual specialisation occurs both on the phonemic and syllabic level; strikingly, for phonemes, specialisation is associated with an enhancement of gamma-band entrainment (Ortiz-Mantilla, Hamalainen, Realpe-Bonilla, & Benasich, 2016), whereas for syllables, it is associated with an amplification of theta-band entrainment (Bosseler et al., 2013; Ortiz-Mantilla et al., 2016).

Second, a compatible behavioural trajectory is known for abstract linguistic information: In the developing ability to parse words into syntactic phrases, infants start out with the ability to perceive – and later infer – these phrases' boundaries through exogenous acoustic amplitude cues (Isobe, 2007; Männel & Friederici, 2009; Wiedmann & Winkler, 2015); yet, after six years of age, amplitude cues are not necessary for syntactic phrasing anymore (Männel, Schipke, & Friederici, 2013) and perhaps become overridden by information that manifests as endogenous syntactic preferences in adults – in association with a decreased entrainment by amplitude cues that contradict these endogenous preferences (Meyer et al., 2016).

Conclusion

We have argued that segmenting speech into representations with structure and meaning calls on two forms of synchronised neural oscillations: Entrainment proper occurs in response to exogenous stimulus, such as acoustic speech envelopes, and likely to other acoustic features of speech. Intrinsic synchronicity may disguise as entrainment, yet stems from the generation of linguistic meaning based on the synthesis of the exogenous acoustic signal with (pre-)activated endogenous representations through perceptual inference.

Acknowledgements

We thank Angela D. Friederici, Antje S. Meyer, and David Poeppel for helpful comments; we also thank three anonymous reviewers for helping our terminological precision, rigor, and logical stringency, and for raising a number of critical issues that we did not consider initially.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

LM, YS, and AEM were supported by the Max Planck Society. LM was supported by the Max Planck Research Group “Language Cycles”. AEM was supported by the Max Planck Research Group “Language and Computation in Neural Systems” and by the Netherlands Organization for Scientific Research (grant 016.Vidi.188.029).

References

- Addante, R. J., Watrous, A. J., Yonelinas, A. P., Ekstrom, A. D., & Ranganath, C. (2011). Prestimulus theta activity predicts correct source memory retrieval. *Proceedings of the National Academy of Sciences*, 108, 10702–10707.
- Aitchison, J. (2012). *Words in the mind: An introduction to the mental lexicon*. Chichester: John Wiley & Sons.
- Arnal, L. H., Doelling, K. B., & Poeppel, D. (2015). Delta-beta coupled oscillations underlie temporal prediction accuracy. *Cerebral Cortex*, 25(9), 3077–3085.
- Arnal, L. H., & Giraud, A. L. (2012). Cortical oscillations and sensory predictions. *Trends in Cognitive Sciences*, 16(7), 390–398.
- Arnal, L. H., Wyart, V., & Giraud, A. L. (2011). Transitions in neural oscillations reflect prediction errors generated in audiovisual speech. *Nature Neuroscience*, 14(6), 797–801.
- Baddeley, A., Hitch, G., & Allen, R. (2009). Working memory and binding in sentence recall. *Journal of Memory and Language*, 61, 438–456.
- Becker, R., Ritter, P., & Villringer, A. (2008). Influence of ongoing alpha rhythm on the visual evoked potential. *Neuroimage*, 39, 707–716.
- Bever, T. G., & Poeppel, D. (2010). Analysis by synthesis: A (re-)emerging program of research for language and vision. *Biolinguistics*, 4(2–3), 174–200.
- Bonhage, C. E., Meyer, L., Gruber, T., Friederici, A. D., & Mueller, J. L. (2017). Oscillatory EEG dynamics underlying automatic chunking during sentence processing. *NeuroImage*, 152, 647–657.
- Bosker, H. R. (2017). Accounting for rate-dependent category boundary shifts in speech perception. *Attention, Perception, & Psychophysics*, 79, 333–343.
- Bosseler, A. N., Taulu, S., Pihko, E., Makela, J. P., Imada, T., Ahonen, A., & Kuhl, P. K. (2013). Theta brain rhythms index perceptual narrowing in infant speech perception. *Frontiers in Psychology*, 4, 690.
- Boucher, V. J., Gilbert, A. C., & Jemel, B. (2019). The role of low-frequency neural oscillations in speech processing: Revisiting delta entrainment. *Journal of Cognitive Neuroscience*, 31(8), 1–11.
- Bourguignon, M., De Tieghe, X., de Beeck, M. O., Ligot, N., Paquier, P., Van Bogaert, P., ... Jousmaki, V. (2013). The pace of prosodic phrasing couples the listener's cortex to the reader's voice. *Human Brain Mapping*, 34(2), 314–326.
- Buzsaki, G. (2006). *Rhythms of the brain*. New York, NY: Oxford University Press.
- Buzsaki, G. (2019). *The brain from inside out*. New York, NY: Oxford University Press.
- Calderone, D. J., Lakatos, P., Butler, P. D., & Castellanos, F. X. (2014). Entrainment of neural oscillations as a modifiable substrate of attention. *Trends in Cognitive Sciences*, 18(6), 300–309.

- Chao, Z. C., Takaura, K., Wang, L., Fujii, N., & Dehaene, S. (2018). Large-scale cortical networks for hierarchical prediction and prediction error in the primate brain. *Neuron*, *100*(5), 1252–1266.e3.
- Cummins, F. (2012). Oscillators and syllables: A cautionary note. *Frontiers in Psychology*, *3*, 364.
- 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.
- Dikker, S., Rabagliati, H., & Pyllkänen, L. (2009). Sensitivity to syntax in visual cortex. *Cognition*, *110*, 293–321.
- Dilley, L. C., & Pitt, M. A. (2010). Altering context speech rate can cause words to appear or disappear. *Psychological Science*, *21*, 1664–1670.
- Ding, N., & He, H. (2016). Rhythm of silence. *Trends in Cognitive Sciences*, *20*(2), 82–84.
- 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.
- Ding, N., & Simon, J. Z. (2014). Cortical entrainment to continuous speech: Functional roles and interpretations. *Frontiers in Human Neuroscience*, *8*, 311.
- Dingemans, M., Blasi, D. E., Lupyan, G., Christiansen, M. H., & Monaghan, P. (2015). Arbitrariness, iconicity, and systematicity in language. *Trends in Cognitive Sciences*, *19*(10), 603–615.
- Doelling, K., Arnal, L., Ghizta, O., & Poeppel, D. (2014). Acoustic landmarks drive delta-theta oscillations to enable speech comprehension by facilitating perceptual parsing. *Neuroimage*, *85*(2), 761–768.
- Engel, A. K., & Fries, P. (2010). Beta-band oscillations—Signalling the status quo? *Current Opinion in Neurobiology*, *20*(2), 156–165.
- Fernald, A., & McRoberts, G. (1996). Prosodic bootstrapping: A critical analysis of the argument and the evidence. In J. L. Morgan & K. Demuth (Eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 365–388). New York, NY: Psychology Press.
- Fodor, J. D. (1998). Learning to parse? *Journal of Psycholinguistic Research*, *27*(2), 285–319.
- 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.
- Friederici, A., Wang, Y., Herrmann, C., Maess, B., & Oertel, U. (2000). Localization of early syntactic processes in frontal and temporal cortical areas: A magnetoencephalographic study. *Human Brain Mapping*, *11*, 1–11.
- Friederici, A. D., & Singer, W. (2015). Grounding language processing on basic neurophysiological principles. *Trends in Cognitive Sciences*, *19*, 329–338.
- Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *360*(1456), 815–836.
- Gervain, J. (2015). Plasticity in early language acquisition: The effects of prenatal and early childhood experience. *Current Opinion in Neurobiology*, *35*, 13–20.
- Ghitza, O. (2011). Linking speech perception and neurophysiology: Speech decoding guided by cascaded oscillators locked to the input rhythm. *Frontiers in Psychology*, *2*, 130.
- Ghitza, O., & Greenberg, S. (2009). On the possible role of brain rhythms in speech perception: Intelligibility of time-compressed speech with periodic and aperiodic insertions of silence. *Phonetica*, *66*(1–2), 113–126.
- Giraud, A. L., & Poeppel, D. (2012). Cortical oscillations and speech processing: Emerging computational principles and operations. *Nature Neuroscience*, *15*(4), 511–517.
- Goswami, U., & Leong, V. (2013). Speech rhythm and temporal structure: Converging perspectives? *Laboratory Phonology*, *4*(1), 2013–0004.
- Greenberg, S. (2001). What are the essential cues for understanding spoken language? *The Journal of the Acoustical Society of America*, *109*(5), 2382–2382.
- Grillo, N., Costa, J., Fernandes, B., & Santi, A. (2015). Highs and lows in English attachment. *Cognition*, *144*, 116–122.
- 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.
- Hale, J. (2001). *A probabilistic Earley parser as a psycholinguistic model*. Proceedings of the second meeting of the North American Chapter of the Association for Computational Linguistics on Language Technologies.
- Hale, J., Dyer, C., Kuncoro, A., & Brennan, J. R. (2018). *Finding syntax in human encephalography with beam search*. Proceedings of the 56th annual meeting of the Association for Computational Linguistics, Melbourne, Australia (pp. 2727–2736). ACL.
- Halle, M., & Stevens, K. (1962). Speech recognition: A model and a program for research. *IEEE Transactions on Information Theory*, *8*(2), 155–159.
- Hanslmayr, S., Aslan, A., Staudigl, T., Klimesch, W., Herrmann, C. S., & Bäuml, K.-H. (2007). Prestimulus oscillations predict visual perception performance between and within subjects. *Neuroimage*, *37*, 1465–1473.
- Hemforth, B., Fernandez, S., Clifton, C., Frazier, L., Konieczny, L., & Walter, M. (2015). Relative clause attachment in German, English, Spanish and French: Effects of position and length. *Lingua*, *166*, 43–64.
- Herrmann, B., Maess, B., Hasting, A. S., & Friederici, A. D. (2009). Localization of the syntactic mismatch negativity in the temporal cortex: An MEG study. *Neuroimage*, *48*, 590–600.
- Herrmann, C. S., Munk, M. H., & Engel, A. K. (2004). Cognitive functions of gamma-band activity: Memory match and utilization. *Trends in Cognitive Sciences*, *8*(8), 347–355.
- Hickok, G., Farahbod, H., & Saberi, K. (2015). The rhythm of perception entrainment to acoustic rhythms induces subsequent perceptual oscillation. *Psychological Science*, *26*(7), 1006–1013.
- Hwang, H., & Steinhauer, K. (2011). Phrase length matters: The interplay between implicit prosody and syntax in Korean “garden path” sentences. *Journal of Cognitive Neuroscience*, *23*(11), 3555–3575.
- Iemi, L., Busch, N. A., Laudini, A., Haegens, S., Samaha, J., Villringer, A., & Nikulin, V. V. (2019). Multiple mechanisms link prestimulus neural oscillations to sensory responses. *Elife*, *8*, e43620.
- Isobe, M. (2007). *The acquisition of nominal compounding in Japanese: A parametric approach*. Proceedings of the 2nd conference on Generative Approaches to Language Acquisition North America.
- Jusczyk, P. W. (1997). *The discovery of spoken language*. Cambridge: MIT Press.

- Kayser, C., Wilson, C., Safaai, H., Sakata, S., & Panzeri, S. (2015). Rhythmic auditory cortex activity at multiple timescales shapes stimulus-response gain and background firing. *Journal of Neuroscience*, 35(20), 7750–7762.
- Keitel, A., & Gross, J. (2016). Individual human brain areas can be identified from their characteristic spectral activation fingerprints. *PLoS Biology*, 14(6), e1002498.
- Kelso, J. A., Saltzman, E. L., & Tuller, B. (1986). The dynamical perspective on speech production: Data and theory. *Journal of Phonetics*, 14(1), 29–59.
- Kim, J. S., & Chung, C. K. (2008). Language lateralization using MEG beta frequency desynchronization during auditory oddball stimulation with one-syllable words. *Neuroimage*, 42, 1499–1507.
- Kösem, A., Basirat, A., Azizi, L., & van Wassenhove, V. (2016). High-frequency neural activity predicts word parsing in ambiguous speech streams. *Journal of Neurophysiology*, 116(6), 2497–2512.
- Kösem, A., Bosker, H. R., Takashima, A., Meyer, A., Jensen, O., & Hagoort, P. (2018). Neural entrainment determines the words we hear. *Current Biology*, 28(18), 2867–2875.e3.
- Kuhl, P. K. (2000). A new view of language acquisition. *Proceedings of the National Academy of Sciences of the United States of America*, 97(22), 11850–11857.
- Kuhl, P. K., Ramirez, R. R., Bosseler, A., Lin, J. F., & Imada, T. (2014). Infants' brain responses to speech suggest analysis by synthesis. *Proceedings of the National Academy of Sciences of the United States of America*, 111(31), 11238–11245.
- 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.
- Lakatos, P., Shah, A. S., Knuth, K. H., Ulbert, I., Karmos, G., & Schroeder, C. E. (2005). An oscillatory hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex. *Journal of Neurophysiology*, 94(3), 1904–1911.
- Lany, J., & Saffran, J. R. (2010). From statistics to meaning: Infants' acquisition of lexical categories. *Psychological Science*, 21(2), 284–291.
- 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.
- Leonard, M. K., Baud, M. O., Sjerps, M. J., & Chang, E. F. (2016). Perceptual restoration of masked speech in human cortex. *Nature Communications*, 7, 13619.
- Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, 106(3), 1126–1177.
- Lewis, A., & Bastiaansen, M. (2015). A predictive coding framework for rapid neural dynamics during sentence-level language comprehension. *Cortex*, 68, 155–168.
- Lewis, A., Schoffelen, J.-M., Hoffmann, C., Bastiaansen, M., & Schriefers, H. (2017). Discourse-level semantic coherence influences beta oscillatory dynamics and the N400 during sentence comprehension. *Language, Cognition and Neuroscience*, 32(5), 601–617.
- Lewis, A. G., Schoffelen, J.-M., Schriefers, H., & Bastiaansen, M. (2016). A predictive coding perspective on beta oscillations during sentence-level language comprehension. *Frontiers in Human Neuroscience*, 10, 85.
- Lizarazu, M., Lallier, M., Molinaro, N., Bourguignon, M., Paz-Alonso, P. M., Lerma-Usabiaga, G., & Carreiras, M. (2015). Developmental evaluation of atypical auditory sampling in dyslexia: Functional and structural evidence. *Human Brain Mapping*, 36(12), 4986–5002.
- MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, 101(4), 676–703.
- Mahmoudzadeh, M., Dehaene-Lambertz, G., Fournier, M., Kongolo, G., Goudjil, S., Dubois, J., ... Wallois, F. (2013). Syllabic discrimination in premature human infants prior to complete formation of cortical layers. *Proceedings of the National Academy of Sciences*, 110(12), 4846–4851.
- Mai, G., Minett, J. W., & Wang, W. S. (2016). Delta, theta, beta, and gamma brain oscillations index levels of auditory sentence processing. *Neuroimage*, 133, 516–528.
- Maltseva, I., Geissler, H. G., & Başar, E. (2000). Alpha oscillations as an indicator of dynamic memory operations – anticipation of omitted stimuli. *International Journal of Psychophysiology*, 36(3), 185–197.
- Männel, C., & Friederici, A. D. (2009). Pauses and intonational phrasing: ERP studies in 5-month-old German infants and adults. *Journal of Cognitive Neuroscience*, 21(10), 1988–2006.
- Männel, C., Schipke, C. S., & Friederici, A. D. (2013). The role of pause as a prosodic boundary marker: Language ERP studies in German 3- and 6-year-olds. *Developmental Cognitive Neuroscience*, 5, 86–94.
- Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology*, 10(1), 29–63.
- Martin, A. E. (2016). Language processing as cue integration: Grounding the psychology of language in perception and neurophysiology. *Frontiers in Psychology*, 7, 120.
- Martin, A. E., & Doumas, L. A. (2017). A mechanism for the cortical computation of hierarchical linguistic structure. *PLoS Biology*, 15(3), e2000663.
- Martin, A. E., & Doumas, L. A. (2019). Predicate learning in neural systems: Using oscillations to discover latent structure. *Current Opinion in Behavioral Sciences*, 29, 77–83.
- Mathewson, K. E., Gratton, G., Fabiani, M., Beck, D. M., & Ro, T. (2009). To see or not to see: Prestimulus alpha phase predicts visual awareness. *Journal of Neuroscience*, 29, 2725–2732.
- Mathewson, K. E., Prudhomme, C., Fabiani, M., Beck, D. M., Lleras, A., & Gratton, G. (2012). Making waves in the stream of consciousness: Entraining oscillations in EEG alpha and fluctuations in visual awareness with rhythmic visual stimulation. *Journal of Cognitive Neuroscience*, 24(12), 2321–2333.
- Mesgarani, N., Cheung, C., Johnson, K., & Chang, E. F. (2014). Phonetic feature encoding in human superior temporal gyrus. *Science*, 343(6174), 1006–1010.
- Meyer, L. (2017). The neural oscillations of speech processing and language comprehension: State of the art and emerging mechanisms. *European Journal of Neuroscience*, 48(7), 2609–2621.
- Meyer, L., Elsner, A., Turker, S., Kuhnke, P., & Hartwigsen, G. (2018). Perturbation of left posterior prefrontal cortex modulates topdown processing in sentence comprehension. *NeuroImage*, 181, 598–604.
- Meyer, L., & Gumbert, M. (2018). Synchronization of electrophysiological responses with speech benefits syntactic information processing. *Journal of Cognitive Neuroscience*, 30(8), 1066–1074.
- Meyer, L., Henry, M. L., Schmuck, N., Gaston, P., & Friederici, A. (2016). Linguistic bias modulates interpretation of speech

- via neural delta-band oscillations. *Cerebral Cortex*, 27(9), 4293–4302.
- Molinaro, N., & Lizarazu, M. (2017). Delta (but not theta)-band cortical entrainment involves speech-specific processing. *European Journal of Neuroscience*, 48(7), 2642–2650.
- Molinaro, N., Lizarazu, M., Lallier, M., Bourguignon, M., & Carreiras, M. (2016). Out-of-synchrony speech entrainment in developmental dyslexia. *Human Brain Mapping*, 37, 2767–2783.
- Moon, C., Lagercrantz, H., & Kuhl, P. K. (2013). Language experienced in utero affects vowel perception after birth: A two-country study. *Acta Paediatrica*, 102(2), 156–160.
- Neuling, T., Rach, S., Wagner, S., Wolters, C. H., & Herrmann, C. S. (2012). Good vibrations: Oscillatory phase shapes perception. *Neuroimage*, 63(2), 771–778.
- Nieuwland, M. S., Politzer-Ahles, S., Heyselaar, E., Segaert, K., Darley, E., Kazanina, N., ... Mézière, D. (2018). Large-scale replication study reveals a limit on probabilistic prediction in language comprehension. *eLife*, 7, e33468.
- Nieuwland, M. S., & Van Berkum, J. J. (2006). When peanuts fall in love: N400 evidence for the power of discourse. *Journal of Cognitive Neuroscience*, 18(7), 1098–1111.
- Nourski, K. V., Steinschneider, M., Rhone, A. E., Oya, H., Kawasaki, H., Howard, M. A., 3rd, & 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.
- Obleser, J., & Kayser, C. (2019). Neural entrainment and attentional selection in the listening brain. *Trends in Cognitive Sciences*, 23(11), 913–926.
- Ortiz-Mantilla, S., Hamalainen, J. A., Realpe-Bonilla, T., & Benasich, A. A. (2016). Oscillatory dynamics underlying perceptual narrowing of native phoneme mapping from 6 to 12 months of age. *Journal of Neuroscience*, 36(48), 12095–12105.
- Park, H., Ince, R. 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.
- 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.
- Pellegrino, F., Coupé, C., & Marsico, E. (2011). Across-language perspective on speech information rate. *Language*, 87(3), 539–558.
- Pikovsky, A., Rosenblum, M., & Kurths, J. (2003). *Synchronization: A universal concept in nonlinear sciences* (Vol. 12). Cambridge: Cambridge University Press.
- Poepfel, D., & Monahan, P. J. (2011). Feedforward and feedback in speech perception: Revisiting analysis by synthesis. *Language and Cognitive Processes*, 26(7), 935–951.
- Pöppel, E. (1997). A hierarchical model of temporal perception. *Trends in Cognitive Sciences*, 1(2), 56–61.
- Reinacher, M., Becker, R., Villringer, A., & Ritter, P. (2009). Oscillatory brain states interact with late cognitive components of the somatosensory evoked potential. *Journal of Neuroscience Methods*, 183(1), 49–56.
- Rimmele, J. M., Morillon, B., Poepfel, D., & Arnal, L. H. (2018). Proactive sensing of periodic and aperiodic auditory patterns. *Trends in Cognitive Sciences*, 22(10), 870–882.
- Roll, M., Lindgren, M., Alter, K., & Horne, M. (2012). Time-driven effects on parsing during reading. *Brain and Language*, 121, 267–272.
- Roopun, A. K., Kramer, M. A., Carracedo, L. M., Kaiser, M., Davies, C. H., Traub, R. D., ... Whittington, M. A. (2008). Period concatenation underlies interactions between gamma and beta rhythms in neocortex. *Frontiers in Cellular Neuroscience*, 1, 1.
- Sarnthein, J., Petsche, H., Rappelsberger, P., Shaw, G. L., & von Stein, A. (1998). Synchronization between prefrontal and posterior association cortex during human working memory. *Proceedings of the National Academy of Sciences of the United States of America*, 95(12), 7092–7096.
- Saussure, F. D. (1916). *Course in general linguistics* (W. Baskin, Trans.). London: Fontana/Collins, 74.
- Schremm, A., Horne, M., & Roll, M. (2015). Brain responses to syntax constrained by time-driven implicit prosodic phrases. *Journal of Neurolinguistics*, 35, 68–84.
- Schroeder, C. E., & Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. *Trends in Neurosciences*, 32(1), 9–18.
- Schroeder, C. E., Wilson, D. A., Radman, T., Scharfman, H., & Lakatos, P. (2010). Dynamics of active sensing and perceptual selection. *Current Opinion in Neurobiology*, 20, 172–176.
- Sedivy, J., & Spivey-Knowlton, M. (1994). The use of structural, lexical, and pragmatic information in parsing attachment ambiguities. In C. Clifton Jr., L. Frazier, & K. Rayner (Eds.), *Perspectives on sentence processing*. Hove: Psychology Press.
- Seidl, A. (2007). Infants' use and weighting of prosodic cues in clause segmentation. *Journal of Memory and Language*, 57(1), 24–48.
- Sherman, M. T., Kanai, R., Seth, A. K., & VanRullen, R. (2016). Rhythmic influence of top-down perceptual priors in the phase of prestimulus occipital alpha oscillations. *Journal of Cognitive Neuroscience*, 28(9), 1318–1330.
- 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.
- Swets, B., Desmet, T., Hambrick, D. Z., & Ferreira, F. (2007). The role of working memory in syntactic ambiguity resolution: A psychometric approach. *Journal of Experimental Psychology*, 136, 64–81.
- Szymanski, F. D., Rabinowitz, N. C., Magri, C., Panzeri, S., & Schnupp, J. W. (2011). The laminar and temporal structure of stimulus information in the phase of field potentials of auditory cortex. *Journal of Neuroscience*, 31(44), 15787–15801.
- VanRullen, R. (2016). Perceptual cycles. *Trends in Cognitive Sciences*, 20(10), 723–735.
- 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, 60.
- von Stein, A., Rappelsberger, P., Sarnthein, J., & Petsche, H. (1999). Synchronization between temporal and parietal cortex during multimodal object processing in man. *Cerebral Cortex*, 9(2), 137–150.
- Wales, R., & Toner, H. (1979). *Intonation and ambiguity. Sentence processing: Psycholinguistic studies presented to Merrill Garrett*. Hillsdale, NJ: Erlbaum.
- Wang, L., Jensen, O., van den Brink, D., Weder, N., Schoffelen, J. M., Magyari, L., ... Bastiaansen, M. (2012). Beta oscillations relate to the N400m during language comprehension. *Human Brain Mapping*, 33, 2898–2912.
- Webman-Shafran, R., & Fodor, J. D. (2016). Phrase length and prosody in on-line ambiguity resolution. *Journal of Psycholinguistic Research*, 45(3), 447–474.

- Weiss, S., & Mueller, H. M. (2012). "Too many betas do not spoil the broth": The role of beta brain oscillations in language processing. *Frontiers in Psychology*, 3, 201.
- Weisz, N., Wühle, A., Monittola, G., Demarchi, G., Frey, J., Popov, T., & Braun, C. (2014). Prestimulus oscillatory power and connectivity patterns predispose conscious somatosensory perception. *Proceedings of the National Academy of Sciences*, 111, E417–E425.
- Wiedmann, N., & Winkler, S. (2015). The influence of prosody on children's processing of ambiguous sentences. In S. Winkler (Ed.), *Ambiguity: Language and communication* (pp. 185–197). Berlin: De Gruyter.
- Wöstmann, M., Waschke, L., & Obleser, J. (2019). Prestimulus neural alpha power predicts confidence in discriminating identical auditory stimuli. *European Journal of Neuroscience*, 49, 94–105.
- Zoefel, B., & VanRullen, R. (2015). Selective perceptual phase entrainment to speech rhythm in the absence of spectral energy fluctuations. *Journal of Neuroscience*, 35(5), 1954–1964.