

Cortical Tracking of Speech: Toward Collaboration between the Fields of Signal and Sentence Processing

Eleonora J. Beier¹, Suphasiree Chantavarin^{1,2}, Gwendolyn Rehrig¹,
Fernanda Ferreira¹, and Lee M. Miller¹

Abstract

■ In recent years, a growing number of studies have used cortical tracking methods to investigate auditory language processing. Although most studies that employ cortical tracking stem from the field of auditory signal processing, this approach should also be of interest to psycholinguistics—particularly the subfield of sentence processing—given its potential to provide insight into dynamic language comprehension processes. However, there has been limited collaboration between these fields, which we suggest is partly because of differences in theoretical background and methodological constraints, some mutually exclusive. In this paper, we first review the theories and methodological constraints that have historically been prioritized in each field and provide concrete examples of how some of these constraints may be reconciled. We then elaborate on how further collaboration

between the two fields could be mutually beneficial. Specifically, we argue that the use of cortical tracking methods may help resolve long-standing debates in the field of sentence processing that commonly used behavioral and neural measures (e.g., ERPs) have failed to adjudicate. Similarly, signal processing researchers who use cortical tracking may be able to reduce noise in the neural data and broaden the impact of their results by controlling for linguistic features of their stimuli and by using simple comprehension tasks. Overall, we argue that a balance between the methodological constraints of the two fields will lead to an overall improved understanding of language processing as well as greater clarity on what mechanisms cortical tracking of speech reflects. Increased collaboration will help resolve debates in both fields and will lead to new and exciting avenues for research. ■

INTRODUCTION

Recent years have seen a growing interest in the cortical tracking of speech as a potential measure of acoustic, linguistic, and cognitive processing (Meyer, Sun, & Martin, 2020; Obleser & Kayser, 2019; Kösem & van Wassenhove, 2017; Meyer, 2018; see Tyler, 2020). The terms “cortical tracking” or “speech tracking” loosely refer to continuous neural activity that is somehow time-locked to ongoing events in the speech signal. According to one common interpretation, cortical tracking reflects the tendency for neural oscillations to align, or phase-lock, with quasiperiodic features in the speech signal. These quasiperiodic elements of speech can be acoustic, such as the fluctuations in the amplitude envelope associated with syllables (Doelling, Arnal, Ghitza, & Poeppel, 2014; Peelle, Gross, & Davis, 2013) or linguistic representations generated in the mind of the listener, such as syntactic phrase boundaries (Meyer, Henry, Gaston, Schmuck, & Friederici, 2017; Ding, Melloni, Zhang, Tian, & Poeppel, 2016). Researchers adopting this approach often refer to cortical tracking as “neural entrainment” (Obleser & Kayser,

2019; see later sections for further discussion of terminology and debates in this field). It has been proposed that entrainment may contribute to improved speech processing and language comprehension—for instance, by instantiating temporal predictions that enable segmentation of the continuous speech signal into units at several timescales (Keitel, Gross, & Kayser, 2018; Kösem et al., 2018; Meyer & Gumbert, 2018; Meyer et al., 2017; Ding et al., 2016; Zoefel & VanRullen, 2015; Doelling et al., 2014; Peelle et al., 2013; Giraud & Poeppel, 2012; Peelle & Davis, 2012; Ahissar et al., 2001).

Cortical tracking methods should therefore be of great interest to researchers studying sentence processing from a psycholinguistic perspective. Sentence processing research makes frequent use of ERPs to draw inferences about neural responses to isolated, discrete events such as word onsets or sentence boundaries (Swaab, Ledoux, Camblin, & Boudewyn, 2012; Kutas & Federmeier, 2011; Kutas, Van Petten, & Kluender, 2006) as well as time-frequency analyses of EEG oscillatory power at specific bands (Prystauka & Lewis, 2019). However, the field of sentence processing has yet to fully incorporate cortical tracking as a tool to investigate language processing mechanisms continuously, rather than at discrete epochs.

¹University of California, Davis, ²Chulalongkorn University, Bangkok, Thailand

As we will argue, combining measures of cortical tracking with typical psycholinguistic paradigms may help resolve long-standing debates and distinguish between competing theories of language processing, while also making use of continuous EEG data that are typically treated as noise in ERP paradigms (a point we discuss further in the section titled Contributions to Psycholinguistics).

In general, there has been limited collaboration between signal processing neuroscientists who use cortical tracking paradigms and psycholinguists, despite the fact that both fields share the common goal of elucidating how listeners process spoken language. To be more specific, the research areas that are most relevant for our purposes are the study of human sentence processing, which is a subfield of psycholinguistics, and the study of auditory signal processing, which is often carried out by perceptual neurophysiologists and engineers who are interested in how the brain transforms auditory signals and who may implement cortical tracking in their research methods. These research topics intersect when the auditory signal comprises spoken sentences. For convenience, we will refer to these research areas as “sentence processing” and “signal processing” throughout, with the caveat that the fields are not mutually exclusive; there are psycholinguists who employ cortical tracking methods to study sentence processing (e.g., Song & Iverson, 2018; Martin & Dumas, 2017; Meyer et al., 2017), signal processing researchers who are interested in the linguistic properties of the speech signal (e.g., Ding et al., 2016), and investigators with clear interest in both fields who already conduct studies that incorporate the methodological compromises we suggest later on (e.g., Giraud & Poeppel, 2012; Peelle & Davis, 2012; Obleser & Kotz, 2011).

Despite these notable examples of interaction, the two fields remain largely independent. Part of the reason for this limited collaboration stems from the different ways these two fields conceptualize and define language processing; whereas sentence processing research focuses on the cognitive and linguistic representations formed during comprehension, signal processing research treats speech as an example of a complex auditory signal and seeks to characterize the system that transforms that input signal into an output signal or response (e.g., Rimmele, Morillon, Poeppel, & Arnal, 2018; Morillon & Schroeder, 2015). An additional hurdle to collaboration stems from the largely different methodological constraints that the two fields are primarily concerned with. Some of these constraints are because of the theoretical grounds upon which research is conducted as well as limitations in the current methods for data acquisition; these are sometimes at odds across fields and can be difficult to reconcile. However, we argue that most constraints are reconcilable and that both fields would benefit from incorporating aspects of each other’s methodology and theoretical constructs. Just as sentence processing research would be able to answer more detailed questions

about linguistic processing by using more of the continuous EEG data through cortical tracking methods, signal processing research would gain a better understanding of the role of cortical tracking in the processing of speech by controlling and manipulating the linguistic features of the stimulus, which are often underspecified in current studies.

In this paper, we first review the theoretical background and the methodological constraints that have historically been prioritized in the study of sentence processing and the neuroscience of auditory signal processing. We then explore in more detail what further collaboration between these two fields could bring, highlighting how cortical tracking methods could be used to improve our understanding of continuous sentence comprehension and how paradigms from psycholinguistics may in turn improve our understanding of the transformations listeners apply to speech as an input signal. Finally, we will provide concrete ideas for how to reconcile each field’s constraints to reach these goals. We conclude by arguing that, although no perfect experiment can be constructed that will satisfy all constraints, a better mutual understanding of each field’s approach will greatly improve experimental designs in both fields and open up new exciting avenues for research.

OVERVIEW OF RESEARCH ON SIGNAL AND SENTENCE PROCESSING

Cortical Tracking in Signal Processing Research

Cortical tracking refers to the observed alignment of rhythmic neural activity with an external periodic or quasi-periodic stimulus. It has been observed in response to both visual and auditory stimuli (Besle et al., 2011; Gomez-Ramirez et al., 2011; Luo, Liu, & Poeppel, 2010; Lakatos, Karmos, Mehta, Ulbert, & Schroeder, 2008; Lakatos et al., 2005). Although this phenomenon has received growing interest in the past decade, there is still considerable debate as to the neural causes of cortical tracking and the role it may play in various aspects of cognition (e.g., attention, temporal prediction, speech processing). Specifically, debate has centered around whether cortical tracking results from the phase-locking of ongoing, endogenous oscillations (Calderone, Lakatos, Butler, & Castellanos, 2014; Doelling et al., 2014; Giraud & Poeppel, 2012; Peelle & Davis, 2012) or whether it is the epiphenomenal result of repeated evoked responses (e.g., reflexive responses to the stimulus). There is also debate about whether cortical tracking serves a functional role in attention and speech processing, such as by acting as a mechanism for temporal prediction and segmentation (Köseme et al., 2018; Morillon & Schroeder, 2015; Calderone et al., 2014; Doelling et al., 2014; Giraud & Poeppel, 2012; Peelle & Davis, 2012; Lakatos et al., 2008), or whether it is merely a passive response to other mechanisms (Rimmele et al., 2018; Ding & Simon, 2014).

These different viewpoints are sometimes reflected in the terminology used to describe cortical tracking: Studies that argue or assume that cortical tracking involves the synchronization of ongoing oscillations often use the term “neural entrainment,” as opposed to the more neutral terms cortical/neural tracking, phase coding, speech tracking, and so on (Obleser & Kayser, 2019). In this paper, we are agnostic about the neural origins and cognitive role of this phenomenon, and we therefore use the neutral term “cortical tracking” throughout.

Cortical Tracking of Speech

In this section, we briefly review the empirical and theoretical background on cortical tracking of speech. Interested readers are encouraged to explore reviews by Meyer (2018), Kösem and van Wassenhove (2017), and Obleser and Kayser (2019) for comprehensive explanations.

Much of the empirical work on cortical tracking of speech has been based on the view that it arises from the phase-locking of ongoing oscillations to events in the speech signal and that it represents an attentional or attention-like mechanism whereby periods of high neuronal excitability align with the temporal occurrence of the stimulus events to maximize processing efficiency (Giraud & Poeppel, 2012; Peelle & Davis, 2012; Lakatos et al., 2008). The idea of cortical tracking as an attentional mechanism has been theorized to support beat perception in music through the Dynamic Attending Theory (Large & Jones, 1999; Large & Kolen, 1994). According to this account, the perception of rhythm relies on the dynamic allocation of attention to points in time when the next beat is predicted to occur. The synchronization of neural oscillations to the beat is therefore a mechanism that allows for temporal predictions. In addition, cortical tracking has been observed for imagined groupings of acoustically identical periodic beats (Nozaradan, Peretz, Missal, & Mouraux, 2011), providing further evidence for its role in the perception of hierarchically organized rhythmic patterns, or meter. It also indicates that cortical tracking may reflect the segmentation of stimuli into larger units not necessarily represented acoustically (but see Meyer et al., 2020, and Tyler, 2020, for important considerations on the difference between cortical tracking of acoustic features as opposed to endogenously generated representations).

It has been argued that this type of attentional mechanism, allowing for temporal predictions through the synchronization of ongoing neural oscillations, is used for the perception of speech as well (Meyer, 2018; Ding et al., 2016; Doelling et al., 2014; Giraud & Poeppel, 2012; Peelle & Davis, 2012). Speech is a temporal signal consisting of quasiperiodic events at multiple timescales. Cortical tracking has been observed in response to many acoustic and linguistic properties of speech, including

syllabic rate (Doelling et al., 2014; Peelle et al., 2013; Luo & Poeppel, 2007), the presence of prosodic intonational boundaries (Bourguignon et al., 2013), and the presence of syntactic phrases (Meyer & Gumbert, 2018; Meyer et al., 2017; Ding et al., 2016). Given that speech requires extremely fast processing, often in noisy environments, it would be beneficial for the listener to be able to preallocate attention to the points in time where important bits of signal will likely occur. In this way, the listener can ensure that this information will coincide with maximal neuronal excitability and therefore more efficient processing (Meyer & Gumbert, 2018; Morillon & Schroeder, 2015; Peelle & Davis, 2012). Speech and music would not be unique in this respect, as temporal predictions are thought to modulate attention and facilitate processing of events at predicted time locations more broadly (Nobre & van Ede, 2018), and temporal predictions in the auditory domain in particular have demonstrated the involvement of delta oscillations (Stefanics et al., 2010) similarly to the domain of language (Ding et al., 2016).

This ability to predict where important linguistic information is likely to occur could also support the temporal segmentation of speech into its linguistic units (Doelling et al., 2014; Giraud & Poeppel, 2012), such as the formation of syntactic boundaries (Meyer et al., 2017; Ding et al., 2016). In a foundational study, Ding et al. (2016) showed evidence of cortical tracking not only to periodically presented monosyllabic words but also to the two-word phrases and the four-word sentences that these words combined into. Importantly, syntactic boundaries were not marked acoustically, suggesting that the observed cortical tracking reflected a mental representation rather than an acoustic property, similar to what has been found for an imaginary meter (Nozaradan et al., 2011). If cortical tracking plays a functional role in actively predicting temporal events and in segmenting speech into units, then it may be an essential aspect of speech perception and language comprehension (Schwartz & Kotz, 2013; Peelle & Davis, 2012; Kotz & Schwartz, 2010; Ghitza & Greenberg, 2009).

Although this view has recently gained popularity, some have pointed out that what may appear as the entrainment of intrinsic oscillations to speech may in fact result from a series of evoked responses (e.g., Nora et al., 2020) or the by-product of attentional gain mechanisms (e.g., Kerlin, Shahin, & Miller, 2010; for discussions, see Ding & Simon, 2014, and Kösem & van Wassenhove, 2017). Cortical tracking would therefore consist of a passive response that plays no functional role in comprehension. Despite the difficulty of ruling out this possibility, several recent studies have provided evidence for oscillatory models that entail a more active role of cortical tracking in speech comprehension (Keitel et al., 2018; Kösem et al., 2018; Meyer & Gumbert, 2018; Zoefel, Archer-Boyd, & Davis, 2018; Meyer et al., 2017; Ding et al., 2016; Ding, Chatterjee, & Simon, 2014;

Doelling et al., 2014; Peelle et al., 2013). The idea that cortical tracking entirely reflects evoked responses is also inconsistent with the finding that neural oscillations persist at the stimulus frequency for several cycles even after the stimulus ends (Calderone et al., 2014). A third possibility is that cortical tracking does reflect the phase-locking of endogenous neuronal oscillations but that this is not itself a temporal prediction mechanism; rather, neural oscillations may constitute a processing constraint, and phase reset is induced by subcortical structures involved in the top-down temporal prediction of both periodic and aperiodic signals (Rimmele et al., 2018).

With respect to speech in particular, debate surrounds the question of whether cortical tracking reflects only low-level perception of acoustic properties of speech or whether it is actively involved in top-down speech comprehension, tracking language-specific features beyond acoustic features. One common way to address this question has been to vary the degree of speech intelligibility through acoustic manipulations, which has led to mixed results (e.g., Baltzell, Srinivasan, & Richards, 2017; Zoefel & VanRullen, 2016; Millman, Johnson, & Prendergast, 2015; Doelling et al., 2014; Peelle et al., 2013; Howard & Poeppel, 2010; Ahissar et al., 2001). One possibility is that multiple mechanisms are involved, such that cortical tracking may play different roles and track different features depending on the frequency band and the neuroanatomical source (Ding & Simon, 2014; Kösem & van Wassenhove, 2017; Zoefel & VanRullen, 2015). For instance, it has been recently proposed that cortical tracking consists of both “entrainment proper” (phase-locking to acoustic periodicities of the signal) and “intrinsic synchronicities” reflecting the endogenous generation of linguistic structure and predictions (see Meyer et al., 2020, for a clear distinction of these terms). Thus, although cortical tracking of speech has recently received much attention, its source and role are still debated, and there is still much to learn regarding its functional role in language comprehension. We will argue that cortical tracking is a useful tool for exploring psycholinguistic questions about language comprehension regardless of what neural mechanisms it may reflect and that psycholinguistic paradigms may in fact help elucidate the potential role(s) of cortical tracking in language processing and cognition more broadly.

Psycholinguistic Issues in Sentence Processing Research

Next, we will focus on the kinds of questions that have been asked in sentence processing research and the general classes of theories that have been proposed to account for psycholinguistic performance. The purpose of this section is not to provide an exhaustive review but rather to set the stage for the discussion to follow regarding how sentence processing research conceptualizes the

important considerations that go into designing empirical studies.

The fundamental question that sentence processing theories try to address is how humans understand language in real time. (Of course language production is an important area of investigation as well but is beyond the scope of this review.) As the written or spoken signal unfolds, the comprehender assigns an interpretation at a number of different levels of linguistic representation: prosodic, syntactic, semantic, and pragmatic. A core assumption is that the system is “incremental,” meaning that interpretations are assigned as the input is received and at all levels of representation (but see Christiansen & Chater, 2015; Bever & Townsend, 2001). Thus, upon hearing the word “The” at the start of an utterance, the syllable is categorized as an instance of the word “the,” it is assigned the syntactic category “determiner,” and a syntactic structure is projected positing the existence of a subject noun phrase and perhaps even an entire clause (i.e., so-called “left-corner” parsing; Abney & Johnson, 1991; Johnson-Laird, 1983). Incremental interpretation supports efficient processing because input is categorized as it is received, which avoids the need for backtracking and for holding unanalyzed material in working memory. However, incremental interpretation will often lead to “garden paths” at a number of levels of interpretation: For example, given a sequence such as “The principal spoke to the cafeteria...,” readers tend to spend a long time fixating on “cafeteria” because it is implausible as the object of “speak to,” a confusion that gets resolved once an animate noun such as “manager” is encountered (Staub, Rayner, Pollatsek, Hyönä, & Majewski, 2007); similarly, there is evidence from ERPs that any word can invoke a late positivity similar to a P600 component with cumulative syntactic effort, as measured by the number of parsing steps taken to parse the sentence before encountering the word (Hale, Dyer, Kuncoro, & Brennan, 2018).

Debate continues regarding the most compelling theoretical framework for explaining these and other processing effects (for a review, see Traxler, 2014). However, psycholinguists do agree that comprehenders eventually make use of all relevant information and that processing is constrained by the architectural properties of the overall cognitive system, including working memory constraints (e.g., Kim, Oines, & Miyake, 2018; Huettig & Janse, 2016; Swets, Desmet, Hambrick, & Ferreira, 2007). Recent models emphasize the need to account for the language system’s tendency to construct shallow, incomplete, and occasionally nonveridical representations of the input (Ferreira & Lowder, 2016; Gibson, Bergen, & Piantadosi, 2013; Ferreira, 2003). Related approaches highlight the rational nature of comprehension, which assume that readers and listeners optimally combine the input with their rational expectations to arrive at an optimal construction of the linguistic signal, which may allow for alterations to the input in accordance with

noisy channel models (Futrell, Gibson, & Levy, 2020; Gibson et al., 2013). These models are often rooted in computational algorithms that assign surprisal (the degree to which the word is expected given the preceding context) and entropy (the degree to which the word constrains upcoming linguistic content) values to each word in a sentence, reflecting how easily a word can be integrated given the left context and the overall statistics of the language (Futrell et al., 2020). Neuroimaging evidence indicates that activation in language-related brain areas correlates with difficulty as reflected by these information-theoretic measures (e.g., Russo et al., 2020; Henderson, Choi, Lowder, & Ferreira, 2016). Recent neural models of language processing locate the operation of combining two elements into a syntactic representation in Brodmann's area 44 of Broca's area, which forms a network for processing of syntactic complexity in combination with the superior temporal gyrus (Fedorenko & Blank, 2020; Zaccarella, Schell, & Friederici, 2017; for a competing view, see Matchin & Hickok, 2020).

The theoretical debates outlined above would likely be of interest to those who use cortical tracking to study auditory signal processing, just as the cortical tracking of speech is clearly relevant to psycholinguistic research. Both fields have shown interest in discovering how listeners might assign abstract linguistic structure to continuous acoustic input as it unfolds and in determining the neural correlates of spoken language processing.

BRIDGING THE TWO FIELDS

When spoken language is the signal, the studies of sentence processing and of auditory signal processing share the goal of characterizing the neural and cognitive architecture of language comprehension. However, the two areas have not extensively collaborated to address this question. Research in sentence processing makes extensive use of electrophysiological measures, especially through the use of ERPs, which have contributed greatly to our understanding of language comprehension (Swaab et al., 2012; Kutas & Federmeier, 2011; Kutas et al., 2006). Beyond the now routine use of ERPs, many psycholinguists have also adopted time–frequency analyses of EEG and magnetoencephalography data, as increases or decreases in power at various frequency bands have been found to correlate with several aspects of comprehension (for reviews, see Prystauka & Lewis, 2019; Meyer, 2018; Bastiaansen & Hagoort, 2006). The use of time–frequency analyses has enabled researchers to make fuller use of their data by including both synchronized and desynchronized neural activity, thus not discarding neural activity that is not phase-locked to a stimulus (Bastiaansen, Mazaheri, & Jensen, 2012). Yet, despite the widespread use of ERPs and time–frequency analyses of neural oscillations, the use of cortical tracking methods in particular to answer psycholinguistic questions to date is relatively rare. Importantly, cortical tracking implies a relationship

between the periodicities found in neural activity and those found in the linguistic stimuli, which is different from the types of time–frequency analyses already frequently used in psycholinguistics. Strictly speaking, a power change in a certain frequency band does not imply phase-locking (and by the same token, momentary phase coherence does not imply an ongoing oscillation). As mentioned earlier, most studies that use cortical tracking of speech stem from the fields of signal processing, auditory processing, and neuroscience.

Nonetheless, there are some notable examples of research overlapping the methods and questions of these two fields. For example, Meyer and colleagues have performed several experiments that measure cortical tracking of speech using typical psycholinguistic experimental designs to answer questions about syntactic parsing, some of which we describe later in the Contributions to Psycholinguistics section (e.g., Meyer & Gumbert, 2018; Meyer et al., 2017). Similarly, Martin and Doumas (2017) have proposed a computational model linking cortical tracking to the building of hierarchical representations of linguistic structure.

However, beyond these emerging pockets of research bridging the gap between sentence processing and signal processing, the two fields remain largely independent. One of the reasons for this may be that the two take very different approaches to the study of language processing. Although many researchers who use cortical tracking methods are interested in signal processing more broadly and consider speech to be one of many naturally occurring complex signals (e.g., Rimmele et al., 2018; Morillon & Schroeder, 2015), sentence processing research typically emphasizes the linguistic properties of language and the different levels of cognitive representations that are generated during language processing, as discussed in the previous section.

As is often the case in interdisciplinary research, a major challenge lies in the discrepancies in terminology and definitions across different literatures. In particular, studies in the two fields may sometimes even differ in their definition of language comprehension (e.g., the distinction between speech perception, processing, and comprehension; see Meyer, 2018) or may not specify the degree or level of comprehension being assessed. Language comprehension entails a range of cognitive processes and levels of representations, which are sometimes left underspecified because of shallow processing (Wang, Bastiaansen, Yang, & Hagoort, 2011, 2012; Ferreira & Patson, 2007; Ferreira, 2003; Sanford & Sturt, 2002; Christianson, Hollingworth, Halliwell, & Ferreira, 2001). Thus, researchers attempting to bridge the two fields will need to be aware of how comprehension is conceptualized across studies.

More generally, the limited collaboration may stem from the different methodological constraints that the two fields typically prioritize, because of their different theoretical backgrounds. In the following sections, we

summarize some of the methodological constraints and solutions that are often employed in signal processing studies that use cortical tracking methods and in sentence processing research and note how these constraints are sometimes at odds.

METHODOLOGICAL CONSTRAINTS ACROSS FIELDS

Cortical Tracking Constraints

In cortical tracking studies, there are several experimental constraints that commonly arise. Foremost among them is the requirement for long stretches of continuous, varied speech. This follows from the signal processing view that speech perception is a mathematically estimable operation that transforms speech into brain responses continuously, with various responses generally overlapping one another in time. A diverse family of techniques known as system identification is well suited to characterize such continuous input–output transformations. In the system-identification framework, an input signal x (e.g., a sound) is transformed through a system $f(x)$ (e.g., the brain)—which is not directly observable—to produce an output signal y (e.g., an EEG signal). By presenting a range of systematically varied x inputs to the system and measuring y , researchers can estimate $\hat{f}(x)$ to approximate the system that transforms input to output (e.g., the brain). Notice that these data-driven approaches do not necessarily presume a certain relationship between input (the signal) and output (the neural response) and therefore may be agnostic to the specific processes within the system that may contribute to the transformation. Instead, they tend to approach neural data with few a priori assumptions, to “discover” a relationship between the signal and the neural response. This is achieved by offering the system (the listener) various instances of the input signal of interest (e.g., spoken sentences) and measuring how the output signal (e.g., the EEG recording) responds differently at each moment. Characterizing the speech–brain system entails modeling this relationship.

Linguistic Stimulus Considerations

Most of the constraints signal processing researchers optimize for concern the stimulus, which serves as the input signal. The first such constraint is that there must be a certain degree of variability in the stimulus. Speech input is often modeled by its slow (less than ~ 16 Hz) power fluctuations in the acoustic envelope. The acoustic envelope reflects the perceptually salient syllabic structure of speech and empirically relates to prominent cortical ERPs, such as the N1 (Sanders & Neville, 2003). Because modeling the speech–brain system is essentially a statistical estimation problem (see Ljung, Chen, & Mu, 2020), the speech input must be varied in its properties

to sample all the possibly relevant values, and it must do so without bias (or the analysis must explicitly correct for any bias). With a lack of variability and naturalism in the speech, the estimation will either be unrepresentative, reflecting the idiosyncrasies of the specific speech corpus, or it will fail to find a relationship at all.

A second constraint is that signal processing approaches may require numerous trials. Just as insufficient stimulus variability can undermine the statistical estimation described earlier, having too little data can lead to invalid estimates or a failure to find a relationship between input and output signals. In principle, there is no limit to the parameter space of this speech–brain system, but here too, many instances of each parameter must be presented to the listener. Furthermore, as in any model estimation problem, the more “free parameters” that must be characterized, the more data are usually required. System-identification approaches therefore offer great flexibility and interpretive power, but at the cost of acquiring more data and ensuring a statistically balanced array of parameters, akin to using a Latin square experimental design. A related constraint that often arises in cortical tracking studies is the need to repeat identical segments of speech multiple times. The motivation here is the same as when creating a traditional, simple ERP: An average response to multiple identical events (say, a tone) will be a more representative estimate with a higher signal-to-noise ratio than any individual response. Unsurprisingly, in single-cell auditory neurophysiology, where a signal-processing mindset has long dominated and influenced many speech-tracking EEG investigators, repeated presentations are de rigeur (e.g., in the venerable poststimulus time histogram). In some cases, the data-driven nature of system-identification techniques might compel such averaging, but better signal-to-noise will help virtually any cortical tracking measure.

A third constraint concerns the periodicity of the auditory stimulus itself. In contrast to system identification, another class of influential cortical tracking experiments manipulates the speech signal’s acoustic and linguistic structures to be artificially periodic (e.g., Ding et al., 2016) and may result in stronger cortical tracking (Meyer et al., 2020; Alexandrou, Saarinen, Kujala, & Salmelin, 2018). From the signal processing perspective, this is beneficial because it allows more straightforward analysis of how the periodicity of speech input is “tracked” by the brain. Specifically, frequency-domain measures allow the investigator to focus only on those periodicities of interest with relatively high statistical power. However, linguistic events are not strictly periodic in time (Nolan & Jeon, 2014; see Beier & Ferreira, 2018, for a discussion). This tension has further relevance to questions of entrainment, particularly the hypothesis that intrinsic brain oscillations become phase-reset or otherwise temporally aligned with informative speech features. Experimentally, to identify entrainment, it may be useful to have well-defined periodicities as opposed to natural

speech dynamics. However, introducing such periodicities limits the ability to determine whether entrainment occurs for natural speech and, by extension, whether entrainment plays an active role in speech comprehension outside the laboratory.

Behavioral Task Considerations

In cortical tracking paradigms, the main consideration regarding the inclusion of a behavioral task is that it should not impede continuous EEG recording or contribute significant noise to the data. Cortical tracking studies may employ a passive listening paradigm (e.g., Keitel, Ince, Gross, & Kayser, 2017; Gross et al., 2013) to obtain EEG recording that is not continually interrupted by motor movements (e.g., button presses) from behavioral tasks on the assumption that spoken language is automatically processed even if there is no offline behavioral task (e.g., in ERP studies presenting auditory sentences; van Berkum, 2004), an assumption that is sometimes held in sentence processing research as well. Relatedly, investigators may omit a behavioral task to prevent unnatural processing strategies that may add noise to the EEG data (for further discussion, see Hamilton & Huth, 2020).

In some cases, however, signal processing studies do include a behavioral task to encourage participants to attend to the auditory signal. For instance, participants may be given probe words and asked to indicate whether they heard those words on a previous trial (e.g., Keitel et al., 2018; Falk, Lanzilotti, & Schön, 2017), or they may be asked to detect semantic anomalies in the sentence (Meyer & Gumbert, 2018). Alternatively, participants may be asked to count the number of words or syllables in the presented sentences (Batterink & Paller, 2017) or to press a button every *n*th-word (e.g., fourth-word) sentence (Getz, Ding, Newport, & Poeppel, 2018). To ensure attention to the acoustic properties of the signal, cortical tracking studies may include acoustic or temporal deviation tasks, which require participants to indicate when or whether the pitch or loudness changed in the speech they heard (Zoefel et al., 2018; Rimmele, Golumbic, Schröger, & Poeppel, 2015). In investigations of language comprehension as opposed to low-level acoustic perception, an offline task is sometimes included to ensure listeners interpreted the utterance successfully, such as a self-report of the number of words participants understood in the signal (Baltzell et al., 2017; Peelle et al., 2013) or comprehension questions (Weissbart, Kandylaki, & Reichenbach, 2020; Biau, Torralba, Fuentemilla, de Diego Balaguer, & Soto-Faraco, 2015). In summary, because signal processing studies are primarily concerned with characterizing the processes that lead to comprehension in real time, behavioral tasks are often viewed as tools to encourage participants to pay attention to an input signal.

Although the experimental requirements of cortical tracking studies tend to be rather technical in nature, they do illustrate why signal processing research has long placed such great emphasis on the continuous nature of speech processing: not only because this is likely how the brain works but also because this is mathematically inherent to the most common techniques (including both time series system-identification and frequency-domain analyses). The constraints also reflect the data-driven nature of signal processing approaches, which can be theory agnostic, in part because they might require a representative and unbiased sampling of the speech-brain activity to achieve a valid result.

Psycholinguistic Constraints

In a typical sentence processing experiment, participants read or listen to sentences with manipulations relevant to the theoretical question that the experiment is designed to evaluate (e.g., sentences with grammatical, semantic, or pragmatic anomalies). Researchers then compare processing of those sentences to sentences in a baseline condition that do not contain any anomaly but are otherwise identical to the experimental sentences (i.e., “minimal pairs” that are controlled for lexical, semantic, and syntactic features). A difference in averaged behavioral response between the experimental and baseline conditions (e.g., longer RTs or reading times) would indicate processing effects that are because of the linguistic manipulation. Importantly, the experiment is designed to control for extraneous variables and to ensure that the measures reflect specific cognitive and neural mechanisms that underlie language comprehension. To address these concerns, a set of guidelines for designing the experimental stimuli and the behavioral tasks has become mainstream in psycholinguistics over the years.

Linguistic Stimulus Considerations

In sentence processing research, linguistic stimuli are typically controlled for factors that are known to influence processing to rule out potential confounds. For example, the length and frequency of words used can affect the magnitude and timing of ERPs (Strijkers, Costa, & Thierry, 2010; Hauk & Pulvermüller, 2004; King & Kutas, 1998), and therefore, word frequency is either controlled at the stimulus creation stage (e.g., selecting words with a similar frequency from a database) or statistically controlled by including frequency in a model at the analysis stage. Relatedly, function words (e.g., prepositions, determiners) evoke different neural responses than do content words (e.g., nouns, verbs): The former evokes a left-lateralized negative shift that the latter does not (Brown, Hagoort, & ter Keurs, 1999). The amplitude of the N400 response to concrete words is larger than that for abstract words, and there is greater right-hemisphere activity for concrete words (Kounios &

Holcomb, 1994). Words with a higher orthographic neighborhood density (e.g., the number of words with a similar orthographic representation, such as “lose” and “rose”) or a phonological neighborhood density (e.g., “cat” and “kit”) evoke greater N400 negativity than those with a low neighborhood density (Winsler, Midgley, Grainger, & Holcomb, 2018; Holcomb, Grainger, & O’Rourke, 2002).

The position of the target word in a sentence can also affect processing and therefore must be taken into consideration when designing stimuli. For example, the N400 amplitude is larger for words that occur earlier, and the effect is attenuated by word frequency (Van Petten & Kutas, 1990). Low transition probability from word to word, or even syllable to syllable, can evoke the N400 response (Teinonen & Huotilainen, 2012; Kutas & Federmeier, 2011; Cunillera, Toro, Sebastián-Gallés, & Rodríguez-Fornells, 2006). To control for transition probability, stimuli are typically cloze-normed (Taylor, 1953). In the cloze procedure, participants read fragments of the experimental sentences and are asked to provide the word(s) that best completes the sentence. The proportion of a given response out of all responses provided is the cloze probability for that response and is thought to index its predictability for the preceding context. For example, if participants read the sentence “It was a breezy day so the boy went outside to fly a _____” and 90% of them responded with the word “kite”, then the word “kite” has a cloze probability of 90% and would be considered a highly predictable sentence continuation. A sentence that violates phrase structure rules or is otherwise ungrammatical (e.g., agreement errors, such as “The doctors is late for surgery”) can trigger an early left anterior negativity (ELAN) as well as the P600 component (Friederici & Meyer, 2004; Friederici, 2002). By norming stimuli for acceptability, which is a proxy for sentence grammaticality that is more accessible to naive raters (Huang & Ferreira, 2020), stimuli that do or do not evoke ERPs linked to structural violations can be selected, depending on the research question. In addition, stimuli are frequently normed for typicality, plausibility, and naturalness. Stimuli that are atypical or implausible will likely evoke N400 responses, and unnatural stimuli could evoke N400 or P600 components, depending on what aspect of the linguistic content leads raters to indicate that they seem unnatural.

Behavioral Task Considerations

In addition to carefully controlling the experimental stimuli, psycholinguistic experiments typically include an offline behavioral task to verify that participants comprehended the stimuli, such as employing true–false questions (Brothers, Swaab, & Traxler, 2017), semantic judgments (Wang, Hagoort, & Jensen, 2018), or asking participants to evaluate each sentence or narrative based on their grammaticality, acceptability, or plausibility (see

Myers, 2009). For research investigating lower-level language processing, such as acoustic representations, researchers have used simple detection tasks such as asking participants to monitor a particular phoneme and to press a response key when they hear that phoneme, but the use of detection tasks has declined in sentence processing research for various reasons (for a review, see Ferreira & Anes, 1994).

Some studies of speech comprehension do not include any explicit tasks besides passive listening (e.g., van Berkum, Brown, Zwitserlood, Kooijman, & Hagoort, 2005; see van Berkum, 2004, for a discussion). These paradigms circumvent issues associated with metalinguistic judgments and may prevent unnatural processing strategies (see Hamilton & Huth, 2020). They are also particularly advantageous when investigating language processing in special populations who may have impaired ability to perform additional behavioral tasks, such as some autistic individuals and children (Brennan, 2016). In studies of typical language processing in adults, however, comprehension tasks are often used because they allow researchers to directly measure the interpretation that participants have generated after processing the stimuli (Ferreira & Yang, 2019). It can be argued that comprehension tasks are crucial because passively presenting the linguistic stimuli to participants does not guarantee that they have fully analyzed the linguistic material and have generated an interpretation for it. Instead, they may engage in shallow processing and come away with an incomplete or even incorrect interpretation of the sentence (Ferreira & Patson, 2007; Ferreira, 2003; Christianson et al., 2001). Omitting a comprehension task may also encourage participants to adopt idiosyncratic goals during the experiment (Salverda, Brown, & Tanenhaus, 2011). Different task demands have additionally been shown to affect the extent to which people engage in basic linguistic processing such as resolving anaphors, structure building, and inferencing (Foertsch & Gernsbacher, 1994) as well as lexical prediction (Brothers et al., 2017). Finally, the behavioral task participants engage in can systematically influence language-related ERP components, including the N400 (Chwilla, Brown, & Hagoort, 1995; Bentin, Kutas, & Hillyard, 1993; Deacon, Breton, Ritter, & Vaughan, 1991) and the P600 (Schacht, Sommer, Shmuilovich, Martienz, & Martín-Loeches, 2014; Gunter & Friederici, 1999).

The presence and type of behavioral task will vary depending on the research questions and goals of the study. Nonetheless, it may be necessary to consider the kind and depth of language processing that is induced in the experiment, as motivation and strategies are known to affect language comprehension (Ferreira & Yang, 2019; Alexopoulou, Michel, Murakami, & Meurers, 2017). Well-controlled linguistic stimuli and behavioral tasks enable sentence processing researchers to draw conclusions about the specific cognitive and neural mechanisms that support language comprehension,

including the time course of these processes as well as the generated interpretation resulting from comprehension.

Reconciling Constraints across Fields

Sentence processing researchers who wish to include cortical tracking as a method and signal processing researchers who wish to employ more linguistic control in their stimuli both face the challenge of taking into account the methodological constraints of both fields. Some of the methodological constraints that psycholinguists must satisfy are difficult to reconcile with the constraints researchers who use cortical tracking methods reckon with. For example, psycholinguists carefully control the linguistic content of their stimuli (e.g., surprisal, plausibility, acceptability), which is tractable when the number of items is relatively small, and they avoid repeating the same item in an experimental session because of the effects of priming and learning. However, signal processing studies can require numerous trials, which is sometimes achieved through repeated exposure to the same item, which listeners could habituate to or overlearn. Similarly, signal processing studies often require long stretches of signal. It is difficult to create the number of unique experimental items following psycholinguistic conventions (e.g., stimulus norming) to generate the type of dataset that signal processing studies may require, and it is similarly difficult to control the linguistic properties of speech in lengthy recordings. Likewise, without an explicit comprehension task, the level of comprehension that took place and the participants' motivation are unclear, but signal processing studies may need to minimize interruptions and motor movements during EEG recording.

All interdisciplinary work requires researchers to draw from the methods used across multiple fields and ultimately agree on a common methodology. Because some methodological choices are mutually exclusive with others, no experiment can realistically meet all methodological standards. Instead, it is necessary to evaluate constraints from each field and then prioritize those that are most applicable to the research question at hand. In the event that conflicting constraints cannot be reconciled, it is worth explicitly acknowledging the validity of the constraints that could not be applied and briefly stating why they were not given priority (e.g., we could not use naturalistic stimuli because it was critical to our design to eliminate prosodic cues to syntax).

Although some constraints may be mutually exclusive—for example, the preference for more periodicity in the acoustic signal to measure cortical tracking, which conflicts with the pressure for speech to sound as natural as possible—there are steps that can be taken to reconcile other constraints without posing an undue burden on researchers. For example, even if stimuli cannot be normed or constructed to control for linguistic content, measures of some constructs (such as word frequency,

surprisal, and entropy) can be obtained for existing stimuli using computational models (Hamilton & Huth, 2020; Weissbart et al., 2020; Brennan, 2016; Willems, Frank, Nijhof, Hagoort, & van den Bosch, 2016) and then controlled for statistically. Such an approach allows researchers to quantify these measures for naturalistic stimuli, which can be used to assess cortical tracking in everyday language comprehension (Alexandrou, Saarinen, Kujala, & Salmelin, 2020; Alexandrou et al., 2018).

It is important to note that both sentence and signal processing experiments tend to use stimuli that differ from the kinds of sentences and utterances that occur in everyday language. Experiments in both areas tend to make use of read speech, which gives researchers good control over the linguistic content of the stimuli but is easier to comprehend (Uchanski, Choi, Braida, Reed, & Durlach, 1996; Payton, Uchanski, & Braida, 1994), is produced at a lower speech rate (Hirose & Kawanami, 2002; Picheny, Durlach, & Braida, 1985; Crystal & House, 1982), and has exaggerated acoustic features relative to spontaneous speech (Gross et al., 2013; Finke & Rogina, 1997; Nakajima & Allen, 1993; see Alexandrou et al., 2020, for a discussion), which may affect processing. Further work using naturalistic utterances is required both to determine the degree to which cortical tracking occurs when listening to spontaneous speech and to assess whether sentence processing models generalize to everyday speech. Nonetheless, the estimable linguistic parameters available using computational models are compatible with the system-identification techniques commonly used in signal processing research. Parameters such as spectrotemporal, phonetic/phonemic, or linguistic properties can be included in the statistical model, for instance, in approaches that elaborate on multiple regression (Sassenhagen, 2019; Di Liberto & Lalor, 2017; Crosse, Di Liberto, Bednar, & Lalor, 2016). These can take continuous values, as with the acoustic envelope, or categorical ones, as with phonemic class. Furthermore, they can be modeled as changing smoothly through time (power) or occurring only at specific, transitory moments (e.g., sentence onsets, syntactic boundaries, prosodic breaks). However, as noted before, increasing the number of parameters also requires more data.

When the research question mandates the use of linguistically controlled stimuli, as in many sentence processing studies as well as signal processing experiments that rely on stimuli with a fixed structure, it is worth noting that the stimuli need not be generated from scratch. It has been standard practice in psycholinguistics for decades to share stimuli in an appendix or in other supplementary materials, which means entire lists of well-controlled stimuli that have already been normed are readily available in published papers. In addition, there are large freely available stimulus sets that researchers have compiled for general use, for example, a data set of cloze norms for over 3,000 English sentences (Peelle

et al., 2020). Audio files for prerecorded stimuli may be found in online repositories such as the Open Science Framework or may be made available by authors upon request. Stimulus crowdsourcing—asking users of a crowdsourcing platform to generate stimuli that meet study-relevant constraints—is an additional strategy researchers can employ to alleviate the burden of stimulus generation. However, whatever the source of the corpus, the challenge of controlling stimuli, balancing factorial designs, and including filler items may lead to small numbers of trials. Although it is possible to conduct a rigorous study using cortical tracking methods with relatively few stimuli (e.g., Meyer et al., 2017, $N_{\text{items}} = 40$), achieving adequate statistical power might require compensation through using extended recording times, multiple sessions, or a larger number of test participants.

In addition, filler items are often helpful for providing listeners with a greater variability of stimulus types while at the same time reducing the likelihood that participants will overlearn the stimulus structure and, as a result, engage in shallow comprehension of the stimuli. This is particularly relevant for the assumption that neural data reflect a continuous process, as repetitive stimuli may induce shallow or atypical processing of the speech signal, in which case the function estimated by system-identification techniques may not correspond to more naturalistic processing. We return to this idea in the following section.

The behavioral task constraints across sentence processing and signal processing research can also be reconciled to quantify the level of comprehension more explicitly in cortical tracking paradigms. Adding a comprehension task to an experiment is a relatively easy way to motivate participants to engage in detailed comprehension of the stimuli. For instance, simple yes/no questions about the meaning of the stimuli encourage participants to construct elaborated semantic representations rather than attending to only the surface structure. Comprehension questions can be presented to participants in between blocks, or on the filler items only, to avoid introducing neural responses related to decision-making on the experimental trials and to ensure that motor movements do not interfere with the EEG recording. If the paradigm would not allow a comprehension task to be intermingled with the experimental trials, experimenters can motivate careful attention to the stimuli by telling participants at the start of the experiment to anticipate a memory test at the end of the session.

Speech is a continuous signal, and psycholinguists who study sentence processing are ultimately interested in understanding how listeners interpret that continuous signal. Nevertheless, there is a tendency to examine language processing through the analysis of discrete events, in part because of the conventions and analysis approaches used historically in the field (Jewett & Williston, 1971). To understand how listeners process real-time spoken input, sentence processing would benefit from adopting the methodological and analysis techniques employed in the

study of signal processing for working with continuous data. Furthermore, it is far less common in sentence processing work for researchers to consider the periodicity of the signal, or variability in the amplitude envelope, which can affect neural signals. By understanding how these acoustic features impact EEG recordings, psycholinguists conducting sentence processing experiments can better separate the relative contribution of acoustic and linguistic properties of their experimental stimuli, which would allow them to draw stronger conclusions about how linguistic features (e.g., lexical ambiguity) ultimately influence comprehension.

Finally, although cortical tracking is an inherently temporal phenomenon, linguistic attributes may strongly affect which cortical areas are involved, and thus the “spatial” pattern of tracking time series across EEG scalp channels. Many neuroimaging studies using fMRI and PET show spatial cortical activation patterns that distinguish lexical category or semantics (nouns vs. verbs, concrete vs. abstract), syntax (argument structure), and numerous other features (for examples, see Rodd, Vitello, Woollams, & Adank, 2015; Moseley & Pulvermüller, 2014; Price, 2012; Friederici, 2011). Insofar as the EEG scalp activation pattern reflects (indirectly) the locations and orientations of cortical sources, controlling such linguistic variables should lead to more consistent and representative tracking analyses.

In summary, although many of the constraints associated with conducting signal processing and sentence processing research may appear to be at odds, there are reasonable compromises that can be made to reconcile methodologies from both fields. As the difficulty of collaboration between these areas is partly because of methodological differences, some of these solutions may make it easier for both sentence processing and signal processing researchers to use cortical tracking to better understand the neural and cognitive processes underlying language comprehension. In the following sections, we further elaborate on what these two fields have to gain from this collaboration and provide more detailed examples of ways to incorporate each field’s methods and standards.

Contribution to Signal Processing Research

What can the study of signal processing, using cortical tracking methods, gain from developing stimuli that satisfy certain psycholinguistic constraints? Stimuli that are implausible, anomalous, or otherwise unnatural in some manner elicit ERP components (e.g., N400s, P600s, ELAN), which will affect oscillations if they occur in the same frequency band and therefore could contribute unwanted noise if not intentionally manipulated. Repetition of the same (or highly similar) sentence or the same syntactic frame throughout the study could also have unintended processing effects. Syntactic similarity across sentences produces structural priming, in which

structural similarity between previous sentences facilitates processing of the current sentence (Tooley, Traxler, & Swaab, 2009; Pickering & Ferreira, 2008; Bock, 1986). Unintended priming effects should be avoided because it is unclear how structural priming of this sort might influence cortical tracking of speech or alter the EEG signal in unintended ways. Another concern is that, when a sentence template is used frequently, the listener can overlearn the template and employ a behavioral strategy that undermines the study. For example, if the task is to detect a word and the target word repeatedly occurs in the same location in the sentence, listeners could successfully circumvent the intended purpose of the comprehension task by attending only to the target region of the sentence. A shallow processing strategy of this sort would allow for high performance on the task without the need to comprehend the sentence. This problem could be avoided by including filler items with different sentence structures and varying the location of the target word within the experimental items, when possible. When the syntactic structure is highly predictable because of overlearning, it may additionally attenuate the EEG signal (Tooley et al., 2009).

Importantly, controlling for the linguistic aspects of stimuli may also help researchers determine whether cortical tracking reflects evoked responses or intrinsic oscillations. If stimuli are controlled such that we can determine when and where larger ERPs should occur, variation introduced by ERPs may be more readily dissociated from variation because of intrinsic oscillations. Through the availability of computational models, many linguistic factors can be controlled for statistically (Hamilton & Huth, 2020; Weissbart et al., 2020; Brennan, 2016; Willems et al., 2016), which would allow researchers to use lengthy, naturalistic auditory stimuli that are often required in signal processing experiments and still account for linguistic constraints. Controlling for linguistic factors that are known to induce processing difficulty or to otherwise affect language processing will yield cleaner data and will provide greater context for interpreting variations in the EEG signal.

In addition to carefully controlling the stimuli, it may be worthwhile to include an explicit comprehension task to ensure participants engage in detailed comprehension while listening to the stimuli, especially if the research aim is to test the role of cortical tracking in comprehension. As discussed in the Psycholinguistic Constraints section, listeners' strategies for comprehension can vary depending on the goal and the task demands (see Ferreira & Yang, 2019), and shallow processing can sometimes lead to underspecified or even incorrect representations (Ferreira & Patson, 2007; Sanford & Sturt, 2002), thus potentially adding noise to the neural data corresponding to these cognitive processes. It is important to acknowledge that naturalistic paradigms have numerous advantages (see Hamilton & Huth, 2020; Brennan, 2016), and it is indeed not necessary to include

a comprehension task if the study's goal does not pertain to higher-level language comprehension per se (e.g., using cortical tracking methods to investigate the sequential grouping of syllables into words, not sentence or discourse-level comprehension). Nevertheless, even in this case, it is worthwhile to consider how task effects impact EEG data because any neural response that has not been accounted for has the potential to add noise. As mentioned previously, language-related ERPs have been shown to vary depending on the level of processing induced by different tasks (e.g., Chwilla et al., 1995; Bentin et al., 1993).

In selecting a behavioral task that addresses comprehension, there are a number of considerations regarding the kind of processing that is induced by the task. When appropriate, comprehension questions are ideal because they enable researchers to quantify the level of comprehension that took place, and they may be a better alternative to self-reported intelligibility because they circumvent unconscious biases. Word detection and anomaly detection tasks are useful in encouraging participants to attend to the sentences, but participants may not necessarily engage in detailed comprehension because these tasks tap into memory for the surface structure rather than the overall meaning of the sentence. Temporal or acoustic deviation tasks, in which participants indicate when or whether the pitch, loudness, or timing changed in the speech, have similar limitations to detection tasks because they only index attention to the acoustic properties of the speech signal, rather than tapping into the processing of the linguistic content of the speech stream. Furthermore, ERPs have been shown to be influenced by whether participants are instructed to pay attention to speech rhythm or syntax (Schmidt-Kassow & Kotz, 2009a, 2009b), which may also lead to overall noisier and potentially misleading EEG data.

In summary, signal processing research using cortical tracking can reap various benefits from designing stimuli and behavioral tasks that fulfill the previously described psycholinguistic constraints. If cortical tracking of speech potentially serves a functional role in speech comprehension, it would be crucial to ensure that the electrophysiological recordings reflect comprehension of the linguistic material, in which participants build syntactic structures, commit to a sentence interpretation, resolve anaphors and ambiguity, and make inferences when applicable. To this aim, including comprehension questions yields a direct measure of linguistic processing and encourages a more detailed analysis of the sentence structure and meaning. Comprehension tasks also provide an explicit goal of comprehension for participants and prevent idiosyncratic goals and strategies, which reduce noise in the data from these extraneous factors. In the data analysis stage, the use of information-theoretic measures in statistical control can be easily implemented to account for systematic noise concerning syntactic and semantic processing. A key advantage of this computational approach is that it can be used on large stretches of

naturalistic uncontrolled stimuli, bolstering the goal of investigating naturalistic language processing that is an emerging trend in both signal processing and sentence processing research (see Alexandrou et al., 2018, 2020; Hamilton & Huth, 2020; Alday, 2019; Brennan, 2016). More generally, the computational modeling approach can also elucidate the role of cortical tracking in instantiating temporal predictions, as information-theoretic modeling can identify the rich linguistic information in the signal that is coded by the brain. Signal processing researchers who are interested in using cortical tracking to study predictive coding can benefit from quantifying the depth of processing that took place because predictive processing will depend on how deeply the linguistic material was processed, which is in turn influenced by the presence and type of behavioral task (for further discussion, see Kuperberg & Jaeger, 2016). Overall, the endeavor of studying auditory signal processing can be greatly augmented by accounting for linguistic aspects in the stimuli when spoken language constitutes the signal and by employing behavioral tasks that enable explicit assessment of the depth of comprehension that took place.

Contributions to Psycholinguistics

Sentence processing research has long studied syntactic ambiguity to differentiate between contrasting theoretical accounts of cognitive parsing mechanisms. In a recent study, Meyer et al. (2017) presented ambiguous sentences such as “The client sued the murderer with the corrupt lawyer” that either did or did not include a disambiguating prosodic break before the prepositional phrase. Cortical tracking in delta-band oscillations reflected syntactic phrase groupings, which frequently—but not always—corresponded to the prosodic grouping (Bögels, Schriefers, Vonk, & Chwilla, 2011; Clifton, Carlson, & Frazier, 2002; Cutler, Dahan, & van Donselaar, 1997; Shattuck-Hufnagel & Turk, 1996; Ferreira, 1993), generating new evidence that syntactic grouping biases can override acoustic grouping cues. Cortical tracking methods could be applied further using temporarily ambiguous sentences to help differentiate between sentence parsing models.

For example, Ding et al. (2016) found that listeners showed cortical tracking to syntactic phrase boundaries (e.g., cortical tracking reflects the subject noun phrase and verb phrase boundary). If tracking of syntactic boundaries generalizes beyond the stimulus materials that Ding et al. used, then using cortical tracking to temporarily ambiguous sentences should reveal the parsing mechanisms at play. Consider the temporarily ambiguous garden-path sentence “The government plans to raise taxes failed.” The sentence fragment “The government plans to raise taxes” is ambiguous because the subject of the sentence is ambiguous (1).¹ “The government” could be the subject of the verb “plans” (1a), or “The

government plans” could be the subject of a sentence in which “government plans” is a compound noun (1b).

1. a) [S [NP The government] [VP plans ...]
- b) [S [NP The government plans] [VP ...]]

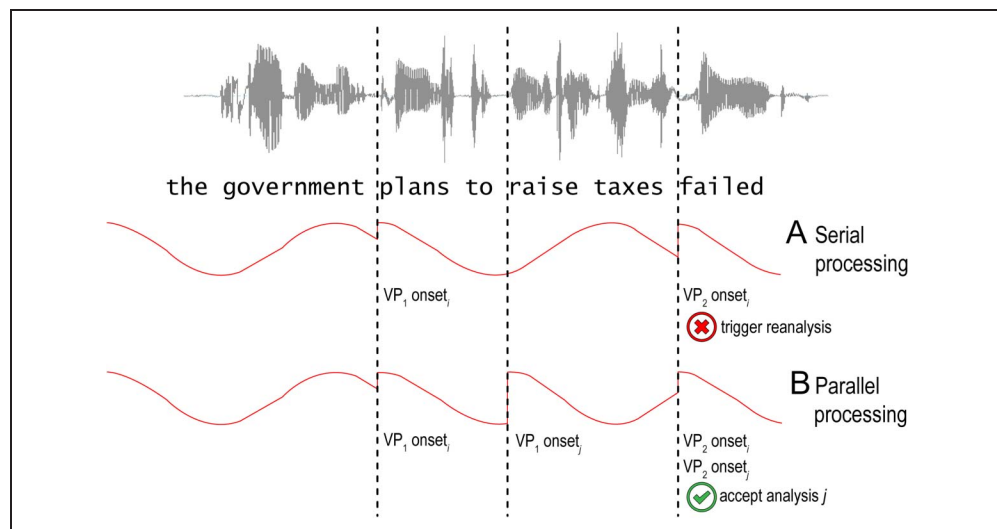
Before the disambiguating word (“failed”), either interpretation of the sentence is viable. Garden-path effects suggest that comprehenders initially assume the structure in 1a (MacDonald, 1993; Frazier & Rayner, 1987). The structure is initially favored at least in part because “plans” occurs more frequently as a verb (59 occurrences) than as a noun (two occurrences) in this particular context (Corpus of Contemporary American English; Davies, 2008).

Sentence processing theories disagree with respect to whether multiple structures are considered simultaneously and on where in the sentence the parser will encounter difficulty. Serial processing models (e.g., Frazier, 1987; Frazier & Fodor, 1978) build only one structure at a time, and reanalysis only occurs when the parser attempts to integrate a syntactic unit that is not compatible with the structure. In the sentence under consideration, encountering the verb “failed” would trigger reanalysis. Parallel processing models (e.g., MacDonald, Pearlmutter, & Seidenberg, 1994; Trueswell & Tanenhaus, 1994), as the name implies, generate multiple structures and narrow down the field of candidates as the parser encounters more and more disambiguating information, which means the parser should encounter the greatest difficulty during the ambiguous region of the sentence (before “failed”).

Under a parallel processing model, during the temporarily ambiguous region of a sentence, at least two competing parses (1a and 1b) are actively under consideration. Crucially, the syntactic phrase boundaries differ between the two structures early on in the sentence. We would expect to see cortical tracking to phrase boundaries corresponding to each of the competing parses during the ambiguous portion of the sentence as the parser considers multiple viable candidates. In contrast, under a serial processing model, only one parse (e.g., 1a) would be considered at a time, and the delta-band oscillatory phase should indicate the parse under consideration. We would therefore predict cortical tracking to syntactic phrase boundaries that are consistent with the parse under consideration only, and we would expect delta-band oscillatory phase reset to occur once contradictory evidence is encountered. Thus, cortical tracking methods provide us with a unique opportunity to resolve some theoretical issues that have proven difficult to disentangle using common behavioral methods such as the recording of eye movements during reading (Figure 1).

In addition, signal processing studies that compare cortical tracking to attended versus unattended speech suggest that we might be able to study depth of

Figure 1. Illustration of predicted delta-band oscillation responses under (A) serial and (B) parallel processing accounts of sentence parsing. Cortical tracking of verb phrase boundaries indicated by a reset in oscillatory phase. The approximation of delta oscillations at 1 Hz is simplified (relative to an actual EEG recording) for clarity.



processing by measuring the degree of cortical tracking of speech. There is evidence that listeners employ shallow processing to efficiently construct a “good-enough” interpretation of the sentence (Ferreira & Patson, 2007; Ferreira, 2003; Christianson et al., 2001). For example, Ferreira (2003) presented listeners with sentences describing transitive events that either were plausible or implausible, and had active or passive syntax, and found a tendency for listeners to transform implausible passive sentences (e.g., “The dog was bitten by the man”) into actives (e.g., “The dog bit the man”), thereby “correcting” the noncanonical nature of both the syntax and meaning of the sentence. The degree of cortical tracking to speech may be able to predict whether or not the listener used a heuristic strategy when processing the sentence. Specifically, we might expect weak cortical tracking to “The dog was bitten by the man” to predict a listener arriving at the incorrect but more felicitous “The dog bit the man” interpretation.

Cortical tracking would supplement not only behavioral methods but also the measures of neural activity already in use in the field of sentence processing. Cortical tracking goes beyond the use of ERPs to study language processing in that it can reveal processes occurring continuously, rather than being constrained by neural responses to discrete events. This could facilitate the process of generating linguistic stimuli, which are often required to be built around specific target words in many current designs; using cortical tracking methods, sentence processing researchers may be able to expand to more naturalistic stimuli. Cortical tracking methods also go beyond the time–frequency analyses currently in use in sentence processing research by observing neural activity that is phase-aligned to periodicities in the stimuli. As we have shown, this property may be exploited to measure how comprehenders deal with stimuli presenting ambiguous structures. Whereas the types of time–

frequency analyses already in use add an invaluable piece to our understanding of language comprehension (Prystauka & Lewis, 2019), cortical tracking tools will undoubtedly add to the types of linguistic questions and paradigms that can be addressed through the recording of EEG and magnetoencephalography data.

In summary, there are exciting opportunities to investigate psycholinguistic theories by studying cortical tracking of speech and to use psycholinguistic methods to further elucidate the relationship between cortical tracking and cognitive processes associated with language processing and comprehension. As we have argued, cortical tracking may help resolve long-standing debates such as whether parsing occurs in a serial or parallel fashion, which have been left unresolved by behavioral methods and the measures of neural activity currently employed in this field.

Conclusion

The fields of sentence and signal processing both seek to understand how listeners process speech, yet collaboration between the two fields has been limited. We outlined several barriers to collaboration, with the primary ones being the different methods used across fields as well as differences in the constraints that experiments in each field must satisfy. Although some of those constraints are at odds with each other, many can be reconciled. We advocate for further collaboration across fields, which would require researchers in each area to acknowledge the experimental constraints of the other and to integrate interdisciplinary methods in their own work, whenever possible. We believe both sentence processing and signal processing research would benefit as a result, because (1) avoiding linguistic stimulus confounds would help determine whether cortical tracking reflects evoked

responses or neural entrainment, (2) psycholinguists could pursue research questions that current methods (e.g., ERPs) are not well suited to address, and (3) language processing models in psycholinguistics could be better informed by incorporating findings from signal processing work. More broadly, both fields would be able to make a fuller use of their data. Signal processing researchers could reduce unwanted noise by controlling and manipulating linguistic features of their stimuli that are often overlooked and ensuring that full comprehension takes place. Sentence processing researchers could better interpret real-time processing by measuring continuous neural activity corresponding to the structure of the stimuli, rather than limiting themselves to observations of neural responses to discrete events such as particular target words. Further collaboration will give rise to new and exciting scientific discoveries of interest to both research communities.

Acknowledgments

We thank the two anonymous reviewers for their insightful comments, which helped us greatly improve our paper. Opinions, interpretations, conclusions, and recommendations are those of the authors and are not necessarily endorsed by the Department of Defense.

Reprint requests should be sent to Eleonora J. Beier, Psychology Department, University of California, Davis, One Shields Ave., Davis, CA 95616-5270, or via e-mail: ejbeier@ucdavis.edu.

Funding Information

We acknowledge support from the National Science Foundation (<http://dx.doi.org/10.13039/100000001>) GRFP number 1650042 awarded to E. J. B., National Science Foundation (<http://dx.doi.org/10.13039/100000001>) grant BCS-1650888 awarded to F. F., and National Institutes of Health grant 1R01HD100516 awarded to F. F. This work was also supported by the Office of the Assistant Secretary of Defense for Health Affairs (<http://dx.doi.org/10.13039/100000005>) through the Hearing Restoration Research Program, under award no. W81XWH-20-1-0485, to L. M. M.; the National Institutes of Health (<http://dx.doi.org/10.13039/100000002>), with grant no. R56 AG053346-02 awarded to G. R.; and the Chulalongkorn University (<http://dx.doi.org/10.13039/501100002873>), with grant no. CU_GIF_62_01_38_01 awarded to S. C.

Diversity in Citation Practices

A retrospective analysis of the citations in every article published in this journal from 2010 to 2020 has revealed a persistent pattern of gender imbalance: Although the proportions of authorship teams (categorized by estimated gender identification of first author/last author) publishing in the *Journal of Cognitive Neuroscience*

(*JoCN*) during this period were M(an)/M = .408, W(oman)/M = .335, M/W = .108, and W/W = .149, the comparable proportions for the articles that these authorship teams cited were M/M = .579, W/M = .243, M/W = .102, and W/W = .076 (Fulvio et al., *JoCN*, 33:1, pp. 3–7). Consequently, *JoCN* encourages all authors to consider gender balance explicitly when selecting which articles to cite and gives them the opportunity to report their article's gender citation balance. The authors of this article report its proportions of citations by gender category to be as follows: M/M = .433, W/M = .133, M/W = .167, and W/W = .267.

Note

1. For ease of explanation, we opted to show simplified syntactic structures and only the relevant syntactic phrase boundaries in this example.

REFERENCES

- Abney, S. P., & Johnson, M. (1991). Memory requirements and local ambiguities of parsing strategies. *Journal of Psycholinguistic Research*, 20, 233–250. DOI: <https://doi.org/10.1007/BF01067217>
- Ahissar, E., Nagarajan, S., Ahissar, M., Protopapas, A., Mahncke, H., & Merzenich, M. M. (2001). Speech comprehension is correlated with temporal response patterns recorded from auditory cortex. *Proceedings of the National Academy of Sciences, U.S.A.*, 98, 13367–13372. DOI: <https://doi.org/10.1073/pnas.201400998>, PMID: 11698688, PMCID: PMC60877
- Alday, P. M. (2019). M/EEG analysis of naturalistic stories: A review from speech to language processing. *Language, Cognition and Neuroscience*, 34, 457–473. DOI: <https://doi.org/10.1080/23273798.2018.1546882>
- Alexandrou, A. M., Saarinen, T., Kujala, J., & Salmelin, R. (2018). Cortical tracking of global and local variations of speech rhythm during connected natural speech perception. *Journal of Cognitive Neuroscience*, 30, 1704–1719. DOI: https://doi.org/10.1162/jocn_a_01295, PMID: 29916785
- Alexandrou, A. M., Saarinen, T., Kujala, J., & Salmelin, R. (2020). Cortical entrainment: What we can learn from studying naturalistic speech perception. *Language, Cognition and Neuroscience*, 35, 681–693. DOI: <https://doi.org/10.1080/23273798.2018.1518534>
- Alexopoulou, T., Michel, M., Murakami, A., & Meurers, D. (2017). Task effects on linguistic complexity and accuracy: A large-scale learner corpus analysis employing natural language processing techniques. *Language Learning*, 67, 180–208. DOI: <https://doi.org/10.1111/lang.12232>
- Baltzell, L. S., Srinivasan, R., & Richards, V. M. (2017). The effect of prior knowledge and intelligibility on the cortical entrainment response to speech. *Journal of Neurophysiology*, 118, 3144–3151. DOI: <https://doi.org/10.1152/jn.00023.2017>, PMID: 28877963, PMCID: PMC5814715
- Bastiaansen, M., Mazaheri, A., & Jensen, O. (2012). Beyond ERPs: Oscillatory neuronal dynamics. In S. J. Luck & E. S. Kappenman (Eds.), *The Oxford handbook of event-related potential components* (pp. 31–49). New York: Oxford University Press.

- Bastiaansen, M., & Hagoort, P. (2006). Oscillatory neuronal dynamics during language comprehension. *Progress in Brain Research, 159*, 179–196. **DOI:** [https://doi.org/10.1016/S0079-6123\(06\)59012-0](https://doi.org/10.1016/S0079-6123(06)59012-0), **PMID:** 17071231
- Batterink, L. J., & Paller, K. A. (2017). Online neural monitoring of statistical learning. *Cortex, 90*, 31–45. **DOI:** <https://doi.org/10.1016/j.cortex.2017.02.004>, **PMID:** 28324696, **PMCID:** PMC5438777
- Beier, E. J., & Ferreira, F. (2018). The temporal prediction of stress in speech and its relation to musical beat perception. *Frontiers in Psychology, 9*, 431. **DOI:** <https://doi.org/10.3389/fpsyg.2018.00431>, **PMID:** 29666600, **PMCID:** PMC5892344
- Bentin, S., Kutas, M., & Hillyard, S. A. (1993). Electrophysiological evidence for task effects on semantic priming in auditory word processing. *Psychophysiology, 30*, 161–169. **DOI:** <https://doi.org/10.1111/j.1469-8986.1993.tb01729.x>, **PMID:** 8434079
- Besle, J., Schevon, C. A., Mehta, A. D., Lakatos, P., Goodman, R. R., McKhann, G. M., et al. (2011). Tuning of the human neocortex to the temporal dynamics of attended events. *Journal of Neuroscience, 31*, 3176–3185. **DOI:** <https://doi.org/10.1523/JNEUROSCI.4518-10.2011>, **PMID:** 21368029, **PMCID:** PMC3081726
- Bever, T. G., & Townsend, D. J. (2001). Some sentences on our consciousness of sentences. In D. J. Townsend & T. G. Bever (Eds.), *Language, brain, and cognitive development: Essays in honor of Jacques Mehler* (pp. 143–155). Cambridge, MA: MIT Press.
- Biau, E., Torralba, M., Fuentemilla, L., de Diego Balaguer, R., & Soto-Faraco, S. (2015). Speaker's hand gestures modulate speech perception through phase resetting of ongoing neural oscillations. *Cortex, 68*, 76–85. **DOI:** <https://doi.org/10.1016/j.cortex.2014.11.018>, **PMID:** 25595613
- Bock, J. K. (1986). Syntactic persistence in language production. *Cognitive Psychology, 18*, 355–387. **DOI:** [https://doi.org/10.1016/0010-0285\(86\)90004-6](https://doi.org/10.1016/0010-0285(86)90004-6)
- Bögels, S., Schriefers, H., Vonk, W., & Chwilla, D. J. (2011). Pitch accents in context: How listeners process accentuation in referential communication. *Neuropsychologia, 49*, 2022–2036. **DOI:** <https://doi.org/10.1016/j.neuropsychologia.2011.03.032>, **PMID:** 21458470
- Bourguignon, M., De Tiège, X., Op de Beeck, M., Ligot, N., Paquier, P., Van Bogaert, P., et al. (2013). The pace of prosodic phrasing couples the listener's cortex to the reader's voice. *Human Brain Mapping, 34*, 314–326. **DOI:** <https://doi.org/10.1002/hbm.21442>, **PMID:** 22392861, **PMCID:** PMC6869855
- Brennan, J. (2016). Naturalistic sentence comprehension in the brain. *Language and Linguistics Compass, 10*, 299–313. **DOI:** <https://doi.org/10.1111/lnc3.12198>
- Brothers, T., Swaab, T. Y., & Traxler, M. J. (2017). Goals and strategies influence lexical prediction during sentence comprehension. *Journal of Memory and Language, 93*, 203–216. **DOI:** <https://doi.org/10.1016/j.jml.2016.10.002>
- Brown, C. M., Hagoort, P., & ter Keurs, M. (1999). Electrophysiological signatures of visual lexical processing: Open- and closed-class words. *Journal of Cognitive Neuroscience, 11*, 261–281. **DOI:** <https://doi.org/10.1162/089892999563382>, **PMID:** 10402255
- 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*, 300–309. **DOI:** <https://doi.org/10.1016/j.tics.2014.02.005>, **PMID:** 24630166, **PMCID:** PMC4037370
- Christiansen, M. H., & Chater, N. (2015). The language faculty that wasn't: A usage-based account of natural language recursion. *Frontiers in Psychology, 6*, 1182. **DOI:** <https://doi.org/10.3389/fpsyg.2015.01182>, **PMID:** 26379567, **PMCID:** PMC4550780
- Christianson, K., Hollingworth, A., Halliwell, J. F., & Ferreira, F. (2001). Thematic roles assigned along the garden path linger. *Cognitive Psychology, 42*, 368–407. **DOI:** <https://doi.org/10.1006/cogp.2001.0752>, **PMID:** 11368528
- Chwilla, D. J., Brown, C. M., & Hagoort, P. (1995). The N400 as a function of the level of processing. *Psychophysiology, 32*, 274–285. **DOI:** <https://doi.org/10.1111/j.1469-8986.1995.tb02956.x>, **PMID:** 7784536
- Clifton, C., Jr., Carlson, K., & Frazier, L. (2002). Informative prosodic boundaries. *Language and Speech, 45*, 87–114. **DOI:** <https://doi.org/10.1177/00238309020450020101>, **PMID:** 12613557
- Crosse, M. J., Di Liberto, G. M., Bednar, A., & Lalor, E. C. (2016). The multivariate temporal response function (mTRF) toolbox: A MATLAB toolbox for relating neural signals to continuous stimuli. *Frontiers in Human Neuroscience, 10*, 604. **DOI:** <https://doi.org/10.3389/fnhum.2016.00604>, **PMID:** 27965557, **PMCID:** PMC5127806
- Crystal, T. H., & House, A. S. (1982). Segmental durations in connected speech signals: Preliminary results. *Journal of the Acoustical Society of America, 72*, 705–716. **DOI:** <https://doi.org/10.1121/1.388251>, **PMID:** 7130529
- Cunillera, T., Toro, J. M., Sebastián-Gallés, N., & Rodríguez-Fornells, A. (2006). The effects of stress and statistical cues on continuous speech segmentation: An event-related brain potential study. *Brain Research, 1123*, 168–178. **DOI:** <https://doi.org/10.1016/j.brainres.2006.09.046>, **PMID:** 17064672
- Cutler, A., Dahan, D., & van Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech, 40*, 141–201. **DOI:** <https://doi.org/10.1177/002383099704000203>, **PMID:** 9509577
- Davies, M. (2008). The corpus of contemporary American English (COCA): 560 million words, 1990–present. Accessed May 14, 2019: <https://corpus.byu.edu/coca/>.
- Deacon, D., Breton, F., Ritter, W., & Vaughan, H. G., Jr. (1991). The relationship between N2 and N400: Scalp distribution, stimulus probability, and task relevance. *Psychophysiology, 28*, 185–200. **DOI:** <https://doi.org/10.1111/j.1469-8986.1991.tb00411.x>, **PMID:** 1946885
- Di Liberto, G. M., & Lalor, E. C. (2017). Indexing cortical entrainment to natural speech at the phonemic level: Methodological considerations for applied research. *Hearing Research, 348*, 70–77. **DOI:** <https://doi.org/10.1016/j.heares.2017.02.015>, **PMID:** 28246030
- Ding, N., Chatterjee, M., & Simon, J. Z. (2014). Robust cortical entrainment to the speech envelope relies on the spectro-temporal fine structure. *Neuroimage, 88*, 41–46. **DOI:** <https://doi.org/10.1016/j.neuroimage.2013.10.054>, **PMID:** 24188816, **PMCID:** PMC4222995
- Ding, N., Melloni, L., Zhang, H., Tian, X., & Poeppel, D. (2016). Cortical tracking of hierarchical linguistic structures in connected speech. *Nature Neuroscience, 19*, 158–164. **DOI:** <https://doi.org/10.1038/nn.4186>, **PMID:** 26642090, **PMCID:** PMC4809195
- Ding, N., & Simon, J. Z. (2014). Cortical entrainment to continuous speech: Functional roles and interpretations. *Frontiers in Human Neuroscience, 8*, 311. **DOI:** <https://doi.org/10.3389/fnhum.2014.00311>, **PMID:** 24904354, **PMCID:** PMC4036061
- Doelling, K. B., Arnal, L. H., Ghitz, O., & Poeppel, D. (2014). Acoustic landmarks drive delta–theta oscillations to enable speech comprehension by facilitating perceptual parsing. *Neuroimage, 85*, 761–768. **DOI:** <https://doi.org/10.1016/j.neuroimage.2013.06.035>, **PMID:** 23791839, **PMCID:** PMC3839250

- Falk, S., Lanzilotti, C., & Schön, D. (2017). Tuning neural phase entrainment to speech. *Journal of Cognitive Neuroscience*, 29, 1378–1389. **DOI:** https://doi.org/10.1162/jocn_a_01136, **PMID:** 28430043
- Fedorenko, E., & Blank, I. A. (2020). Broca's area is not a natural kind. *Trends in Cognitive Sciences*, 24, 270–284. **DOI:** <https://doi.org/10.1016/j.tics.2020.01.001>, **PMID:** 32160565, **PMCID:** PMC7211504
- Ferreira, F. (1993). Creation of prosody during sentence production. *Psychological Review*, 100, 233–253. **DOI:** <https://doi.org/10.1037/0033-295X.100.2.233>, **PMID:** 8483983
- Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology*, 47, 164–203. **DOI:** [https://doi.org/10.1016/s0010-0285\(03\)00005-7](https://doi.org/10.1016/s0010-0285(03)00005-7), **PMID:** 12948517
- Ferreira, F., & Anes, M. (1994). Why study spoken language? In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 33–56). San Diego, CA: Academic Press.
- Ferreira, F., & Lowder, M. W. (2016). Prediction, information structure, and good-enough language processing. In B. H. Ross (Ed.), *The psychology of learning and motivation* (Vol. 65, pp. 217–247). New York: Academic Press.
- Ferreira, F., & Patson, N. D. (2007). The 'good enough' approach to language comprehension. *Language and Linguistics Compass*, 1, 71–83. **DOI:** <https://doi.org/10.1111/j.1749-818X.2007.00007.x>
- Ferreira, F., & Yang, Z. (2019). The problem of comprehension in psycholinguistics. *Discourse Processes*, 56, 485–495. **DOI:** <https://doi.org/10.1080/0163853X.2019.1591885>
- Finke, M., & Rogina, I. (1997). Wide context acoustic modeling in read vs. spontaneous speech. In *1997 IEEE International Conference on Acoustics, Speech, and Signal Processing* (Vol. 3, pp. 1743–1746). Los Alamitos, CA: IEEE Computer Society Press.
- Foertsch, J., & Gernsbacher, M. A. (1994). In search of complete comprehension: Getting "minimalists" to work. *Discourse Processes*, 18, 271–296. **DOI:** <https://doi.org/10.1080/01638539409544896>, **PMID:** 25520530, **PMCID:** PMC4266472
- Frazier, L. (1987). Sentence processing: A tutorial review. In M. Coltheart (Ed.), *Attention and performance 12: The psychology of reading* (pp. 559–586). Hillsdale, NJ: Erlbaum.
- Frazier, L., & Fodor, J. D. (1978). The sausage machine: A new two-stage parsing model. *Cognition*, 6, 291–325. **DOI:** [https://doi.org/10.1016/0010-0277\(78\)90002-1](https://doi.org/10.1016/0010-0277(78)90002-1)
- Frazier, L., & Rayner, K. (1987). Resolution of syntactic category ambiguities: Eye movements in parsing lexically ambiguous sentences. *Journal of Memory and Language*, 26, 505–526. **DOI:** [https://doi.org/10.1016/0749-596X\(87\)90137-9](https://doi.org/10.1016/0749-596X(87)90137-9)
- Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*, 6, 78–84. **DOI:** [https://doi.org/10.1016/s1364-6613\(00\)01839-8](https://doi.org/10.1016/s1364-6613(00)01839-8), **PMID:** 15866191
- Friederici, A. D. (2011). The brain basis of language processing: From structure to function. *Physiological Reviews*, 91, 1357–1392. **DOI:** <https://doi.org/10.1152/physrev.00006.2011>, **PMID:** 22013214
- Friederici, A. D., & Meyer, M. (2004). The brain knows the difference: Two types of grammatical violations. *Brain Research*, 1000, 72–77. **DOI:** <https://doi.org/10.1016/j.brainres.2003.10.057>, **PMID:** 15053954
- Futrell, R., Gibson, E., & Levy, R. P. (2020). Lossy-context surprisal: An information-theoretic model of memory effects in sentence processing. *Cognitive Science*, 44, e12814. **DOI:** <https://doi.org/10.1111/cogs.12814>, **PMID:** 32100918, **PMCID:** PMC7065005
- Getz, H., Ding, N., Newport, E. L., & Poeppel, D. (2018). Cortical tracking of constituent structure in language acquisition. *Cognition*, 181, 135–140. **DOI:** <https://doi.org/10.1016/j.cognition.2018.08.019>, **PMID:** 30195135, **PMCID:** PMC6201233
- 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, 113–126. **DOI:** <https://doi.org/10.1159/000208934>, **PMID:** 19390234
- Gibson, E., Bergen, L., & Piantadosi, S. T. (2013). Rational integration of noisy evidence and prior semantic expectations in sentence interpretation. *Proceedings of the National Academy of Sciences, U.S.A.*, 110, 8051–8056. **DOI:** <https://doi.org/10.1073/pnas.1216438110>, **PMID:** 23637344, **PMCID:** PMC3657782
- Giraud, A.-L., & Poeppel, D. (2012). Cortical oscillations and speech processing: Emerging computational principles and operations. *Nature Neuroscience*, 15, 511–517. **DOI:** <https://doi.org/10.1038/nn.3063>, **PMID:** 22426255, **PMCID:** PMC4461038
- Gomez-Ramirez, M., Kelly, S. P., Molholm, S., Sehatpour, P., Schwartz, T. H., & Foxe, J. J. (2011). Oscillatory sensory selection mechanisms during intersensory attention to rhythmic auditory and visual inputs: A human electrocorticographic investigation. *Journal of Neuroscience*, 31, 18556–18567. **DOI:** <https://doi.org/10.1523/JNEUROSCI.2164-11.2011>, **PMID:** 22171054, **PMCID:** PMC3298747
- Gunter, T. C., & Friederici, A. D. (1999). Concerning the automaticity of syntactic processing. *Psychophysiology*, 36, 126–137. **DOI:** <https://doi.org/10.1017/s004857729997155x>, **PMID:** 10098388
- Gross, J., Hoogenboom, N., Thut, G., Schyns, P., Panzeri, S., Belin, P., et al. (2013). Speech rhythms and multiplexed oscillatory sensory coding in the human brain. *PLoS Biology*, 11, e1001752. **DOI:** <https://doi.org/10.1371/journal.pbio.1001752>, **PMID:** 24391472, **PMCID:** PMC3876971
- Hale, J., Dyer, C., Kuncoro, A., & Brennan, J. (2018). Finding syntax in human encephalography with beam search. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (pp. 2727–2736). Melbourne, Australia: Association for Computational Linguistics. **DOI:** <https://doi.org/10.18653/v1/P18-1254>
- Hamilton, L. S., & Huth, A. G. (2020). The revolution will not be controlled: Natural stimuli in speech neuroscience. *Language, Cognition and Neuroscience*, 35, 573–582. **DOI:** <https://doi.org/10.1080/23273798.2018.1499946>, **PMID:** 32656294, **PMCID:** PMC7324135
- Hauk, O., & Pulvermüller, F. (2004). Effects of word length and frequency on the human event-related potential. *Clinical Neurophysiology*, 115, 1090–1103. **DOI:** <https://doi.org/10.1016/j.clinph.2003.12.020>, **PMID:** 15066535
- Henderson, J. M., Choi, W., Lowder, M. W., & Ferreira, F. (2016). Language structure in the brain: A fixation-related fMRI study of syntactic surprisal in reading. *Neuroimage*, 132, 293–300. **DOI:** <https://doi.org/10.1016/j.neuroimage.2016.02.050>, **PMID:** 26908322
- Hirose, K., & Kawanami, H. (2002). Temporal rate change of dialogue speech in prosodic units as compared to read speech. *Speech Communication*, 36, 97–111. **DOI:** [https://doi.org/10.1016/S0167-6393\(01\)00028-0](https://doi.org/10.1016/S0167-6393(01)00028-0)
- Holcomb, P. J., Grainger, J., & O'Rourke, T. (2002). An electrophysiological study of the effects of orthographic neighborhood size on printed word perception. *Journal of Cognitive Neuroscience*, 14, 938–950. **DOI:** <https://doi.org/10.1162/089892902760191153>, **PMID:** 12191460
- Howard, M. F., & Poeppel, D. (2010). Discrimination of speech stimuli based on neuronal response phase patterns depends on acoustics but not comprehension. *Journal of*

- Neurophysiology*, 104, 2500–2511. **DOI:** <https://doi.org/10.1152/jn.00251.2010>, **PMID:** 20484530, **PMCID:** PMC2997028
- Huang, Y., & Ferreira, F. (2020). The application of signal detection theory to acceptability judgments. *Frontiers in Psychology*, 11, 73. **DOI:** <https://doi.org/10.3389/fpsyg.2020.00073>, **PMID:** 32082223, **PMCID:** PMC7005104
- Huetig, F., & Janse, E. (2016). Individual differences in working memory and processing speed predict anticipatory spoken language processing in the visual world. *Language, Cognition and Neuroscience*, 31, 80–93. **DOI:** <https://doi.org/10.1080/23273798.2015.1047459>
- Jewett, D. L., & Williston, J. S. (1971). Auditory-evoked far fields averaged from the scalp of humans. *Brain*, 94, 681–696. **DOI:** <https://doi.org/10.1093/brain/94.4.681>, **PMID:** 5132966
- Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Cambridge, MA: Harvard University Press.
- Keitel, A., Gross, J., & Kayser, C. (2018). Perceptually relevant speech tracking in auditory and motor cortex reflects distinct linguistic features. *PLoS Biology*, 16, e2004473. **DOI:** <https://doi.org/10.1371/journal.pbio.2004473>, **PMID:** 29529019, **PMCID:** PMC5864086
- Keitel, A., Ince, R. A. A., Gross, J., & Kayser, C. (2017). Auditory cortical delta-entrainment interacts with oscillatory power in multiple fronto-parietal networks. *Neuroimage*, 147, 32–42. **DOI:** <https://doi.org/10.1016/j.neuroimage.2016.11.062>, **PMID:** 27903440, **PMCID:** PMC5315055
- Kerlin, J. R., Shahin, A. J., & Miller, L. M. (2010). Attentional gain control of ongoing cortical speech representations in a “cocktail party.” *Journal of Neuroscience*, 30, 620–628. **DOI:** <https://doi.org/10.1523/JNEUROSCI.3631-09.2010>, **PMID:** 20071526, **PMCID:** PMC2832933
- Kim, A. E., Oines, L., & Miyake, A. (2018). Individual differences in verbal working memory underlie a tradeoff between semantic and structural processing difficulty during language comprehension: An ERP investigation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 44, 406–420. **DOI:** <https://doi.org/10.1037/xlm0000457>, **PMID:** 28933902
- King, J. W., & Kutas, M. (1998). Neural plasticity in the dynamics of human visual word recognition. *Neuroscience Letters*, 244, 61–64. **DOI:** [https://doi.org/10.1016/s0304-3940\(98\)00140-2](https://doi.org/10.1016/s0304-3940(98)00140-2), **PMID:** 9572585
- Kösem, A., Bosker, H. R., Takashima, A., Meyer, A., Jensen, O., & Hagoort, P. (2018). Neural entrainment determines the words we hear. *Current Biology*, 28, 2867–2875. **DOI:** <https://doi.org/10.1016/j.cub.2018.07.023>, **PMID:** 30197083
- Kösem, A., & van Wassenhove, V. (2017). Distinct contributions of low- and high-frequency neural oscillations to speech comprehension. *Language, Cognition and Neuroscience*, 32, 536–544. **DOI:** <https://doi.org/10.1080/23273798.2016.1238495>
- Kotz, S. A., & Schwartze, M. (2010). Cortical speech processing unplugged: A timely subcortico-cortical framework. *Trends in Cognitive Sciences*, 14, 392–399. **DOI:** <https://doi.org/10.1016/j.tics.2010.06.005>, **PMID:** 20655802
- Kounios, J., & Holcomb, P. J. (1994). Concreteness effects in semantic processing: ERP evidence supporting dual-coding theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 804–823. **DOI:** <https://doi.org/10.1037/0278-7393.20.4.804>, **PMID:** 8064248
- Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience*, 31, 32–59. **DOI:** <https://doi.org/10.1080/23273798.2015.1102299>, **PMID:** 27135040, **PMCID:** PMC4850025
- Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*, 62, 621–647. **DOI:** <https://doi.org/10.1146/annurev.psych.093008.131123>, **PMID:** 20809790, **PMCID:** PMC4052444
- Kutas, M., Van Petten, C. K., & Kluender, R. (2006). Psycholinguistics electrified II (1994–2005). In M. J. Traxler & M. A. Gernsbacher (Eds.), *Handbook of psycholinguistics* (pp. 659–724). New York: Academic Press. **DOI:** <https://doi.org/10.1016/B978-012369374-7/50018-3>
- 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, 110–113. **DOI:** <https://doi.org/10.1126/science.1154735>, **PMID:** 18388295
- 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, 1904–1911. **DOI:** <https://doi.org/10.1152/jn.00263.2005>, **PMID:** 15901760
- Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review*, 106, 119–159. **DOI:** <https://doi.org/10.1037/0033-295X.106.1.119>
- Large, E. W., & Kolen, J. F. (1994). Resonance and the perception of musical meter. *Connection Science*, 6, 177–208. **DOI:** <https://doi.org/10.1080/09540099408915723>
- Ljung, L., Chen, T., & Mu, B. (2020). A shift in paradigm for system identification. *International Journal of Control*, 93, 173–180. **DOI:** <https://doi.org/10.1080/00207179.2019.1578407>
- 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, e1000445. **DOI:** <https://doi.org/10.1371/journal.pbio.1000445>, **PMID:** 20711473, **PMCID:** PMC2919416
- Luo, H., & Poeppel, D. (2007). Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. *Neuron*, 54, 1001–1010. **DOI:** <https://doi.org/10.1016/j.neuron.2007.06.004>, **PMID:** 17582338, **PMCID:** PMC2703451
- MacDonald, M. C. (1993). The interaction of lexical and syntactic ambiguity. *Journal of Memory and Language*, 32, 692–715. **DOI:** <https://doi.org/10.1006/jmla.1993.1035>
- MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, 101, 676–703. **DOI:** <https://doi.org/10.1037/0033-295x.101.4.676>, **PMID:** 7984711
- Martin, A. E., & Doumas, L. A. A. (2017). A mechanism for the cortical computation of hierarchical linguistic structure. *PLoS Biology*, 15, e2000663. **DOI:** <https://doi.org/10.1371/journal.pbio.2000663>, **PMID:** 28253256, **PMCID:** PMC5333798
- Matchin, W., & Hickok, G. (2020). The cortical organization of syntax. *Cerebral Cortex*, 30, 1481–1498. **DOI:** <https://doi.org/10.1093/cercor/bhz180>, **PMID:** 31670779, **PMCID:** PMC7132936
- Meyer, L. (2018). The neural oscillations of speech processing and language comprehension: State of the art and emerging mechanisms. *European Journal of Neuroscience*, 48, 2609–2621. **DOI:** <https://doi.org/10.1111/ejn.13748>, **PMID:** 29055058
- Meyer, L., & Gumbert, M. (2018). Synchronization of electrophysiological responses with speech benefits syntactic information processing. *Journal of Cognitive Neuroscience*, 30, 1066–1074. **DOI:** https://doi.org/10.1162/jocn_a_01236, **PMID:** 29324074

- Meyer, L., Henry, M. J., Gaston, P., Schmuck, N., & Friederici, A. D. (2017). Linguistic bias modulates interpretation of speech via neural delta-band oscillations. *Cerebral Cortex*, *27*, 4293–4302. **DOI:** <https://doi.org/10.1093/cercor/bhw228>, **PMID:** 27566979
- Meyer, L., Sun, Y., & Martin, A. E. (2020). Synchronous, but not entrained: Exogenous and endogenous cortical rhythms of speech and language processing. *Language, Cognition and Neuroscience*, *35*, 1089–1099. **DOI:** <https://doi.org/10.1080/23273798.2019.1693050>
- Millman, R. E., Johnson, S. R., & Prendergast, G. (2015). The role of phase-locking to the temporal envelope of speech in auditory perception and speech intelligibility. *Journal of Cognitive Neuroscience*, *27*, 533–545. **DOI:** https://doi.org/10.1162/jocn_a_00719, **PMID:** 25244119
- Morillon, B., & Schroeder, C. E. (2015). Neuronal oscillations as a mechanistic substrate of auditory temporal prediction. *Annals of the New York Academy of Sciences*, *1337*, 26–31. **DOI:** <https://doi.org/10.1111/nyas.12629>, **PMID:** 25773613, **PMCID:** PMC4363099
- Moseley, R. L., & Pulvermüller, F. (2014). Nouns, verbs, objects, actions, and abstractions: Local fMRI activity indexes semantics, not lexical categories. *Brain and Language*, *132*, 28–42. **DOI:** <https://doi.org/10.1016/j.bandl.2014.03.001>, **PMID:** 24727103, **PMCID:** PMC4029073
- Myers, J. (2009). Syntactic judgment experiments. *Language and Linguistics Compass*, *3*, 406–423. **DOI:** <https://doi.org/10.1111/j.1749-818X.2008.00113.x>
- Nakajima, S., & Allen, J. F. (1993). A study on prosody and discourse structure in cooperative dialogues. *Phonetica*, *50*, 197–210. **DOI:** <https://doi.org/10.1159/000261940>
- Nobre, A. C., & van Ede, F. (2018). Anticipated moments: Temporal structure in attention. *Nature Reviews Neuroscience*, *19*, 34–48. **DOI:** <https://doi.org/10.1038/nrn.2017.141>, **PMID:** 29213134
- Nolan, F., & Jeon, H.-S. (2014). Speech rhythm: A metaphor? *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences*, *369*, 20130396. **DOI:** <https://doi.org/10.1098/rstb.2013.0396>, **PMID:** 25385774, **PMCID:** PMC4240963
- Nora, A., Faisal, A., Seol, J., Renvall, H., Formisano, E., & Salmelin, R. (2020). Dynamic time-locking mechanism in the cortical representation of spoken words. *eNeuro*, *7*, ENEURO.0475-19.2020. **DOI:** <https://doi.org/10.1523/ENEURO.0475-19.2020>, **PMID:** 32513662, **PMCID:** PMC7470935
- Nozaradan, S., Peretz, I., Missal, M., & Mouraux, A. (2011). Tagging the neuronal entrainment to beat and meter. *Journal of Neuroscience*, *31*, 10234–10240. **DOI:** <https://doi.org/10.1523/JNEUROSCI.0411-11.2011>, **PMID:** 21753000, **PMCID:** PMC6623069
- Obleser, J., & Kotz, S. A. (2011). Multiple brain signatures of integration in the comprehension of degraded speech. *Neuroimage*, *55*, 713–723. **DOI:** <https://doi.org/10.1016/j.neuroimage.2010.12.020>, **PMID:** 21172443
- Obleser, J., & Kayser, C. (2019). Neural entrainment and attentional selection in the listening brain. *Trends in Cognitive Sciences*, *23*, 913–926. **DOI:** <https://doi.org/10.1016/j.tics.2019.08.004>, **PMID:** 31606386
- Payton, K. L., Uchanski, R. M., & Braida, L. D. (1994). Intelligibility of conversational and clear speech in noise and reverberation for listeners with normal and impaired hearing. *Journal of the Acoustical Society of America*, *95*, 1581–1592. **DOI:** <https://doi.org/10.1121/1.408545>, **PMID:** 8176061
- Peelle, J. E., & Davis, M. H. (2012). Neural oscillations carry speech rhythm through to comprehension. *Frontiers in Psychology*, *3*, 320. **DOI:** <https://doi.org/10.3389/fpsyg.2012.00320>, **PMID:** 22973251, **PMCID:** PMC3434440
- 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*, 1378–1387. **DOI:** <https://doi.org/10.1093/cercor/bhs118>, **PMID:** 22610394, **PMCID:** PMC3643716
- Peelle, J. E., Miller, R. L., Rogers, C. S., Spehar, B., Sommers, M. S., & Van Engen, K. J. (2020). Completion norms for 3085 English sentence contexts. *Behavior Research Methods*, *52*, 1795–1799. **DOI:** <https://doi.org/10.3758/s13428-020-01351-1>, **PMID:** 31993960, **PMCID:** PMC7406521
- Picheny, M. A., Durlach, N. I., & Braida, L. D. (1985). Speaking clearly for the hard of hearing I: Intelligibility differences between clear and conversational speech. *Journal of Speech, Language, and Hearing Research*, *28*, 96–103. **DOI:** <https://doi.org/10.1044/jshr.2801.96>, **PMID:** 3982003
- Pickering, M. J., & Ferreira, V. S. (2008). Structural priming: A critical review. *Psychological Bulletin*, *134*, 427–459. **DOI:** <https://doi.org/10.1037/0033-2909.134.3.427>, **PMID:** 18444704, **PMCID:** PMC2657366
- Price, C. J. (2012). A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. *Neuroimage*, *62*, 816–847. **DOI:** <https://doi.org/10.1016/j.neuroimage.2012.04.062>, **PMID:** 22584224, **PMCID:** PMC3398395
- Prystauka, Y., & Lewis, A. G. (2019). The power of neural oscillations to inform sentence comprehension: A linguistic perspective. *Language and Linguistics Compass*, *13*, e12347. **DOI:** <https://doi.org/10.1111/lnc3.12347>, **PMID:** 33042211, **PMCID:** PMC7546279
- Rimmele, J. M., Golumbic, E. Z., Schröger, E., & Poeppel, D. (2015). The effects of selective attention and speech acoustics on neural speech-tracking in a multi-talker scene. *Cortex*, *68*, 144–154. **DOI:** <https://doi.org/10.1016/j.cortex.2014.12.014>, **PMID:** 25650107, **PMCID:** PMC4475476
- Rimmele, J. M., Morillon, B., Poeppel, D., & Arnal, L. H. (2018). Proactive sensing of periodic and aperiodic auditory patterns. *Trends in Cognitive Sciences*, *22*, 870–882. **DOI:** <https://doi.org/10.1016/j.tics.2018.08.003>, **PMID:** 30266147
- Rodd, J. M., Vitello, S., Woollams, A. M., & Adank, P. (2015). Localising semantic and syntactic processing in spoken and written language comprehension: An activation likelihood estimation meta-analysis. *Brain and Language*, *141*, 89–102. **DOI:** <https://doi.org/10.1016/j.bandl.2014.11.012>, **PMID:** 25576690
- Russo, A. G., De Martino, M., Mancuso, A., Iaconetta, G., Manara, R., Elia, A., et al. (2020). Semantics-weighted lexical surprisal modeling of naturalistic functional MRI time-series during spoken narrative listening. *Neuroimage*, *222*, 117281. **DOI:** <https://doi.org/10.1016/j.neuroimage.2020.117281>, **PMID:** 32828929
- Salverda, A. P., Brown, M., & Tanenhaus, M. K. (2011). A goal-based perspective on eye movements in visual world studies. *Acta Psychologica*, *137*, 172–180. **DOI:** <https://doi.org/10.1016/j.actpsy.2010.09.010>, **PMID:** 21067708, **PMCID:** PMC3109199
- Sanders, L. D., & Neville, H. J. (2003). An ERP study of continuous speech processing: I. Segmentation, semantics, and syntax in native speakers. *Cognitive Brain Research*, *15*, 228–240. **DOI:** [https://doi.org/10.1016/S0926-6410\(02\)00195-7](https://doi.org/10.1016/S0926-6410(02)00195-7), **PMID:** 12527097
- Sanford, A. J., & Sturt, P. (2002). Depth of processing in language comprehension: Not noticing the evidence. *Trends in Cognitive Sciences*, *6*, 382–386. **DOI:** [https://doi.org/10.1016/S1364-6613\(02\)01958-7](https://doi.org/10.1016/S1364-6613(02)01958-7), **PMID:** 12200180
- Sassenhagen, J. (2019). How to analyse electrophysiological responses to naturalistic language with time-resolved multiple regression. *Language, Cognition and*

- Neuroscience*, 34, 474–490. DOI: <https://doi.org/10.1080/23273798.2018.1502458>
- Schacht, A., Sommer, W., Shmuilovich, O., Martíenz, P. C., & Martín-Loeches, M. (2014). Differential task effects on N400 and P600 elicited by semantic and syntactic violations. *PLoS One*, 9, e91226. DOI: <https://doi.org/10.1371/journal.pone.0091226>, PMID: 24614675, PMCID: PMC3948820
- Schmidt-Kassow, M., & Kotz, S. A. (2009a). Event-related brain potentials suggest a late interaction of meter and syntax in the P600. *Journal of Cognitive Neuroscience*, 21, 1693–1708. DOI: <https://doi.org/10.1162/jocn.2008.21153>, PMID: 18855546
- Schmidt-Kassow, M., & Kotz, S. A. (2009b). Attention and perceptual regularity in speech. *NeuroReport*, 20, 1643–1647. DOI: <https://doi.org/10.1097/WNR.0b013e328333b0c6>, PMID: 19907350
- Schwartz, M., & Kotz, S. A. (2013). A dual-pathway neural architecture for specific temporal prediction. *Neuroscience & Biobehavioral Reviews*, 37, 2587–2596. DOI: <https://doi.org/10.1016/j.neubiorev.2013.08.005>, PMID: 23994272
- Shattuck-Hufnagel, S., & Turk, A. E. (1996). A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research*, 25, 193–247. DOI: <https://doi.org/10.1007/BF01708572>, PMID: 8667297
- Song, J., & Iverson, P. (2018). Listening effort during speech perception enhances auditory and lexical processing for non-native listeners and accents. *Cognition*, 179, 163–170. DOI: <https://doi.org/10.1016/j.cognition.2018.06.001>, PMID: 29957515
- Staub, A., Rayner, K., Pollatsek, A., Hyönä, J., & Majewski, H. (2007). The time course of plausibility effects on eye movements in reading: Evidence from noun–noun compounds. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33, 1162–1169. DOI: <https://doi.org/10.1037/0278-7393.33.6.1162>, PMID: 17983320
- 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, 13578–13585. DOI: <https://doi.org/10.1523/JNEUROSCI.0703-10.2010>, PMID: 20943899, PMCID: PMC4427664
- Strijkers, K., Costa, A., & Thierry, G. (2010). Tracking lexical access in speech production: Electrophysiological correlates of word frequency and cognate effects. *Cerebral Cortex*, 20, 912–928. DOI: <https://doi.org/10.1093/cercor/bhp153>, PMID: 19679542
- Swaab, T. Y., Ledoux, K., Camblin, C. C., & Boudewyn, M. A. (2012). Language-related ERP components. In S. J. Luck & E. S. Kappenman (Eds.), *Oxford handbook of event-related potential components* (pp. 397–440). Oxford: Oxford University Press. DOI: <https://doi.org/10.1093/oxfordhb/9780195374148.013.0197>
- 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: General*, 136, 64–81. DOI: <https://doi.org/10.1037/0096-3445.136.1.64>, PMID: 17324085
- Taylor, W. L. (1953). “Cloze procedure”: A new tool for measuring readability. *Journalism & Mass Communication Quarterly*, 30, 415–433. DOI: <https://doi.org/10.1177/107769905303000401>
- Teinonen, T., & Huottilainen, M. (2012). Implicit segmentation of a stream of syllables based on transitional probabilities: An MEG study. *Journal of Psycholinguistic Research*, 41, 71–82. DOI: <https://doi.org/10.1007/s10936-011-9182-2>, PMID: 21993901
- Tooley, K. M., Traxler, M. J., & Swaab, T. Y. (2009). Electrophysiological and behavioral evidence of syntactic priming in sentence comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35, 19–45. DOI: <https://doi.org/10.1037/a0013984>, PMID: 19210079
- Traxler, M. J. (2014). Trends in syntactic parsing: Anticipation, Bayesian estimation, and good-enough parsing. *Trends in Cognitive Sciences*, 18, 605–611. DOI: <https://doi.org/10.1016/j.tics.2014.08.001>, PMID: 25200381, PMCID: PMC6814003
- Trueswell, J. C., & Tanenhaus, M. K. (1994). Toward a lexicalist framework of constraint-based syntactic ambiguity resolution. In C. Clifton Jr., L. Frazier, & K. Rayner (Eds.), *Perspectives on sentence processing* (pp. 155–179). Hillsdale, NJ: Erlbaum.
- Tyler, L. K. (Ed.). (2020). Meyer forum [special issue]. *Language, Cognition and Neuroscience*, 35, 1089–1222. DOI: <https://doi.org/10.1080/23273798.2019.1693050>
- Uchanski, R. M., Choi, S. S., Braid, L. D., Reed, C. M., & Durlach, N. I. (1996). Speaking clearly for the hard of hearing IV: Further studies of the role of speaking rate. *Journal of Speech, Language, and Hearing Research*, 39, 494–509. DOI: <https://doi.org/10.1044/jshr.3903.494>, PMID: 8783129
- van Berkum, J. J. A. (2004). Sentence comprehension in a wider discourse: Can we use ERPs to keep track of things? In M. Carreiras & C. Clifton Jr. (Eds.), *The on-line study of sentence comprehension: Eyetracking, ERPs and beyond* (pp. 229–270). New York: Psychology Press.
- van Berkum, J. J. A., Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31, 443–467. DOI: <https://doi.org/10.1037/0278-7393.31.3.443>, PMID: 15910130
- Van Petten, C., & Kutas, M. (1990). Interactions between sentence context and word frequency in event-related brain potentials. *Memory & Cognition*, 18, 380–393. DOI: <https://doi.org/10.3758/bf03197127>, PMID: 2381317
- Wang, L., Bastiaansen, M., Yang, Y., & Hagoort, P. (2011). The influence of information structure on the depth of semantic processing: How focus and pitch accent determine the size of the N400 effect. *Neuropsychologia*, 49, 813–820. DOI: <https://doi.org/10.1016/j.neuropsychologia.2010.12.035>, PMID: 21195102
- Wang, L., Bastiaansen, M., Yang, Y., & Hagoort, P. (2012). Information structure influences depth of syntactic processing: Event-related potential evidence for the Chomsky illusion. *PLoS One*, 7, e47917. DOI: <https://doi.org/10.1371/journal.pone.0047917>, PMID: 23110131, PMCID: PMC3480462
- Wang, L., Hagoort, P., & Jensen, O. (2018). Language prediction is reflected by coupling between frontal gamma and posterior alpha oscillations. *Journal of Cognitive Neuroscience*, 30, 432–447. DOI: https://doi.org/10.1162/jocn_a_01190, PMID: 28949823
- Weissbart, H., Kandylaki, K. D., & Reichenbach, T. (2020). Cortical tracking of surprisal during continuous speech comprehension. *Journal of Cognitive Neuroscience*, 32, 155–166. DOI: https://doi.org/10.1162/jocn_a_01467, PMID: 31479349
- Willems, R. M., Frank, S. L., Nijhof, A. D., Hagoort, P., & van den Bosch, A. (2016). Prediction during natural language comprehension. *Cerebral Cortex*, 26, 2506–2516. DOI: <https://doi.org/10.1093/cercor/bhv075>, PMID: 25903464
- Winsler, K., Midgley, K. J., Grainger, J., & Holcomb, P. J. (2018). An electrophysiological megastudy of spoken word recognition. *Language, Cognition and Neuroscience*, 33, 1063–1082. DOI: <https://doi.org/10.1080/23273798.2018.1455985>
- Zaccarella, E., Schell, M., & Friederici, A. D. (2017). Reviewing the functional basis of the syntactic Merge mechanism for

- language: A coordinate-based activation likelihood estimation meta-analysis. *Neuroscience & Biobehavioral Reviews*, *80*, 646–656. **DOI:** <https://doi.org/10.1016/j.neubiorev.2017.06.011>, **PMID:** 28743620
- Zoefel, B., Archer-Boyd, A., & Davis, M. H. (2018). Phase entrainment of brain oscillations causally modulates neural responses to intelligible speech. *Current Biology*, *28*, 401–408. **DOI:** <https://doi.org/10.1016/j.cub.2017.11.071>, **PMID:** 29358073, **PMCID:** PMC5807089
- Zoefel, B., & VanRullen, R. (2015). The role of high-level processes for oscillatory phase entrainment to speech sound. *Frontiers in Human Neuroscience*, *9*, 651. **DOI:** <https://doi.org/10.3389/fnhum.2015.00651>, **PMID:** 26696863, **PMCID:** PMC4667100
- Zoefel, B., & VanRullen, R. (2016). EEG oscillations entrain their phase to high-level features of speech sound. *Neuroimage*, *124*, 16–23. **DOI:** <https://doi.org/10.1016/j.neuroimage.2015.08.054>, **PMID:** 26341026