Neural Oscillations: Probing the Temporal Dynamics of Syntactic Parsing

Authors: Qi Ruiying, Feng Ye, Si Fuzhen, sealing blade

Date: 2024-10-27T17:23:16+00:00

Abstract

Searching for correspondences between linguistic phenotypes and neural mechanisms—the so-called mapping problem (the mappingproblem)—constitutes a major focus of current research. Among these, the neural mechanisms underlying syntactic parsing are particularly challenging, as they involve identifying, within neural activity, processes that correspond to syntactic structure building, which is key to unraveling the mystery of human language capacity. Recent research on neural oscillatory activity has not only provided strong evidence for the psychological reality of syntactic processing during syntactic parsing, but has also demonstrated the feasibility of utilizing neural oscillations to elucidate the neural coding activity that underlies syntactic parsing processes. Furthermore, theoretical models of syntactic computation from the Minimalist Program in theoretical linguistics can be cross-validated with experimental studies on neural oscillations in neuroscience, and such research can offer insights into the temporal dynamics of syntactic construction. Future research may concentrate on four areas: finer-grained alignment between neural oscillations and syntactic processing; the generation mechanisms of neural oscillations and their biological significance; developmental patterns of neural oscillations during child language acquisition; and the neurophysiological basis of language disorders and their rehabilitation applications.

Full Text

Preamble

Neural Oscillations: Glimpsing into the Temporal Dynamics of Syntactic Parsing

Qi Ruiying, Feng Ye, Si Fuzhen
(Department of Linguistics / Lab of Biolinguistics and Brain Sciences, Beijing

Language and Culture University, Beijing 100083, China)

Abstract

Establishing correspondences between linguistic phenotypes and neural mechanisms—the so-called mapping problem—represents a major focus of current research. Among its challenges, the neural mechanisms underlying syntactic parsing are particularly formidable, as they involve identifying neural activities that correspond to the construction of syntactic structure, which is key to unraveling the enigma of human linguistic capacity. Recent investigations into neural oscillatory activity not only provide compelling evidence for the psychological reality of syntactic processing during parsing but also demonstrate the feasibility of using neural oscillations to explicate the neural coding activities that underlie syntactic parsing. The theoretical models of syntactic computation within the Minimalist Program of theoretical linguistics can be corroborated by experimental research on neural oscillations in neuroscience, and through such interdisciplinary studies we can glimpse the temporal dynamics of syntactic construction. Future research should concentrate on four areas: finer-grained alignment between neural oscillations and syntactic processing; the generation mechanisms of neural oscillations and their biological significance; developmental patterns of neural oscillations during child language acquisition; and the neurophysiological basis of language disorders and their rehabilitation applications.

Keywords: neural oscillation, language comprehension, syntactic parsing, Minimalist Program, incremental conversion

1. Introduction

David Poeppel (2012) proposed the mapping problem between linguistics and neuroscience: What is the relationship between the ontological structures of these two fields? At its core, this problem concerns how to directly link different aspects of language processing to neural activities in the brain within an interpretive framework, thereby revealing the specific neural mechanisms of natural language processing. The greatest challenge in addressing this question stems from the granularity mismatch problem: theoretical linguistics tends to adopt a fine-grained perspective in describing and investigating language ontology, whereas neuroscience research often focuses on relatively macroscopic, coarse-grained linguistic behaviors. This difference in perspective makes it difficult to establish correspondences between linguistic representations and neural activities, causing language-brain correlation research to remain superficial and preventing identification of the true neural mechanisms underlying linguistic phenomena. Solving the mapping problem is crucial for exploring human linguistic capacity, and consequently, the academic community has begun investigating the neural coding activities of linguistic constructs such as syntax and semantics (Hale et al., 2022).

Currently, an increasing number of studies have revealed the feasibility of using low-frequency neural oscillations as an index of syntactic hierarchical structure construction, making them a strong candidate for elucidating the neural mechanisms relevant to syntactic parsing. In particular, these studies have discovered a non-isomorphic mapping relationship between the timing of exogenous stimuli and the timing represented by endogenous oscillations, as well as a dissociation between grammatical knowledge and semantic statistical cues or lexical properties. These findings reveal the existence of a priori grammatical knowledge dedicated to syntactic construction in the human brain, providing evidence for the psychological reality of syntactic processing. Furthermore, the potential relationship between phase coherence of neural oscillations and different linguistic hierarchical structures has gradually attracted attention, with promising theoretical models emerging in related fields.

2. Glimpsing Syntactic Parsing: Incremental Conversion from Sequence to Hierarchy

The hierarchical nature of syntactic structure is a crucial feature distinguishing natural language from other communication systems, endowing humans with the capacity to generate infinite linguistic expressions using finite means (Chomsky, 1957, 1965). Syntactic parsing—the incremental process of converting linear speech sequences into abstract syntactic hierarchical structures—has long been a central topic in linguistics and cognitive science (Traxler, 2014). Within the framework of the mapping problem, both linguistic theory and neuroscience research have explored and described this incremental conversion process.

2.1 Syntactic Parsing in Linguistic Theory: The Merge Operation

The Minimalist Program, a theoretical framework recently proposed by Noam Chomsky, seeks the most simplified forms of linguistic structure and explanation (Chomsky, 1995). According to this framework, syntactic parsing is essentially a process that takes lexical items as basic atomic units and constructs binary-branching hierarchical sets from accessible syntactic objects in the workspace through the Merge operation. The mathematical expression for Merge is as follows (where WS = workspace; P/Q = syntactic objects; X = additional elements; { } marks phrase structure; [ ] marks features) (Chomsky et al., 2023; Marcolli, Chomsky et al., 2023):

$$WS = \{P, Q, ...\}$$

$$MERGE(P, Q, WS) = WS' = \{\{P, Q\}, X_1, \dots, X_n\}$$

External Merge (directly calling lexical items from the lexicon to generate basic argument structures, etc.) can be expressed as:

$$WS = [X, Y, Z] \rightarrow MERGE(X, Y) \rightarrow WS' = [\{X, Y\}, Z]$$

Internal Merge (i.e., "movement," selecting constituents from within an already constructed structural component and moving them to a new Merge position to check features required by information-structural arrangements) can be expressed as:

$$WS = [\{X, Y\}, Z] \rightarrow MERGE(X, \{X, Y\}) \rightarrow WS' = [\{X, \{X, Y\}\}, Z]$$

In essence, to achieve incremental conversion from linear sequence to hierarchical structure, three aspects must be considered: (1) how representations of syntactic objects available for Merge operations are formed, i.e., how the brain collects sensory information to form lexical representations; (2) how the workspace dynamically updates to store and predict syntactic objects; and (3) how the Merge operation constructs hierarchical structures from lexical representations within the workspace (Chomsky, 1995). The final result of Merge operations is technically represented by syntax trees. According to the Linear Correspondence Axiom, the hierarchical structure of syntax trees maps onto linear sequences, forming phonological linear sequences marked by terminal nodes, which visually displays the layout and procedures of hierarchical syntactic structure construction. Through Merge operations, lower-level syntactic units are integrated into higher-level syntactic representations, achieving incremental conversion from linear sequence to hierarchical structure. From a psycholinguistic perspective, this can be described as: during online input of linear speech sequences, the human brain integrates various types of information to form lexical representations (syntactic objects) after lexical access; these representations are stored in working memory (the workspace) and are subsequently extracted and utilized during inference (Merge) processes to form tree structures with vertical hierarchical relationships, ultimately achieving sentence comprehension.

2.2 Syntactic Parsing in Neuroscience Research: Neural Oscillations and Temporal Structure

Investigating the neural mechanisms of syntactic parsing requires observation of human brain activity. As previously mentioned, within the framework of linguistic theory, syntactic parsing is a process of constructing hierarchical structures from linear sequences. The question of how neural activities in the human brain represent this process constitutes the neural mechanism problem of syntactic parsing. Traditionally, discussions of syntactic parsing mechanisms have focused primarily on Event-Related Potentials (ERPs), such as the N400 associated with semantic violations (Kutas & Federmeier, 2000, 2011), the P600 associated with syntactic violations (Gouvea et al., 2010, p. 600), and the Closure Positive Shift (CPS) related to prosodic boundary identification (Steinhauer et al., 1999). ERP research has achieved significant success, revealing psychological processes that cannot be manifested in behavioral performance through more sensitive electrophysiological indices. However, investigating syntactic processing through evoked potentials has certain limitations: ERP induction relies heavily on violation-based stimulus materials, and its experimental designs inevitably suffer from subjectivity in operational definitions and variable control, leading to major controversies in interpreting its induction mechanisms. Moreover, such research struggles to provide direct temporal information about hierarchical structure processing and cannot reveal the full picture of neural activities during syntactic parsing. In recent years, an increasing number of studies have discovered potential connections between neural oscillations and syntactic parsing and have begun investigating how neural oscillations can explain different psychological processes in syntactic parsing (Bastiaansen et al., 2011), providing an alternative viable path for syntactic parsing research.

Neural oscillations are rhythmic activities exhibited by cortical neural populations across spatiotemporal scales. These activities are typically divided into five main frequency bands: delta (δ: 0.5–4 Hz), theta (: 4–8 Hz), alpha (α: 8–12 Hz), beta (β: 12–30 Hz), low gamma (γ: 30–60 Hz), and high gamma (60–200 Hz) (Prystauka & Lewis, 2019). Based on frequency, neural oscillations can be categorized as low-frequency oscillations (covering δ and bands) and high-frequency oscillations (covering α, β, and γ bands). As periodic neural activities, a prominent feature of neural oscillation-encoded signals is their temporal scale, with different frequency bands associated with information processing at different temporal scales. Consequently, the spatiotemporal properties of neural oscillations make them natural candidate mechanisms for encoding information with salient temporal structural features. Specifically, low-frequency neural oscillations, due to their longer cycles, are potentially relevant for information exchange across larger temporal windows; conversely, high-frequency neural oscillations, with shorter cycles, may be linked to information exchange within smaller temporal windows (Zhang Lixin et al., 2017). This characteristic suggests that the temporal structure of neural oscillations bears some relationship to the temporal dynamics of constructing different linguistic hierarchical structures (words, phrases, and sentences) during natural language processing, implying potential connections between neural oscillations and syntactic parsing at different scales (Hu Ruichen et al., 2019). Furthermore, neural oscillations can provide relatively precise temporal signatures for information processing and transmission within brain systems, which aligns with the requirements for robust syntactic parsing.

Linguistic theory has offered rich discussions of the syntactic parsing process, while neuroscience research indicates that human neural oscillatory activities may have intuitively plausible associations with the temporal dynamics of syntactic parsing. Bridging the gap in the mapping problem requires combining the theories and research paradigms of these two disciplines. Indeed, an increasing number of recent studies have integrated linguistic theory with neuroscience research paradigms, using neural oscillation-related indices to explore the psychological processes of syntactic parsing, providing powerful explanations for the psychological reality of hierarchical structure construction and revealing potential neural coding activities of syntactic parsing, thereby advancing solutions to the mapping problem.

3. Probing Syntactic Parsing: The Psychological Reality of A Priori Grammatical Knowledge

The hierarchical structures formed by Merge operations constitute a core feature of natural language. Before examining the specific incremental conversion process, it is necessary to establish the psychological reality of syntactic processing during language hierarchy generation. The generation of linguistic hierarchical structures relies on representations of a priori grammatical knowledge rather than being merely the result of prosodic, semantic, or other modules. From this perspective, the generation of natural language hierarchical structures is based on a priori grammatical knowledge—that is, listeners rely on existing grammatical knowledge to parse incoming stimuli through a top-down process, which is a necessary condition for syntactic parsing. From the standpoint of syntactic parsing, if evidence demonstrates that the Principle of Structural Dependency (the principle governing hierarchical structure construction) genuinely exists—that is, if hierarchical structure representation during online sentence processing is syntactically independent of prosodic, semantic, and lexical factors—then the psychological reality of syntactic processing can be considered supported.

Low-frequency neural oscillation research has made important contributions to this discussion. Humans can achieve effective language comprehension based on different types of information and cues, and low-frequency oscillatory activities observed during this process show significant correlations with syntactic parsing tasks. However, considerable controversy remains regarding which specific psychological process in sentence processing is reflected by low-frequency oscillatory activity. On one hand, some scholars argue that low-frequency oscillatory activity reflects endogenous neural oscillatory representations of abstract linguistic hierarchical structures—a relatively independent process driven by individuals' prior knowledge (Ding, 2023; Giraud, 2020; Haegens, 2020; Kandylaki & Kotz, 2020; Klimovich-Gray & Molinaro, 2020; Lewis, 2020). This view is known as the Hierarchical Structure Building (HSB) hypothesis. On the other hand, some research suggests that low-frequency oscillatory activity originates from tracking external rhythms (Keitel et al., 2018), predictive processing of statistical cues (Frank & Christiansen, 2018), and lexical representations (Frank & Yang, 2018). Although low-frequency oscillatory activity in actual speech perception results from the interaction of different information sources (Chen Liangjie et al., 2022), researchers can still distinguish syntactic structure representation from other psychological processes through experimental design and variable control, thereby finding evidence for correlations between low-frequency oscillatory activity and syntactic hierarchical structures.

3.1 Grammatical Knowledge and Prosodic Information

Because syntactic structure and prosodic information progress along similar macro-temporal scales, the neural activities they elicit are often easily confounded. Some research suggests that low-frequency oscillatory activity primarily supports speech perception and information extraction by tracking prosodic information such as speech envelopes (Poeppel & Assaneo, 2020). Meyer et al. (2020a, 2020b), however, propose a different view, arguing that such low-frequency activity essentially reflects the inference and tracking of linguistic structure itself—a process of intrinsic synchronization between external input and prior grammatical knowledge. In fact, mounting evidence indicates that attributing low-frequency oscillatory activity solely to tracking external rhythms cannot fully explain the origins of syntactic representation; the consistency between low-frequency oscillations and syntactic computational procedures is more pronounced, suggesting that they more likely represent intrinsic representations of grammatical knowledge itself.

Regarding the prosody-syntax interface, syntactic structure and prosodic information exhibit a non-isomorphic relationship. First, language externalization is a "platter-style" implementation (Si, 2016; Si Fuzhen, 2024), because although "speech production uses internal language, it cannot be identified with it. A computationally tractable theory of language enables us to distinguish the internalized system of linguistic knowledge from the processes that access it, and this is a very important distinction" (Chomsky, 2017, p. 2). Second, the presentation frequency of natural language itself is not fixed. Although research indicates that natural speech is quasi-rhythmic, this does not mean that rhythmic differences between different natural speech samples can be ignored (Kazanina & Tavano, 2023). Additionally, the relationship between prosodic boundaries and syntactic constituent boundaries in natural speech is relatively loose, with no strict correspondence between them (Ma Baopeng & Zhuang Huibin, 2022). This means that segmentation and extraction of different components in natural speech cannot be accomplished solely through acoustic cues. Syntactic categories themselves are not defined by acoustic boundaries, and the compositionality of sentence semantics must be constructed through relationships between syntactic constituents, which cannot be achieved merely by tracking speech envelopes (Meyer et al., 2020b).

Precisely because of this non-isomorphic relationship between syntactic structure and external rhythms, humans can successfully comprehend sentences even in the absence of clear acoustic cues or prosodic boundaries by relying on intrinsic grammatical knowledge. Ding et al. (2016) used a frequency tagging paradigm to present different levels of linguistic structure (words, phrases, sentences) at different stimulation frequencies, finding that even when prosodic boundaries were artificially removed from sentences, δ oscillatory responses at approximately 1, 2, and 4 Hz could still be successfully observed. However, when the same Chinese materials were presented to English native speakers, only the 4 Hz response was observed, demonstrating that low-frequency oscillatory activity essentially represents comprehensible linguistic hierarchical structures rather than tracking the rhythm of stimulus materials. Other studies have also found that when linguistic stimulus materials are presented at nonperiodic frequencies, researchers can still observe 1, 2, and 4 Hz responses representing syntactic structure processing (Ding et al., 2016; Jin et al., 2018). Moreover, even in the absence of external rhythmic stimulation—such as during visual presentation of stimuli (Henke et al., 2023) or imagined speech (Lu et al., 2021)—low-frequency neural oscillatory responses similar to normal syntactic parsing can be observed. Conversely, when external rhythms are normal but syntactic information is inaccessible—such as in word list conditions without syntactic structure (Kaufeld et al., 2020), perception of non-native materials (Ding et al., 2016), or reversed speech conditions (Gross et al., 2013; Mai et al., 2016)—corresponding neural oscillation indices show attenuation.

These empirical studies successfully dissociate neural oscillatory representations of external rhythmic stimulation from those of intrinsic grammatical knowledge: even when external rhythms cannot provide sufficient acoustic cues, researchers can still observe low-frequency oscillatory activities related to sentence-level hierarchies; however, when external rhythmic cues are abundant but grammatical knowledge is inaccessible, corresponding oscillatory responses are not necessarily observed. This indicates that low-frequency neural oscillations reflect the inferential process of intrinsic grammatical knowledge rather than simply bottom-up perception of external rhythms.

3.2 Syntactic Rules and Semantic Statistical Cues

Extensive research demonstrates that humans can predict speech perception content through semantic statistical cues (Dikker et al., 2010; Poeppel et al., 2008). For example, word frequency significantly affects single-word recognition time (Gardner et al., 1987). Transitional probability between words also shows strong correlations with reading time for words in sentences (Smith & Levy, 2013). In noisy environments, sentences with higher transitional probabilities between words are recognized better than those with lower probabilities (Miller et al., 1951). Based on the role of probabilistic cues in speech processing, some scholars argue that syntactic parsing is essentially a product of semantic analysis and statistical probability prediction, suggesting that statistical information conveyed by linear sequences themselves is more important for language comprehension than syntactic hierarchical structures. Therefore, response patterns of low-frequency neural oscillations can also be realized in statistical models targeting only linear sequential relationships (Frank & Christiansen, 2018). However, although statistical probabilities can assist and compensate for language comprehension, this paper argues that rule-based hierarchical syntactic construction is the primary driver of language comprehension, and semantic statistical probabilities alone are insufficient to fully explain the language understanding process.

The hierarchical structure of natural language syntax itself cannot be explained solely through statistical probabilities. Statistical probabilities are essentially based on simple linear sequences, whereas relationships between phrases in natural language are not simple adjacent dependencies but involve complex long-distance dependencies (Berwick et al., 2013). This indicates that syntactic parsing requires syntactic rules for hierarchical structure construction, and relying only on linear probabilities cannot generate natural language with recursive properties (Ding et al., 2017). Indeed, research separating syntactic chunks from specific lexical semantics has found that low-frequency neural oscillations more likely result from hierarchical syntactic construction rather than from lexical semantics and probability analysis (Jin et al., 2020; Y. Lu et al., 2023).

Furthermore, studies show that humans can achieve language comprehension with low dependence on transitional probabilities but cannot understand sentences without relying on syntactic hierarchical structures. Ding et al. (2016) controlled transitional probabilities in their low-frequency oscillation experiments by constructing Markovian Sentence Sets (MSS) and found that the magnitude of transitional probability did not significantly modulate δ oscillatory response strength. Additionally, syntactic structure itself can modulate statistical prediction. Using probabilistic cues for prediction often requires dependence on syntactic hierarchical structure. Slaats et al. (2023) analyzed previous MEG datasets and found that context-driven probabilistic cue effects were significant only in sentence conditions, not in word list conditions. This indicates that low-frequency oscillatory responses during speech perception are modulated jointly by syntactic structure and semantics, not merely by differences in transitional probabilities between lexical semantics. The use of probabilistic cues itself also requires syntactic hierarchical structure as a condition. Similarly, Rafferty et al. (2023) used minimal phrase and Jabberwocky paradigms to find that low-frequency neural oscillatory responses at the phrase level (0.5 Hz) were primarily related to syntactic combination rather than specific lexical semantic properties. In addition to low-frequency oscillation research, other psycholinguistic experiments have also distinguished the neural mechanisms of syntactic and semantic processing (Brennan & Hale, 2019; Pallier et al., 2011), supporting their relative independence.

3.3 Syntactic Hierarchy and Lexical Properties

Beyond excluding contributions from prosodic and semantic modules, Frank and Yang (2018) proposed the lexical representation account regarding the frequency band response patterns of low-frequency oscillations during syntactic parsing. This account suggests that observed frequency band responses are not related to syntactic hierarchical structure but rather to the presentation frequency of lexical items belonging to specific syntactic categories in the stimulus materials. In other words, the 1 Hz low-frequency neural oscillatory response does not reflect the sentence level itself but rather responds to lexical properties presented at 1 Hz frequency in the sentence. For example, in the sentence “老牛吃草” (the old cow eats grass), verbs (“吃”) appear once, nouns (“牛”, “草”) appear twice, and individual character representations (“老”, “牛”, “吃”, “草”) appear four times, corresponding to presentation frequencies of 1, 2, and 4 Hz respectively. Although such theories attribute low-frequency oscillatory activity to lexical syntactic categories, they do not acknowledge the role of syntactic hierarchical structure in syntactic parsing.

In response, Jin (2020) demonstrated through chunking tasks that low-frequency neural oscillatory responses are generated by rule-based chunking processes rather than being byproducts of lexical property combinations. Other studies have used reversed phrases that preserve the frequency and distribution of lexical properties from baseline sentences while removing original syntactic hierarchical structure information, ultimately failing to observe the corresponding 1 Hz neural index under such conditions (Lo et al., 2022). These results contradict the lexical representation account and confirm the necessity of hierarchical structure for low-frequency neural oscillatory responses.

In summary, this paper argues that low-frequency neural oscillatory activity more likely reflects neural mechanisms dedicated to syntactic structure, genuinely representing the “Principle of Structural Dependency” in linguistic hierarchical structure computation rather than being a byproduct of speech processing or probability prediction. Admittedly, during actual speech perception, syntactic parsing results from joint modulation of endogenous oscillations, external rhythms, probability prediction, and lexical semantics, and at certain moments the neural circuits corresponding to low-frequency oscillatory activity are difficult to distinguish anatomically (Meyer et al., 2020a). However, this does not mean that separating these different representations is worthless. On the contrary, such separation helps us dissociate different modules in linguistic hierarchical structure representation and provides explanations for the psychological reality of syntactic processing during parsing.

4. Tracing the Neural Coding Process of Syntactic Parsing

As previously discussed, researchers can dissociate multiple information representations in syntactic parsing through neural oscillations. Most relevant experiments achieve this by observing response strength (power) or amplitude of different neural oscillations. However, to further explain the neural coding process of syntactic parsing and align neural oscillatory activity with linguistic hierarchical structures, a deeper understanding of the more temporally specific physical properties of neural oscillations is required. Given the psychological reality of syntactic processing, how is the integration of different types of information during syntactic parsing realized through coordination between neural oscillations?

Recently, an increasing number of theoretical models and empirical studies have begun using phase coherence of neural oscillations to investigate syntactic parsing processes (Bai et al., 2022; Brennan & Martin, 2019; Calmus et al., 2020; Ding, 2020; Flanagan & Goswami, 2018; Ghitza, 2011; Martin, 2020; Martin &

Doumas, 2017; Murphy, 2024). This approach aligns signal transformations of neural oscillations with the temporal dynamics of linguistic hierarchical structure construction, representing a deeper investigation into the neural coding activities of syntactic parsing. Research suggests that representations of syntactic hierarchical structure by neural populations may be accomplished through phase synchronization and desynchronization processes of oscillations within the same frequency band or across different frequency bands. In other words, different neural populations are gradually activated or inhibited with processing time, generating oscillatory signals with different phase properties to store, maintain, and extract sensory information at different temporal scales, ultimately achieving dynamic coding of hierarchical structure construction.

4.1 Phase Coherence

The spatial distribution characteristics of neural oscillations primarily depend on their phase properties. Phase refers to the position of an oscillatory signal within its periodic waveform. Phase coherence refers to the degree of phase synchronization between oscillatory activities of different neurons and serves as an important index for measuring coordinated activity between neurons. Phase synchronization manifests in various forms, such as phase synchronization between different brain regions within the same frequency band (Fries, 2005) and cross-frequency coupling (CFC) between different frequency bands. During syntactic parsing, neural oscillations within and across frequency bands exhibit varying degrees of phase coherence. Therefore, phase coherence is thought to reflect the process of integrating information across different temporal scales for representations at different linguistic levels (Giraud & Poeppel, 2012).

Specifically, studies using CFC and related indices have found that during neural oscillatory tracking of linguistic structures, systematic phase synchronization phenomena appear both within and across frequency bands, including low-frequency coupling (e.g., δ- coupling) and low-high frequency coupling (e.g., -γ coupling) (Bai et al., 2022; Brennan & Martin, 2019). In this coupling process, low-level sensory information carried by high-frequency neural oscillations is matched with high-level linguistic structures represented by low-frequency neural oscillations, representing a crucial step in merging linguistic linear sequences into hierarchical structures. Although current research has not yet discovered strict alignment relationships between specific linguistic levels and synchronized frequency bands, plausible hypotheses about the syntactic parsing process can still be formulated based on existing findings. As shown in Figure 1 [Figure 1: see original paper], the phase coupling signal patterns of neural oscillations align with both the horizontal representation and vertical structure of syntax tree notation (see Marcolli et al., 2023a, 2023b) and exhibit a non-isomorphic mapping relationship with the phase properties of neural oscillations. That is, the correspondence between neural oscillatory activity signals and syntactic structure may be non-linear, but the two can be transformed through certain mathematical operations (see Murphy, 2024; Coopmans et al., 2023; Kazanina

& Tavano, 2023).

4.2 Theoretical Models

Several theoretical models have recently provided detailed descriptions of neural oscillatory coding of syntactic parsing from different perspectives based on phase coherence (see Table 1). The ROSE (Representations, Operations, Structures, Encoding) model focuses on neural signal transformation patterns such as cross-frequency coupling, emphasizing the establishment of connections between syntactic operations and neural activities (Murphy, 2024). The CNAL (Compositional Neural Architecture for Language) model, based on neural manifolds, provides an explanatory mechanism for transformations between different modalities (phonology, lexicon, syntax, semantics) during semantic composition. SMMM (Structure-based Memory Maintenance Model), as a classic psychological model, is characterized by its clear concretization of the role of syntactic short-term and long-term working memory in maintaining structural representations (Ding, 2020). DORA (Discovery of Relations by Analogy), as a symbolic-connectionist model, proposes a possible generation and prediction mechanism for argument structure (Martin & Doumas, 2017). VS-BIND (Vector-symbolic Sequencing of Binding INstantiating Dependencies) uses circular convolution of vectors (operator “ ”) to simulate -γ oscillatory coupling, thereby representing syntactic dependencies (Calmus et al., 2020).

These models propose hypotheses about syntactic parsing and its neural coding activities from different dimensions, possess strong explanatory power within their applicable scopes, and have received some empirical support. However, due to different disciplinary perspectives and research paradigms, these models are not directly comparable. To more comprehensively introduce the characteristics and similarities/differences of each model, this paper attempts to systematically evaluate them within David Marr's three-level framework.

David Marr (1982)'s three-level theory of human brain information processing mechanisms provides an important reference framework for relevant cognitive science research. The three levels are: computational level, algorithmic level, and implementational level. In the context of this paper, these levels respectively focus on: (1) basic computational functions of language, such as the use of grammatical rules; (2) neural oscillatory activities, examining their temporal processes and regulation in language processing; and (3) underlying biological mechanisms, such as the operation of genes, proteins, neurons, and brain circuits. Following this three-level framework, this paper will discuss the similarities and differences of these models from the perspectives of computational, algorithmic, and implementational levels to assess their explanatory adequacy for syntactic parsing processes (see Figure 2 [Figure 2: see original paper]).

Within this unified theoretical framework, the aforementioned models share three commonalities. First, at the implementational level, these models are similar in explaining the underlying biological mechanisms of syntactic parsing, as they all premise that neural populations are gradually and incrementally activated during syntactic processing to explain neural oscillatory activities representing syntactic structure. That is, as external stimuli are continuously input, different neural populations are successively activated during the formation of linguistic hierarchical structures, generating rhythmic neural signals to achieve dynamic representation of linguistic levels. Notably, the current five models have not yet involved deeper discussions at the neurophysiological and dynamical levels. Second, at the algorithmic level, these models all assume that the key to representing linguistic hierarchical structures lies in the temporal coupling of neural signals, explaining the syntactic parsing process through their phase coherence. Finally, at the computational level, although these five models are based on different linguistic theories and explain different structural types, they all affirm the existence of hierarchical structures in linguistic phenotypes and discuss this point. Therefore, overall, these models reveal the feasibility of neural oscillations as a mechanism for syntactic structure construction, considering neural oscillations as a potential neural mechanism for syntactic parsing.

Meanwhile, these models also exhibit three differences. The most significant difference lies at the algorithmic level. These models offer different interpretations of the specific connections between different frequency bands of neural oscillations and different linguistic levels: ROSE maps full-band neural activities to syntactic computation processes in Minimalist Program theory, assuming that γ oscillatory activity represents the formation of syntactic objects (i.e., lexical items), -γ coupling represents lexical retrieval processes, δ- coupling represents hierarchical structure Merge, α oscillations are related to inhibitory activities, and β oscillations are related to predictive activities. CNAL is similar to ROSE, differing only in that it does not discuss the roles of α and β bands in this process. SMMM and DORA primarily emphasize the central role of δ oscillations in linguistic hierarchy implementation, with less discussion of other frequency bands. VS-BIND explains syntactic processing based solely on -γ coupling.

Based on current empirical findings, δ band research has successfully excluded influences from prosodic, semantic, and lexical factors, demonstrating that δ oscillations are more relevant to syntactic hierarchical structure construction (Ding et al., 2016), while oscillations are more likely to covary with speech envelopes of stimulus materials (Chalas et al., 2023), and γ oscillations are more likely related to acoustic feature structures of stimulus materials (Xu, Zhao et al., 2023). This suggests that -γ coupling may represent lexical formation, a process that occurs before linguistic hierarchical structure construction. From this perspective, VS-BIND is less credible than other models, while the neural mechanisms of the other four models are supported by empirical research. For α and β oscillations, cognitive function-related studies also support their respective associations with inhibition and timing (Klimesch, 2012) and predictive activities (Abbasi & Gross, 2020). Therefore, ROSE provides a more comprehensive explanation of the syntactic parsing process. Additionally, not all models have conducted mathematical-level simulations of these neural signal activity patterns. CNAL uses the neural manifold data model, assuming that signal patterns of neural populations can be reduced to embedded manifolds in some high-dimensional space, and explains how external stimuli are transformed between different modules through coordinate transformations in subspaces. VS-BIND uses circular convolution operations to simulate -γ oscillatory coupling, thereby explaining adjacent and non-adjacent relationships between syntactic constituents. In contrast, ROSE, DORA, and SMMM have not used specific mathematical tools to interpret signal patterns of neural oscillations representing syntactic hierarchies.

Beyond algorithmic differences, these models also differ in the linguistic ontological structures they focus on at the computational level. Specifically, ROSE and CNAL systematically involve the full spectrum from atomic features to hierarchical structure construction, offering interpretations for transformations of external input across phonological, lexical, syntactic, and semantic modules, making them finer-grained models. In contrast, SMMM, DORA, and VS-BIND focus on explaining the process of syntactic hierarchical structure formation itself, discussing how abstract syntactic memory devices maintain and predict hierarchical structures. These are relatively coarse-grained models but more intuitively demonstrate the psychological process of syntactic hierarchical construction.

From the perspective of model categories and core content, the underlying operating mechanisms of these models involve different disciplinary research paradigms with substantial differences. ROSE is based on linguistic theoretical frameworks, systematically mapping neural oscillations across frequency bands and their activities to specific steps in syntactic parsing, emphasizing the relationship between linguistic phenotypes and neural activities. CNAL introduces cutting-edge neural manifold concepts, belonging to computational neuroscience research paradigms. SMMM is a classic psychological mechanism model whose core lies in concretizing the maintenance and representation of syntactic units and highlighting the role of syntactic short-term and long-term memory devices. DORA and VS-BIND are both symbolic-connectionist models that use neural networks to describe syntactic parsing processes. Therefore, these models are to some extent incomparable; their differences do not reflect contradictions but rather explanations of syntactic parsing processes from different dimensions.

In summary, ROSE and CNAL, from a fine-grained perspective, comprehensively map neural oscillatory activities across different frequency bands to syntactic parsing processes, are supported by substantial empirical research, can better solve the mapping problem, and possess good developmental potential. Although the other three models are not as fine-grained, they still possess strong explanatory power in concretizing the behavioral process of syntactic hierarchical construction and provide references for subsequent research.

In conclusion, the potential connection between phase coherence and syntactic hierarchical structures has gradually attracted attention. Although rigorous mathematical proofs and dynamical computational models linking neural oscillations to syntactic parsing processes are currently lacking, the significant temporal distribution characteristics of neural oscillatory phase properties can indeed provide a potentially reasonable explanation for the neural coding activities of syntactic parsing.

5. Summary and Outlook

Hierarchical syntactic structure is an essential property distinguishing natural language from other communication systems. The process of syntactic parsing must explain the incremental conversion between linear speech sequences and syntactic hierarchical structures. During this process, although cues such as prosody and statistical probability participate in processing, they are insufficient for syntactic parsing; the formation of robust syntactic structure representations still requires a priori grammatical knowledge. Reviewing existing neural oscillation research, the field has accumulated rich evidence supporting intrinsic syntactic representations of abstract linguistic structures and demonstrated enormous potential for using neural oscillations to explain the neural coding processes of syntactic parsing. Neural oscillations have become an important entry point for exploring the mapping problem between language and neural primitives. Future research should focus on the following aspects:

First, investigate the causal neural coding activities of syntactic parsing. The exposition of the relationship between neural oscillations and linguistic structures in this paper remains speculative. To further prove finer-grained mapping relationships between the two, additional investigation is needed. Specifically, current research primarily uses amplitude and power of specific neural oscillation frequency bands as observation indices, lacking exploration of the relationship between neural oscillation phase properties and linguistic hierarchical structures. Consequently, most research results may remain at the level of discussing the existence of syntactic knowledge rather than directly proving how neural oscillations across different frequency bands represent linguistic hierarchical structures through signal transformations to achieve incremental conversion. Future research should more deeply explore connections between neural oscillation phase properties and specific psychological processes of syntactic parsing, using phase coupling-related indices (such as CFC) and experimental paradigms that can sufficiently separate different linguistic hierarchical structures (such as frequency tagging paradigms). Additionally, current research primarily relies on variable control methods from traditional psychological experiments, answering research questions through direct observation of results from different condition comparisons, which may be constrained by surface linguistic forms. Future research could encode features of syntactic structures (such as syntax tree depth and nodes) as abstract data structures for modeling and validation (see Gwilliams et al., 2024), thereby bypassing surface linguistic implementation forms to further explain potential connections between neural oscillations and abstract syntactic structures. Finally, the studies reviewed in this paper involve only correlational analyses in experimental design and statistical methods, not yet involving causal inference. With sufficient experimental data in the future, more mathematical and statistical methods can be used for analysis to advance finer-grained causal explanations of syntactic parsing, find mapping relationships between their non-strictly isomorphic structures, and align them with precise neural oscillatory activities. Current linguistic theory has attempted to use abstract algebra to depict activities at the syntactic computational level (Marcolli, Chomsky et al., 2023), and existing syntactic parsing models have attempted to match mathematical models of syntactic hierarchical construction with specific neurophysiology (Kaushik & Martin, 2022). Future research can further explore the neural mechanisms of syntactic parsing under the guidance of these theoretical frameworks.

Second, clarify the generation mechanisms of neural oscillations and their biological significance. The implementational-level mechanisms of neural oscillatory activity remain controversial: First, there is considerable difficulty in distinguishing neural oscillations from evoked responses. Neural oscillations themselves may emerge from combinations of countless evoked responses (Ding, 2022), while evoked responses themselves may arise from average modulation of oscillatory activities (Meyer et al., 2020a). Second, neural oscillation indices observed by different imaging techniques have essential differences, and due to the spatial resolution limitations of EEG and MEG, current research also struggles to map oscillatory activities to macroscopic or microscopic neural structures. Neurophysiological research indicates that the periodicity of neural oscillation signals may be related to activation-inhibition cycling mechanisms of neural populations. This paper focuses on discussing their potential as principles and explanatory mechanisms for syntactic construction and therefore does not systematically review these neural activities themselves. To further investigate syntactic parsing processes in the future, it will be necessary to explore causality among these neural activities, connect the spatial distribution of neural oscillations with their temporal structure, map them to specific patterns of neural population excitation transmission, clarify their specific neurophysiological foundations, or identify gene expressions related to linguistic hierarchical structure through genetic studies, thereby bridging the gap between the algorithmic and implementational levels of syntactic parsing.

Third, focus on developmental changes in neural oscillation patterns during language acquisition. From the perspective of language acquisition, children's acquisition of syntactic structure manifests as a transition from reliance on external rhythms to the formation of intrinsic grammatical knowledge. Early language acquisition in infants exhibits perceptual narrowing (Kuhl, 2004) and developmental weight shifting (Seidl, 2007). This may occur because as relevant cortical areas develop, infants' identification and discrimination of linguistic input gradually shift from dependence on acoustic features to reliance on grammatical knowledge (Meyer et al., 2020b). Currently, relevant hypotheses and research primarily examine the role of neural oscillations in child language acquisition from the perspective of tracking external rhythms (Attaheri et al., 2022; Nallet & Gervain, 2021). Subsequent research could use frequency tagging paradigms and other methods to further separate grammatical knowledge from prosodic information to investigate their different roles in language acquisition.

Fourth, identify the neurophysiological deficits in different language disorder populations. From the perspective of language attrition, age-related decline in language ability in elderly individuals manifests in neural oscillatory activities as enhanced oscillatory responses for lower-level linguistic structures (4 Hz) and weakened responses for higher-level structures (1, 2 Hz) (Xu et al., 2023), reflecting the impact of abstract grammatical knowledge decline on low-frequency neural oscillatory activity. Future research could use low-frequency neural oscillations as an entry point to investigate specific neural manifestations of language attrition in elderly populations. From the perspective of special populations, previous studies have found that compared to healthy populations, language disorder populations show certain variations in neural oscillatory activities (Fridriksson et al., 2015; Meyer et al., 2021; Peter et al., 2023), suggesting that their intrinsic grammatical knowledge representations may have certain deficits. Particularly for individuals with Neural Developmental Disorder (NDD), whose language disorder mechanisms lack clear pathological definitions and often cannot have brain lesion locations identified through high spatial resolution brain imaging techniques (Abbott & Love, 2023), neural oscillation indices could be used to investigate their linguistic behavioral deficits. Existing theories have attempted to explain the neural deficits in Developmental Dyslexia (DD) from the perspective of endogenous oscillatory activities (Goswami, 2011), and future research could expand to Developmental Language Disorder (DLD) populations that are more difficult to define. Research on language disorder populations can not only provide pathological evidence for neural oscillatory mechanisms in syntactic parsing but also help overcome heterogeneity issues in research on some language disorder populations, providing more unified pathological standards for early prediction and clinical intervention of these disorders.

References

Chen, L., Liu, L., Ge, Z., Yang, X., & Li, L. (2022). The role of rhythm in auditory speech comprehension. *Advances in Psychological Science*, 30(8), 1818–1831.

Hu, R., Yuan, P., Jiang, Y., & Wang, Y. (2019). The role of temporal structure information in human perception and its brain mechanisms. *Acta Physiologica Sinica*, 71(1), 105–116.

Jiang, M. (2009). Syntactic autonomy: Controversies and evidence. *Foreign Languages and Literature*, 25(3), 79–86.

Ma, B., & Zhuang, H. (2022). A review and implications of prosody-syntax interface research over the past two decades. *Journal of Foreign Languages*, 45(1), 56–66.

Si, F. (2024). The “Galilean puzzle” in language and brain research. *Journal of Foreign Languages*, 47(2), 2–9.

Yang, L. (2012). A cross-linguistic review of recursion-only hypotheses. *Foreign Language Teaching and Research*, 44(01), 54–64+158.

Zhang, L., Wang, F., Wang, L., Yang, J., & Wan, B. (2017). Research progress on cross-frequency coupling of neural oscillations in cognitive function studies. *Acta Physiologica Sinica*, 69(6), 805–816.

Abbasi, O., & Gross, J. (2020). Beta-band oscillations play an essential role in motor-auditory interactions. *Human Brain Mapping*, 41(3), 656–665.

Abbott, N., & Love, T. (2023). Bridging the divide: Brain and behavior in developmental language disorder. *Brain Sciences*, 13(11), 1606.

Attaheri, A., Choisdealbha, Á. N., Di Liberto, G. M., Rocha, S., Brusini, P., Mead, N., Olawole-Scott, H., Boutris, P., Gibbon, S., Williams, I., Grey, C., Flanagan, S., & Goswami, U. (2022). Delta- and theta-band cortical tracking and phase-amplitude coupling to sung speech by infants. *NeuroImage*, 247, 118698.

Bai, F., Meyer, A. S., & Martin, A. E. (2022). Neural dynamics differentially encode phrases and sentences during spoken language comprehension. *PLoS Biology*, 20(7), e3001713. https://doi.org/10.1371/journal.pbio.3001713

Bastiaansen, M., Mazaheri, A., & Jensen, O. (2011). Beyond ERPs: Oscillatory neuronal dynamics. In E. S. Kappenman & S. J. Luck (Eds.), *The Oxford Handbook of Event-Related Potential Components* (p. 0). Oxford University Press.

Berwick, R. C., Friederici, A. D., Chomsky, N., & Bolhuis, J. J. (2013). Evolution, brain, and the nature of language. *Trends in Cognitive Sciences*, 17(2), 89–98.

Brennan, J. R., & Hale, J. T. (2019). Hierarchical structure guides rapid linguistic predictions during naturalistic listening. *PloS One*, 14(1), e0207741. https://doi.org/10.1371/journal.pone.0207741

Brennan, J. R., & Martin, A. E. (2019). Phase synchronization varies systematically with linguistic structure composition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 375(1791), 20190305.

Calmus, R., Wilson, B., Kikuchi, Y., & Petkov, C. I. (2020). Structured sequence processing and combinatorial binding: Neurobiologically and computationally informed hypotheses. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 375(1791), 20190304.

Chalas, N., Daube, C., Kluger, D. S., Abbasi, O., Nitsch, R., & Gross, J. (2023). Speech onsets and sustained speech contribute differentially to delta and theta speech tracking in auditory cortex. *Cerebral Cortex*, 33(10), 6273–6281.

Chomsky, N. (1957). *Syntactic Structures*. De Gruyter Mouton.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. The MIT Press.

Chomsky, N. (1995). *The Minimalist Program*. The MIT Press.

Chomsky, N. (2017). The Galilean challenge: Architecture and evolution of language. *Journal of Physics: Conference Series*, 880, 012–015.

Chomsky, N., & Moro, A. (2022). *Secrets of Words*. The MIT Press.

Chomsky, N., Seely, T. D., Berwick, R. C., Fong, S., Huybregts, M. a. C., Kitahara, H., McInnerney, A., & Sugimoto, Y. (2023). Merge and the strong minimalist thesis. Cambridge University Press.

Coopmans, C. W., Mai, A., Slaats, S., Weissbart, H., & Martin, A. E. (2023). What oscillations can do for syntax depends on your theory of structure building. *Nature Reviews Neuroscience*, 24(11), 723–723.

Dikker, S., Rabagliati, H., Farmer, T. A., & Pylkkänen, L. (2010). Early occipital sensitivity to syntactic category is based on form typicality. *Psychological Science*, 21(5), 629–634.

Ding, N. (2020). A Structure-based Memory Maintenance Model for Neural Tracking of Linguistic Structures (arXiv:2002.11870). arXiv. https://doi.org/10.48550/arXiv.2002.11870

Ding, N. (2022). The Neural Correlates of Linguistic Structure Building: Comments on Kazanina & Tavano (2022) (arXiv:2212.04219). arXiv. https://doi.org/10.48550/arXiv.2212.04219

Ding, N. (2023). Low-frequency neural parsing of hierarchical linguistic structures. *Nature Reviews Neuroscience*, 24(12), 792–792.

Ding, N., Melloni, L., Tian, X., & Poeppel, D. (2017). Rule-based and word-level statistics-based processing of language: Insights from neuroscience. *Language, Cognition and Neuroscience*, 32(5), 570–575.

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

Flanagan, S., & Goswami, U. (2018). The role of phase synchronisation between low frequency amplitude modulations in child phonology and morphology speech tasks. *The Journal of the Acoustical Society of America*, 143(3), 1366–1375.

Frank, S. L., & Christiansen, M. H. (2018). Hierarchical and sequential processing of language: A response to: Ding, Melloni, Tian, and Poeppel (2017). Rule-based and word-level statistics-based processing of language: insights from neuroscience. *Language, Cognition and Neuroscience*, 33(9), 1213–1218.

Frank, S. L., & Yang, J. (2018). Lexical representation explains cortical entrainment during speech comprehension. *PLoS One*, 13(5), e0197304. https://doi.org/10.1371/journal.pone.0197304

Fridriksson, J., Basilakos, A., Hickok, G., Bonilha, L., & Rorden, C. (2015). Speech entrainment compensates for Broca's area damage. *Cortex*, 69, 68–75.

Fries, P. (2005). A mechanism for cognitive dynamics: Neuronal communication through neuronal coherence. *Trends in Cognitive Sciences*, 9(10), 474–480.

Gardner, M. K., Rothkopf, E. Z., Lapan, R., & Lafferty, T. (1987). The word frequency effect in lexical decision: Finding a frequency-based component. *Memory & Cognition*, 15(1), 24–28.

Ghitza, O. (2011). Linking speech perception and neurophysiology: Speech decoding guided by cascaded oscillators locked to the input rhythm. *Frontiers in Psychology*, 2, 130.

Giraud, A.-L. (2020). Oscillations for all ¯*(ツ)*/¯? A commentary on Meyer, Sun & Martin (2020). *Language, Cognition and Neuroscience*, 35(9), 1106–1113.

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

Goswami, U. (2011). A temporal sampling framework for developmental dyslexia. *Trends in Cognitive Sciences*, 15(1), 3–10.

Gouvea, A., Phillips, C., Kazanina, N., & Poeppel, D. (2010). The linguistic processes underlying the P600. *Language and Cognitive Processes*, 25, 149–188.

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

Gwilliams, L., Marantz, A., Poeppel, D., & King, J. R. (2024). Hierarchical dynamic coding coordinates speech comprehension in the brain. bioRxiv: the preprint server for biology.

Haegens, S. (2020). Entrainment revisited: A commentary on Meyer, Sun, and Martin (2020). *Language, Cognition and Neuroscience*, 35(9), 1119–1123.

Haegens, S., & Zion Golumbic, E. (2018). Rhythmic facilitation of sensory processing: A critical review. *Neuroscience & Biobehavioral Reviews*, 86, 150–165.

Hale, J. T., Campanelli, L., Li, J., Bhattasali, S., Pallier, C., & Brennan, J. R. (2022). Neurocomputational models of language processing. *Annual Review of Linguistics*, 8(1), 427–447.

Henke, L., Lewis, A. G., & Meyer, L. (2023). Fast and slow rhythms of naturalistic reading revealed by combined eye-tracking and electroencephalography. *Journal of Neuroscience*, 43(24), 4533–4545.

Jin, P., Lu, Y., & Ding, N. (2020). Low-frequency neural activity reflects rule-based chunking during speech listening. *eLife*, 9, e55613. https://doi.org/10.7554/eLife.55613

Jin, P., Zou, J., Zhou, T., & Ding, N. (2018). Eye activity tracks task-relevant structures during speech and auditory sequence perception. *Nature Communications*, 9(1), 5374.

Kandylaki, K. D., & Kotz, S. A. (2020). Distinct cortical rhythms in speech and language processing and some more: A commentary on Meyer, Sun, & Martin (2019). *Language, Cognition and Neuroscience*, 35(9), 1124–1128.

Kaufeld, G., Bosker, H. R., Oever, S. ten, Alday, P. M., Meyer, A. S., & Martin, A. E. (2020). Linguistic structure and meaning organize neural oscillations into a content-specific hierarchy. *Journal of Neuroscience*, 40(49), 9467–9475.

Kaushik K. R., Martin A. E. (2022). A mathematical neural process model of language comprehension, from syllable to sentence. PsyArXiv; 2022. doi: 10.31234/osf.io/xs5kr.

Kazanina, N., & Tavano, A. (2023). What neural oscillations can and cannot do for syntactic structure building. *Nature Reviews Neuroscience*, 24(2), 113-128.

Keitel, A., Gross, J., & Kayser, C. (2018). Perceptually relevant speech tracking in auditory and motor cortex reflects distinct linguistic features. *PLoS Biology*, 16(3), e2004473. https://doi.org/10.1371/journal.pbio.2004473

Klimesch, W. (2012). α-band oscillations, attention, and controlled access to stored information. *Trends in Cognitive Sciences*, 16(12), 606–617.

Klimovich-Gray, Anastasia & Molinaro, Nicola. (2020). Synchronising internal and external information: A commentary on Meyer, Sun & Martin (2020). *Language, Cognition and Neuroscience*, 35(9), 1129–1132.

Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience*, 5(11), 831–843.

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences*, 4(12), 463–470.

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.

Lewis, A. G. (2020). Balancing exogenous and endogenous cortical rhythms for speech and language requires a lot of entraining: A commentary on Meyer, Sun & Martin (2020). *Language, Cognition and Neuroscience*, 35(9), 1133–1137.

Lo, C.-W., Tung, T.-Y., Ke, A. H., & Brennan, J. R. (2022). Hierarchy, not lexical regularity, modulates low-frequency neural synchrony during language comprehension. *Neurobiology of Language*, 3(4), 538–555.

Lu, L., Sheng, J., Liu, Z., & Gao, J.-H. (2021). Neural representations of imagined speech revealed by frequency-tagged magnetoencephalography responses. *NeuroImage*, 229, 117724.

Lu, Y., Jin, P., Ding, N., & Tian, X. (2023). Delta-band neural tracking primarily reflects rule-based chunking instead of semantic relatedness between words. *Cerebral Cortex*, 33(8), 4448–4458.

Mai, G., Minett, J. W., & Wang, W. S.-Y. (2016). Delta, theta, beta, and gamma brain oscillations index levels of auditory sentence processing. *NeuroImage*, 133, 516–528.

Marcolli, M., Berwick, R. C., & Chomsky, N. (2023a). Old and New Minimalism: A Hopf algebra comparison (arXiv:2306.10270). arXiv. https://doi.org/10.48550/arXiv.2306.10270

Marcolli, M., Berwick, R. C., & Chomsky, N. (2023b). Syntax-semantics interface: An algebraic model (arXiv:2311.06189). arXiv. https://doi.org/10.48550/arXiv.2311.06189

Marcolli, M., Chomsky, N., & Berwick, R. (2023). Mathematical Structure of Syntactic Merge (arXiv:2305.18278). arXiv. https://doi.org/10.48550/arXiv.2305.18278

Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. W.H. Freeman.

Martin, A. E. (2020). A Compositional Neural Architecture for Language. *Journal of Cognitive Neuroscience*, 32(8), 1407–1427.

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

Martins, P. T., & Boeckx, C. (2019). Language evolution and complexity considerations: The no half-Merge fallacy. *PLoS Biology*, 17(11), e3000389. https://doi.org/10.1371/journal.pbio.3000389

McClamrock, R. (1991). Marr's three levels: A re-evaluation. *Minds and Machines*, 1, 185-196.

Meyer, L., Lakatos, P., & He, Y. (2021). Language dysfunction in Schizophrenia: Assessing neural tracking to characterize the underlying disorder(s)? *Frontiers in Neuroscience*, 15, 640502.

Meyer, L., Sun, Y., & Martin, A. E. (2020a). "Entraining" to speech, generating language? *Language, Cognition and Neuroscience*, 35(9), 1138–1148.

Meyer, L., Sun, Y., & Martin, A. E. (2020b). Synchronous, but not entrained: Exogenous and endogenous cortical rhythms of speech and language processing. *Language, Cognition and Neuroscience*, 35(9), 1089–1099.

Miller, G. A., Heise, G. A., & Lichten, W. (1951). The intelligibility of speech as a function of the context of the test materials. *Journal of Experimental Psychology*, 41(5), 329–335.

Murphy, E. (2024). ROSE: A neurocomputational architecture for syntax. *Journal of Neurolinguistics*, 70, 101180.

Nallet, C., & Gervain, J. (2021). Neurodevelopmental preparedness for language in the neonatal brain. *Annual Review of Developmental Psychology*, 3, 41–58.

Obleser, J., & Kayser, C. (2019). Neural entrainment and attentional selection in the listening brain. *Trends in Cognitive Sciences*, 23(11), 913–926.

Pallier, C., Devauchelle, A.-D., & Dehaene, S. (2011). Cortical representation of the constituent structure of sentences. *Proceedings of the National Academy of Sciences*, 108(6), 2522–2527.

Peter, V., Goswami, U., Burnham, D., & Kalashnikova, M. (2023). Impaired neural entrainment to low frequency amplitude modulations in English-speaking children with dyslexia or dyslexia and DLD. *Brain and Language*, 236, 105217.

Poeppel, D. (2012). The maps problem and the mapping problem: Two challenges for a cognitive neuroscience of speech and language. *Cognitive Neuropsychology*, 29(1–2), 34–55.

Poeppel, D., & Assaneo, M. F. (2020). Speech rhythms and their neural foundations. *Nature Reviews Neuroscience*, 21(6), 322–334.

Poeppel, D., Idsardi, W. J., & van Wassenhove, V. (2008). Speech perception at the interface of neurobiology and linguistics. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 363(1493), 1071–1086.

Prystauka, Y., & Lewis, A. G. (2019). The power of neural oscillations to inform sentence comprehension: A linguistic perspective. *Language and Linguistics Compass*, 13(9), e12347. https://doi.org/10.1111/lnc3.12347

Rafferty, M. B., Saltuklaroglu, T., Reilly, K., Paek, E. J., & Casenhiser, D. M. (2023). Neural synchrony reflects closure of jabberwocky noun phrases but not predictable pseudoword sequences. *The European Journal of Neuroscience*, 57(11), 1834–1847.

Seidl, A. (2007). Infants' use and weighting of prosodic cues in clause segmentation. *Journal of Memory and Language*, 57(1), 24–48.

Slaats, S., Weissbart, H., Schoffelen, J.-M., Meyer, A. S., & Martin, A. E. (2023). Delta-band neural responses to individual words are modulated by sentence processing. *Journal of Neuroscience*, 43(26), 4867–4883.

Smith, N. J., & Levy, R. (2013). The effect of word predictability on reading time is logarithmic. *Cognition*, 128(3), 302–319.

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

Traxler, M. J. (2014). Trends in syntactic parsing: Anticipation, Bayesian estimation, and good-enough parsing. *Trends in Cognitive Sciences*, 18(11), 605–611.

Xu, N., Qin, X., Zhou, Z., Shan, W., Ren, J., Yang, C., Lu, L., & Wang, Q. (2023). Age differentially modulates the cortical tracking of the lower and higher level linguistic structures during speech comprehension. *Cerebral Cortex*, 33(19).

Xu, N., Zhao, B., Luo, L., Zhang, K., Shao, X., Luan, G., Wang, Q., Hu, W., & Wang, Q. (2023). Two stages of speech envelope tracking in human auditory cortex modulated by speech intelligibility. *Cerebral Cortex*, 33(5), 2215–2228.

Note: Figure translations are in progress. See original paper for figures.