Cleaning up the Brickyard: How Theory and Methodology Affect Experimental Outcome in Cognitive Neuroscience of Language

Constantijn L. van der Burght[1], Angela D. Friederici[2], Matteo Maran[2,3], Giorgio Papitto[2,3], Elena Pyatigorskaya[2,3], Joëlle A. M. Schroën[2,3], Patrick C. Trettenbrein[2,3], and Emiliano Zaccarella[2]

[1] Psychology of Language Department, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

[2] Department of Neuropsychology, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

[3] International Max Planck Research School on Neuroscience of Communication: Structure, Function, and Plasticity (IMPRS NeuroCom), Leipzig, Germany

Author Note

Constantijn L. van der Burght https://orcid.org/0000-0001-6461-1730

Angela D. Friederici https://orcid.org/0000-0002-6328-865X

Matteo Maran https://orcid.org/0000-0002-5867-3203

Giorgio Papitto http://orcid.org/0000-0002-0132-2488

Elena Pyatigorskaya https://orcid.org/0000-0002-7320-4464

Joëlle A. M. Schroën https://orcid.org/0000-0001-8811-5790

Patrick C. Trettenbrein https://orcid.org/0000-0003-2233-6720

Emiliano Zaccarella https://orcid.org/0000-0002-5703-1778

All authors contributed equally and are listed in alphabetical order.

We have no known conflicts of interest to disclose.

Correspondence concerning should be addressed to Patrick C. Trettenbrein, Department of Neuropsychology, Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstraße 1a, 04103 Leipzig, Germany. E-mail: trettenbrein@cbs.mpg.de

Abstract

The capacity for language is a defining property of our species, yet despite decades of research, evidence on its neural basis is still mixed and a generalized consensus is difficult to achieve. We suggest that this is partly caused by researchers defining "language" in radically different ways, with focus on a wide range of phenomena, properties, and levels of investigation. Accordingly, there is very little agreement amongst cognitive neuroscientists of language on the operationalization of fundamental concepts to be investigated in neuroscientific experiments. This paper reviews chains of derivation in the cognitive neuroscience of language, focusing on how the hypothesis under consideration is defined by a combination of theoretical and methodological assumptions. We first attempt to disentangle the complex relationship between linguistics, psychology, and neuroscience in the field. We then focus on how conclusions that can be drawn from any experiment are inherently constrained not just by the research techniques and analyses adopted, but also by the theoretical starting point of the study: auxiliary assumptions, both theoretical and methodological, on which the validity of conclusions drawn rely. These issues are discussed in the context of classical experimental manipulations, as well as study designs that employ novel approaches such as naturalistic stimuli and computational methods. We conclude by proposing that a highly interdisciplinary field such as the cognitive neuroscience of language requires researchers to form explicit statements concerning the theoretical definitions, methodological choices, and other constraining factors involved in their work.

Keywords: derivation chains; auxiliary assumptions; linguistic theory; cognitive science; neuroscience; psycholinguistics; language science

Cleaning up the Brickyard: How Theory and Methodology Affect Experimental Outcome in Cognitive Neuroscience of Language

Several decades ago, a letter published in *Science* (Forscher, 1963) put forward that there was "chaos in the brickyard" of scientific inquiry. The letter concluded that it was no longer clear if scientists were still attempting to build edifices (i.e., grand theories) or whether they had simply produced a pile of individual pieces of research (i.e., bricks) that do not necessarily address an overarching research question. The highly interdisciplinary field of cognitive neuroscience of language may also be described as a rather chaotic brickyard. Researchers within this field have defined "language" in radically different ways, leading them to focus on different phenomena, properties, and levels of investigation. Moreover, researchers have used a wide range of methodological approaches to investigate their definition of language. The "brickyard" of cognitive neuroscience of language is therefore filled with idiosyncratically-shaped bricks (i.e., research results and directions) that do not necessarily fit together, making it difficult for the field to move forward.

To better understand how individual bricks may be designed and organized to build an edifice of "language", one can employ the derivation chain framework introduced by Meehl (1990). This framework describes how any theory under investigation depends on several premises, so-called auxiliary assumptions. Accordingly, conclusions drawn based on a given experimental outcome rest on the validity of these (hidden) auxiliary assumptions. If any of the auxiliary assumptions turns out to be false, the derivation chain from theory to statistical inference will be weakened or broken. This weakened derivation chain ultimately compromises the conclusions that can be drawn based on a particular experiment. In other words, the main theory is only as strong as its supporting auxiliary assumptions. These derivation chains, previously discussed in other fields (e.g., Scheel et al., 2021), are highly appropriate to describe many chains of assumptions present in the cognitive neuroscience of language.

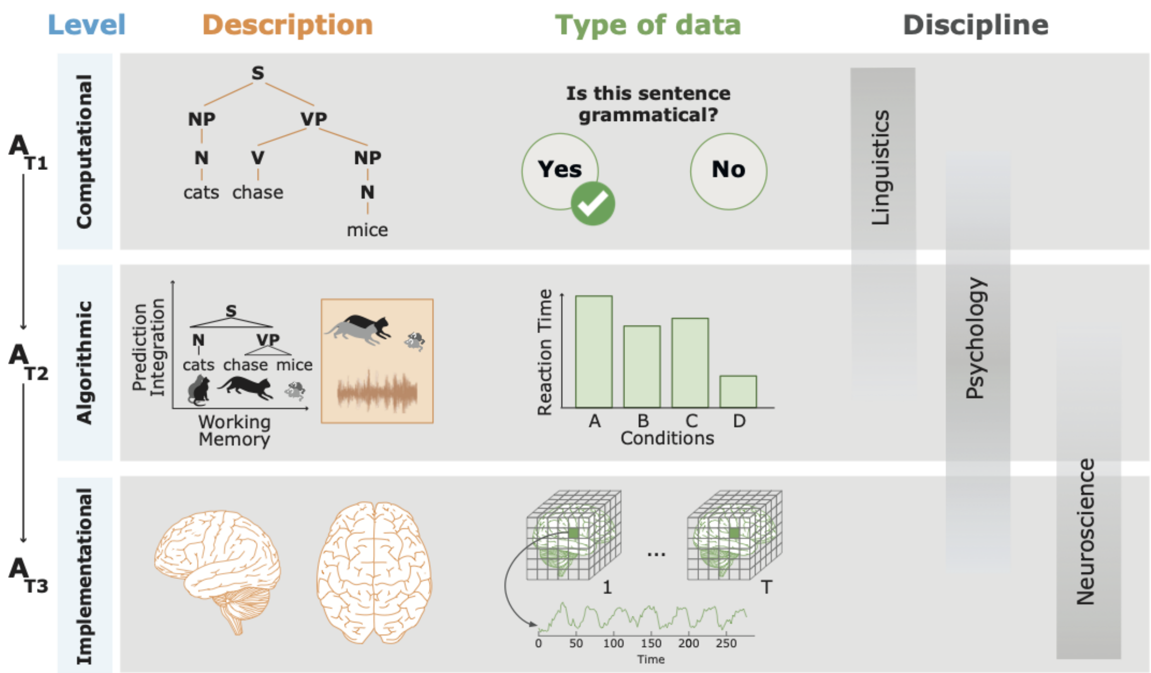
While many research fields interact with other disciplines, unique to the cognitive neuroscience of language is the possibility to build upon theories and insights of an entire academic discipline, that is, modern linguistics. The translation from linguistic theory to a statistical test in a cognitive neuroscience experiment relies, however, on a large number of auxiliary assumptions. Furthermore, additional methodological assumptions are made when designing an experiment, collecting data, and performing statistical analyses. Without paying sufficient attention to the specific elements of the resulting derivation chain, hypothesis tests in the field of cognitive neuroscience of language can be severely limited in informativeness and validity. In the present article, we discuss auxiliary assumptions (shorthand: A) of theoretical (A_T) and methodological (A_M) nature that constrain the implications of research into the cognitive neuroscience of language. Here, we suggest that this derivation chain can be formulated as:

Cognitive neuroscience of language = $A_{T1} \times A_{T2} \times A_{Tn} \times A_{M1} \times A_{M2} \times A_{Mn}$

Throughout this paper, we will give examples of auxiliary assumptions from either linguistic theory or experimental methodology in our field (see Figure 1), and we will argue that particularly in the language domain these (frequently hidden) assumptions require special attention. The main reason for this is the wealth of (psycho-)linguistic theories that the cognitive neuroscience of language has at its disposal, offering a rich base of testable hypotheses about the phenomenon under investigation. At the same time, this extensive availability of theoretical accounts introduces problems that may be unique to language research. For instance, when investigating the neural implementation of linguistic structure-building—the process by which words are combined together to form phrases and sentences—the same cognitive process on the neural level may, for example, be labeled as Merge (Friederici et al., 2017), Unification (Hagoort, 2017), or combinatorial semantics (Pylkkänen, 2019), among other definitions; according to the specific theoretical account under analysis. Here, the choice of the auxiliary theories shapes the way in which evidence for (or against) a particular hypothesis is discussed. Consequently, the presence of a wide range of distinct auxiliary assumptions that attempt to describe the same cognitive process risks adding undesirable complexity to the literature investigating the neural implementation of that process. A similar problem arises when the theoretical starting points are underspecified: Two different studies may refer to the same cognitive process using similar labels, although the underlying auxiliary assumptions are in fact dissimilar.

Figure 1

Derivation chains in cognitive neuroscience of language. Auxiliary assumptions derived from theoretical accounts (A_T) on three distinct levels of investigation: computational, algorithmic, and implementational. Each level relies on distinct descriptions of the phenomena under study (left). A_{T1} (top-row): A-temporal description of the sentence structur. A_{T2} (middle-row): Explanation in terms of general cognitive concepts such as working memory and prediction. A_{T3} (bottom-row): Description in terms of involved brain regions or observed signals. These levels are mapped to data of various nature (middle). A_{T1} (top-row): Subjective judgements of the grammaticality of a sentence. A_{T2} (middle-row): Reaction times for different experimental conditions. A_{T3} (bottom-row): Spatial and temporal time series of a brain signal related to neural activity. Each level roughly aligns with a specific scientific disciplines (right). A_{T1} (top-row): Linguistics. A_{T2} (middle-row): Psychology. A_{T3} (bottom-row): Neuroscience.



Overall, the tightness of the derivation chain in cognitive neuroscience of language depends on how well the construct under investigation is represented in linguistic theory.

The more agreement exists on a given construct or operation among different theoretical accounts, the more likely it is that the transfer from linguistic theory to neuroscience will be successful (Baggio, 2020). Nonetheless, the choice between the formal descriptions of language provided in the field of theoretical linguistics constitutes an additional auxiliary theory in itself. This approach has been adopted by some scholars in the field (e.g., Friederici, 2017; Matchin & Hickok, 2020; Pylkkänen, 2019; Zaccarella & Friederici, 2015), but a debate exists in the field about how well this approach can account for all aspects of language processing (Frank et al., 2012).

In the following, we will first discuss how a wide range of assumptions imported from linguistics, psychology, and neuroscience constrain the theoretical foundations of most of the studies conducted in the interdisciplinary field of cognitive neuroscience of language (section "Theoretical Assumptions"). We will then consider an equally important factor affecting the reliability of the derivation chain, which refers to the assumptions made when translating this interdisciplinary theoretical apparatus into an empirical investigation (section "Methodological Assumptions"). In particular, from the experimental design to the measurements and analysis, researchers incorporate a large number of assumptions in the derivational chain. Each of these decisions shape the contribution of a given study to the field of the cognitive neuroscience of language.

Theoretical Assumptions

The first step in any derivation chain is to identify the construct that one intends to investigate experimentally. While this may sound trivial to scholars from other fields of research, researchers working in the cognitive neuroscience of language have hardly ever agreed on how to strictly define "language" and associated theoretical concepts. For example, some researchers equate language and communication (e.g., Beckner et al., 2009; Scott-Phillips et al., 2009). Other researchers treat language as an abstract formal process working over a spoken, written, or signed signal (e.g., Chomsky, 2017; Friederici et al., 2017). Again others define language as the act of communicating in any form, including not just speech and sign but also communicative gesture (e.g., Arbib et al., 2008; Goldin-Meadow &

Brentari, 2017; Özyürek, 2014) For any particular definition or understanding of language, a certain interpretative context is therefore set and even within a particular definition of language, different aspects of language are of specific focus (see Box 1). Here, the choice of basic definitions directly influences what a researcher will consider as the key constructs of interest, which manipulations will be adopted in a given experimental design, and how the collected data will be analyzed.

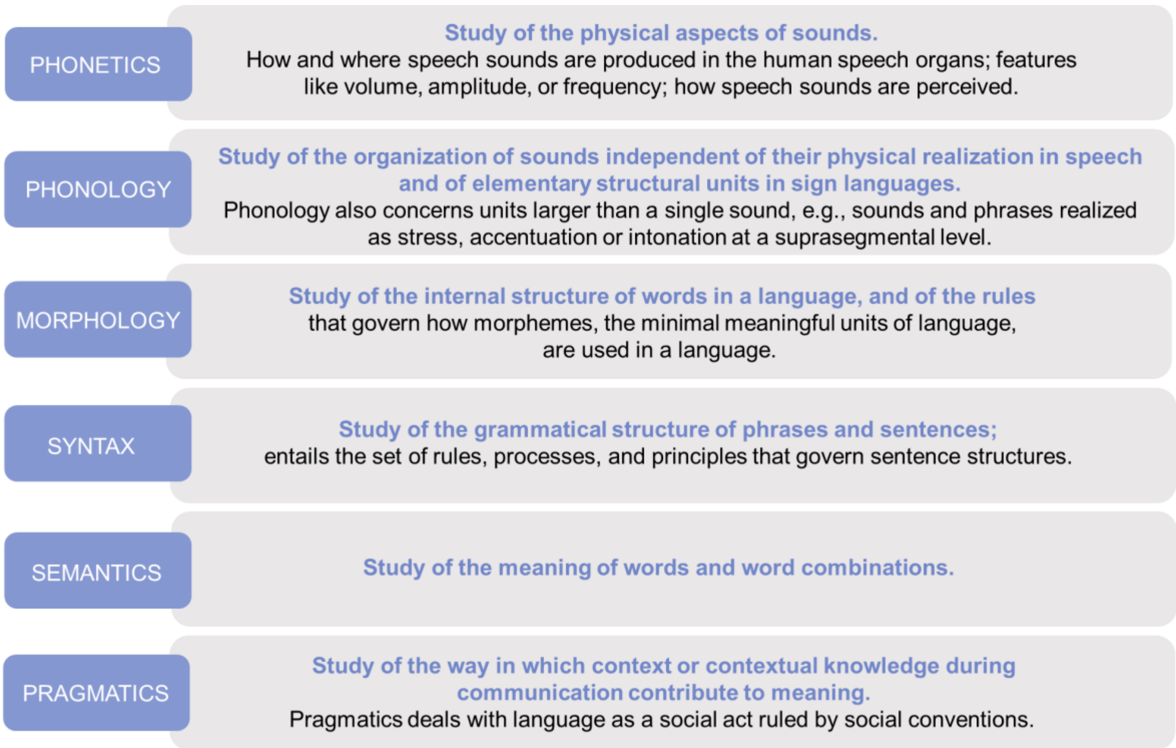
We propose that the lack of agreement on the strict definition of "language" is also partially rooted in the fact that the translation from linguistic theory into observational terms through empirical hypotheses is not always straightforward. It is by now a widely known and established fact that linguistic theories and processing theories do not directly map onto neurobiological processes (for detailed discussions, see Baggio, 2020; Embick & Poeppel, 2015; Martorell, 2018; Poeppel & Embick, 2013). Using Marr's classic definition of the tripartite view of levels of analysis (Marr, 2010), the link between the computational, algorithmic, and implementational levels in our field has not been clearly defined yet (Figure 1). Therefore, researchers have many degrees of freedom in choosing their favorite construct, theoretical framework, processing theory, and hypothesis about the neural implementation.

Here, we do not aim to establish a consensus about how to define "language" or what phenomena to focus on in future research. Instead, we propose that researchers should be explicit in describing the starting point of the derivation chain used in the construction of a particular edifice. Once this is done, the diversity of viewpoints may actually leave some room for integration in the long run. That is, we suggest that the different basic definitions and diverging research interests within the cognitive neuroscience of language might have led to the production of differently-shaped bricks and simultaneously to the construction of edifices that serve radically different purposes (i.e., a "chaotic brickyard"). However, these different bricks in the cognitive neuroscience of language do not necessarily need to contradict each other, and might become more clear and coherent once the underlying auxiliary assumptions are made explicit. For example, findings obtained in the context of studying language as communication may highlight aspects that are disregarded in an approach focusing on language as a computational system, and vice versa. Moreover, similar neural data may receive radically different interpretations by different researchers, depending on basic definitions and assumptions which they explicitly but more often also implicitly choose to adopt.

In the remainder of this article, we will mainly use the study of syntax in linguistics, psychology, and neuroscience as one of the different sub-components of language in focus within the field of the cognitive neuroscience of language (see Box 1). However, we wish to emphasize that the fundamentals we discuss for syntax can be and have been applied equally well to other sub-components such as, for example, semantics or phonology.

Box 1. Subcomponents of Language

No matter what definition of "language" an individual researcher adopts, linguists have tended to describe the language system using a number of subcomponents. At least six major subcomponents of language can be identified, although the boundaries between them tend to be blurry: Phonetics, phonology, morphology, syntax, semantics, and pragmatics. Each subcomponent deals with different linguistic information types, ranging from acoustic information of single speech sounds to complex grammatical relations:



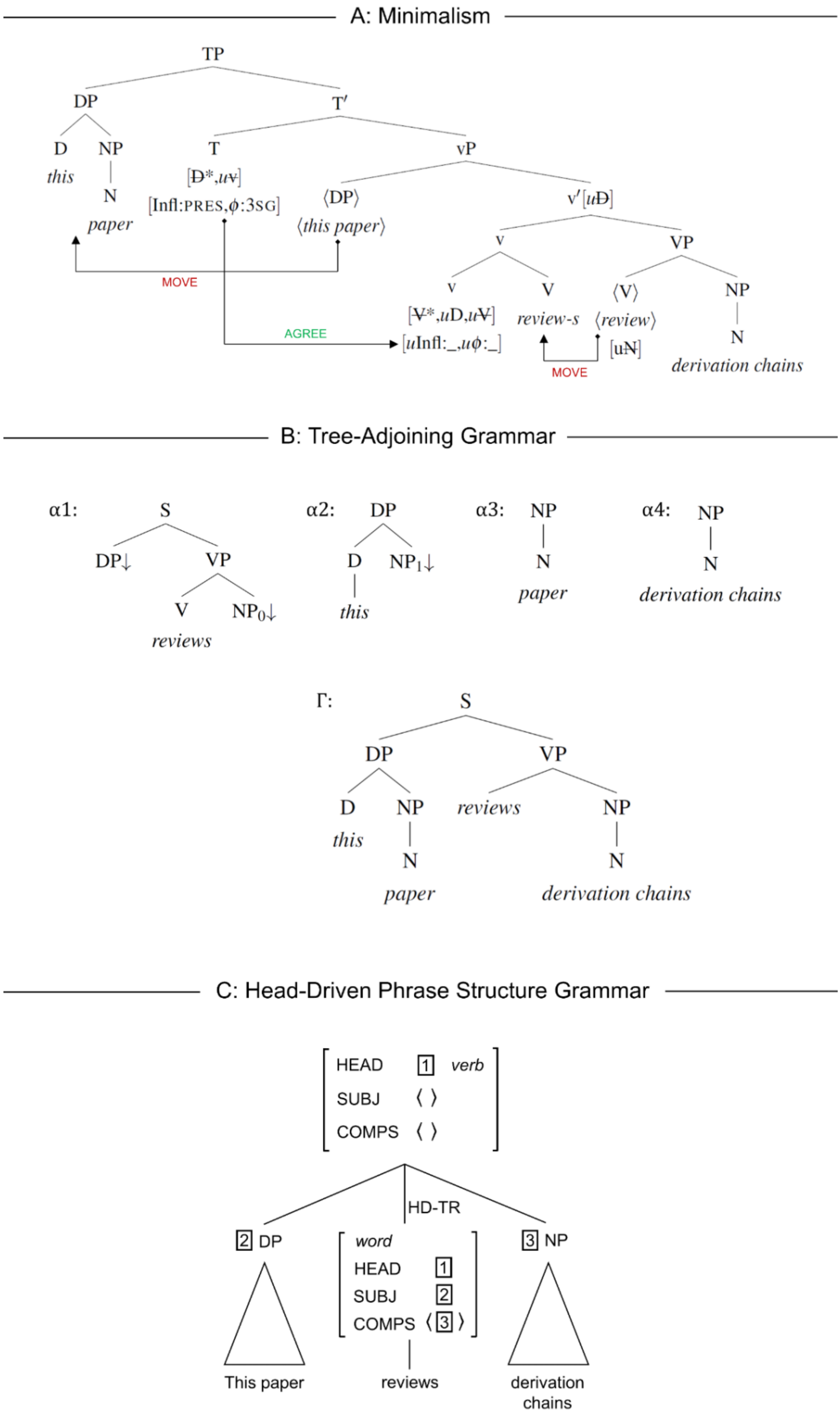
***A_{T1}*: Imports from Linguistics**

Linguistics examines language itself. Each individual researcher decides what elements of linguistics enter into the derivation chain of their experiment. In principle, researchers can choose to disregard modern linguistics altogether and perform neuroscientific experiments without any connection to linguistic theory. For example, a researcher may choose to present participants with written words randomly drawn from a thesaurus and contrast this with a control condition in which participants view natural scenes. There may be good reasons to conduct an experiment along these lines. However, the conclusions about language processing based on this experiment will be limited to the truism that in one condition the participants saw linguistic stimuli, whereas in the other condition they did not. Crucially, even when linguistic theory is disregarded (explicitly), the linguistic view that language is categorical in nature (i.e., there are no half-words) is automatically adopted (Chomsky, 1966/2002). That is, words or signs are treated as discrete units despite the fact that this is never reflected in the nature of the auditory, written, or signed signal. In fact, apart from prosodic phrasal boundaries, there is no acoustic break signaling the begin and the end of individual words in the speech stream, nor holds or pauses between signs in the visual stream of signs: In the case of prosodic phrase boundaries, listeners are known to perceive a boundary between syntactic phrases even without acoustic break (Steinhauer et al., 1999).

A case in point for the current discussion is the way in which different theories describe the grammar reflecting the idealized speakers' abstract knowledge of their language, without being concerned about how utterances are processed in real time (Chomsky, 1965; see *competence* vs. *performance* in "Imports from Psychology" below). As a matter of fact, the syntactic representation of a sentence can be described using many different frameworks, which in turn, come with a number of implicit and explicit auxiliary assumptions. According to different theory-specific proposals, a sentence can therefore be analyzed assuming either binary branching (e.g., Chomsky, 1988, 1995) or *n*-ary branching (Pollard & Sag, 1994), with rather different-looking structural descriptions (see Figures 2A and 2C). Furthermore, a certain linguistic theory may start from the assumption that well-formed sentences have a so-called Deep Structure satisfying some structural requirements and a Surface Structure that complies with positional relationships (as in the early versions of the Principles and Parameters theory within the generative grammar tradition; Chomsky, 1988); or alternatively that it should provide the complete derivation of a structure in its description (e.g., minimalism; Chomsky, 1995) at all times (see Figure 2A); or, conversely, that it should allow for the combination of "pre-assembled" structural elements already stored in memory (e.g., tree adjoining grammars; Joshi et al., 1975; see Figure 2B).

Figure 2

A: *A minimalist derivation of the sentence "This paper reviews derivation chains". The derivation proceeds in a bottom-up fashion, starting from the most deeply embedded element in the sentence. Here "u" indicates uninterpretable features that have to be satisfied and subsequently deleted. The feature [uN] expresses the requirement of the verb ("review") for a nominal argument ("derivation chains"). After the noun phrase (NP) ("derivation chains") was merged, [uN] is removed. The features [uV] and [uD] on the small v express the requirement for a verb phrase (VP) ("review derivation chains") and a determiner phrase (DP) ("this paper"). The VP is merged and the feature [uV] is deleted. The strong feature [V*] requires the movement of the verb "review". The unsatisfied feature [uD] is projected further until a DP "this paper" is merged. The feature [uv] on the functional head T is removed after the vP has been merged. The strong feature [D*] requires the movement of the DP "this paper". The T head assigns tense, person, and number features to the v via an agreement operation.* ***B:*** *A tree-adjoining grammar description of the sentence "This paper reviews derivation chains". Initial trees are labeled α1, α2, α3, and α4. The final derived tree is labeled Γ. Downward-pointing arrows indicate the substitution sites, which are substituted by the respective trees α2, α3, and α4.* ***C:*** *An HPSG tree of the sentence "This paper reviews derivation chains".*


A: Minimalism
TP
DP
T′
D
NP
this
N
paper
T
[Infl:PRES,ϕ:3SG]
vP
⟨DP⟩
⟨this paper⟩
v′[uD]
v
VP
v
V
review-s
⟨V⟩
⟨review⟩
[uN]
NP
N
derivation chains
[uInfl:_,uϕ:_]
MOVE
AGREE
MOVE
B: Tree-Adjoining Grammar
α1:
S
DP↓
VP
V
NP0↓
reviews
α2:
DP
D
NP1↓
this
α3:
NP
N
paper
α4:
NP
N
derivation chains
Γ:
S
DP
VP
D
NP
this
N
paper
reviews
NP
N
derivation chains
C: Head-Driven Phrase Structure Grammar
HEAD 1 verb
SUBJ ⟨ ⟩
COMPS ⟨ ⟩
HD-TR
2 DP
word
HEAD 1
SUBJ 2
COMPS ⟨3⟩
3 NP
This paper
reviews
derivation chains

The different analyses presented here (Figure 2) all accurately capture the structure of the respective sentences according to the theoretical assumptions of each of the corresponding linguistic frameworks. From the point of view of the cognitive neuroscience of language, this however raises two major questions: Firstly, is there any reason to adopt one approach of capturing sentence structure over the other? And secondly, what are the implications of adopting one or the other approach for the design of a cognitive neuroscience experiment?

There are no easy answers to these questions, because researchers may have different reasons to prefer one structural description over another. As long as the phenomenon they are attempting to investigate in their experiment does not critically depend on a particular aspect of a specific theory, the choice of theory and the concomitant theoretical imports (e.g., binary branching and complete derivation) may be arbitrary. However, if an experiment attempts to investigate, for example, the nature of the syntactic combinatorial operations in language, then the choice of a theoretical description directly affects the experimental design. In a minimalist analysis, it is assumed that the combinatorial operation is binary, combining only two elements at a time, whereas more than two elements can be combined at the same time under the assumption of *n*-ary branching. An experiment motivated by linguistic theory attempting to study syntactic combination would have to take this theoretical difference into account.

A_{T2}: Imports from Psychology

Psycholinguistics, at the interface between psychology and linguistics, examines the psychological processes involved in comprehending and producing language. Figure 1 illustrates that linguistic theory does not straightforwardly map to psychological processes. For instance, while a listener (or signer) during online language comprehension encounters the left element first ("*cats*" in the example sentence "*cats chase mice*"), the syntactic analysis of the sentence according to task-independent, atemporal linguistic theories usually starts from the last element in the sequence ("*mice*"). Here, linguistic theory primarily addresses language *competence*, that is the internal linguistic knowledge of the ideal speaker/signer-hearer/viewer (Chomsky, 1965), whereas psycholinguistics deals with language *performance*, the real speaker/signers's actual language use (Harley, 2013).

An intuitive way to understand the difference between competence and performance is that real-life comprehension and production are constrained by cognitive resources and contextual information. These constraining factors are particularly evident in center-embedded structures such as in the example "*The scientist that a journalist that a policeman knows cites is very famous.",* where it becomes very "expensive" (in terms of working memory resources) to link each subject to its corresponding verb without hesitation or rereading. From the theoretical linguistics point of view however, speakers and signers can, in principle, use an infinite number of embeddings (Miller & Chomsky, 1963)—that is, there is no rule in language that prevents the embedding of one sentence into another. Psycholinguistically, however, comprehenders get constrained by factors such as limited working memory performance, as it becomes increasingly difficult to comprehend structures that go beyond two (or three) embeddings (Blaubergs & Braine, 1974; Cecchetto et al., 2016; R. L. Lewis, 1993, 1996).

As a matter of fact, studies in the cognitive neuroscience of language do not only import theoretical linguistic assumptions discussed earlier in their derivation chain. Rather, the derivation chain also adopts a number of performance-related assumptions originating from general psychology: For example, working memory (e.g., Fedorenko et al., 2006; Swets et al., 2007), executive functions (e.g., Baddeley et al., 2009; Shao et al., 2014), predictive processing (e.g., van der Burght et al., 2021; Weber et al., 2006), or automaticity of higher mental processes (Pyatigorskaya et al., 2022). Distinct linguistic theories differ in the strength of their link to these psychological theories. For example, the assumption that linguistic structures or constructions can be stored in the mental lexicon is implemented in the Tree-Adjoining-Grammar formalism (Figure 2B) as well as Construction Grammar (Goldberg, 1995). In contrast, versions of the Minimalist Program (Figure 2A) deny the mental storage of linguistic constructions, which are instead thought to be fully derived. While a more or less direct link from a description of linguistic competence to actual performance factors does not make a certain linguistic theory more "psychologically real" than another, an explicit link to psychologically measurable constructs can be beneficial when investigating certain constructs experimentally.

Performance factors are important to take into account in some situations, but the study of *competence* alone can sometimes help to interpret the available data in a certain way that allows for a distinct level of understanding: For example, the study of language use alone may miss the suitable description of language structures as hierarchical organization of constituents (Frank et al., 2012), whereas this insight arises relatively early when looking only at structural descriptions of linguistic *competence*. Similarly, the degree to which some theoretical linguistic assumptions (e.g., constituency or the full derivation under a Minimalist approach) can (and should) be translated to actual *performance* remains an open question and subject to debate. With the aim of narrowing down the gap between *competence* and *performance*, processing-friendly models of language competence have begun to show that some behavioral facts on language comprehension in real speakers can be better explained if fine-grained theoretical assumptions and specific directionality of derivation are taken into account (e.g., the type of lexical restrictions of the constituents involved in the structure; Chesi & Moro, 2015).

In general, the choice to begin conceptualizing an experiment from the algorithmic level of psychological theories of language processing (Figure 1) or from the abstract analysis of language *competence* may lead not only to a focus on different phenomena, but also to radically different interpretations of the data—unless a more comprehensive comparison of models accounting for different *performance* factors is taken into account (Chesi & Canal, 2019).

A_{T3}: Imports from Neuroscience

While the disconnection between linguistics and psychology is grounded mostly in the distinction between *competence* and *performance*, there does not seem to be such an obvious disconnection between psychology and neuroscience at first. Both fields aim to study how language is processed, but from two different perspectives that may offer different explanations of the data observed. In fact, due to the general disconnection between psychology and neuroscience (e.g., Beste, 2021), the explanations offered by either cannot be straightforwardly translated. For example, it is currently not clear how different clusters observed in a functional magnetic resonance imaging (fMRI) study could be related to a difference in reaction times, or vice versa. There is of course much more to say about the different outlooks and the actual disconnects of psychology and neuroscience in general. For example, while a psychological theory relates to how a particular kind of information (in our case: linguistic information) is being processed, it is currently still unclear how any kind of information is actually represented neurally (Gallistel, 2017; Krakauer et al., 2017; Poeppel & Idsardi, 2022). Similarly, any psychological theory of (language) processing relies on a notion of memory, but it is currently also still unknown how neurons carry forward information in time (Gallistel & King, 2009; James Langille & Randy Gallistel, 2020; Trettenbrein, 2016).

When designing an experiment on language using neuroscientific methods, researchers face a three-way disconnect in their derivation chain (see Figure 1). At present, no general linking theories exist that bridge the gaps between linguistics, psychology, and neuroscience exist, although, recently models have been proposed (e.g., Martin, 2020). As a result, researchers have to deal with substantial degrees of freedom and a myriad of auxiliary assumptions coming from all three fields and levels of analysis. We will illustrate this disconnect with the following example: Given that one of the defining properties of language is its discrete infinity (i.e., the fact that we can effortlessly comprehend and produce sentences that we have never encountered before in our life), different researchers in cognitive neuroscience of language have singled out the generative capacity of language that assembles constituents as their object of investigation. Some works have investigated how word combination is reflected neurally using fMRI, moving from a *competence* perspective (Zaccarella et al., 2015, 2017; Zaccarella & Friederici, 2015a). Conversely, other works have similarly studied how elements are combined in language, yet from the *performance* perspective of psychology (Hagoort, 2013, 2019; Hagoort & Indefrey, 2014).

On a certain level of abstraction, both groups of studies have attempted to understand how linguistic elements are cortically combined. Yet, the interpretations of the neural data, which has shown significant overlap in language-relevant brain regions, have at times diverged quite drastically due to the different derivation chains employed. While the first group of studies moved from theoretical linguistics to capture the neural correlates of an empirical hypothesis from linguistic theory, standardly phrased as Merge operation, Hagoort and colleagues motivated their experiments against a processing theory in which linguistic elements are combined by a Unification operation. Crucially, these theoretical choices directly influenced the respective interpretation of the data by both teams, despite an overlap in the identified brain regions. When viewed within the derivation chain framework, this discrepancy is actually not surprising. The different explicit and implicit theoretical imports from linguistics and psychology by both teams of researchers led them to investigate an abstractly similar phenomenon (i.e., linguistic combinatoriality) in two different ways, despite the overlap regarding the chosen method (i.e., fMRI).

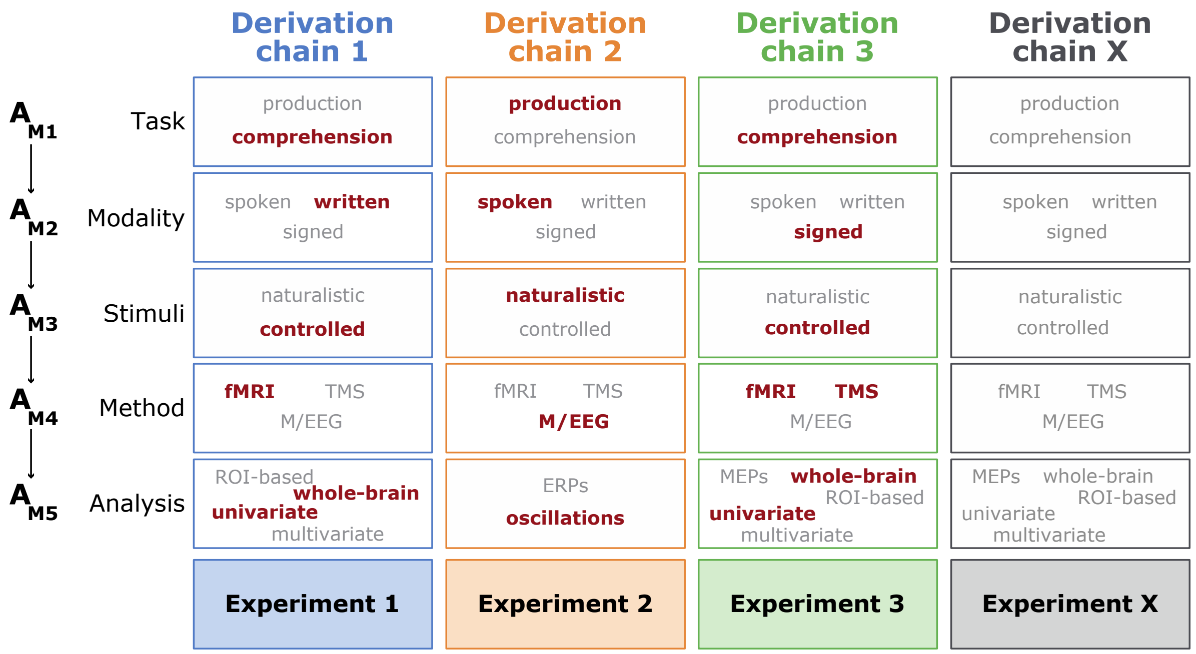
A pressing question that arises against this background is whether and how neural data could be informative in distinguishing between the neural correlates of two distinct operations: Merge and Unification. Both may abstractly appear to fulfill a similar purpose, the combination of two elements. Yet, the theoretical description of Merge in linguistics is more constrained and clearly limited to syntax. In contrast, the concept of Unification from a processing perspective also applies to semantics, phonology, and, more generally, any notion of combining elements retrieved from memory. We suggest that the neural data, at this point in time, remain uniformative with regard to which theoretical definition should be preferred over another. Instead, the a priori selection of the respective auxiliary assumption from either linguistics or psychology seems to define how the neural data are interpreted.

Methodological assumptions

In addition to the auxiliary assumptions on a theoretical level, there are also auxiliary assumptions defined purely by methodological choices of the experimenter (Figure 3). While some of these assumptions also affect other fields of cognitive neuroscience, some of them are specific to the cognitive neuroscience of language.

Figure 3

Examples of methodology-derived auxiliary assumptions. The same hypothesis can involve different derivation chains at the methodological level. The different chains are built up from the chosen (highlighted in red) task, modality, stimuli, method, and analysis, among all the possible ones that can be employed (shaded). Each methodological choice comes with its own auxiliary assumptions, constraining the scope of the research question under investigation. AM: methodology-derived auxiliary assumptions; fMRI: functional magnetic resonance imaging; M/EEG: magneto/electroencephalography; MEPs: muscle evoked potentials; ROI: region of interest; TMS: transcranial magnetic stimulation.



Language and stimuli

A_{M1}: The Language (Modality) Tested is one Where the Relevant Linguistic Phenomenon of Interest can be Tested. An important characteristic that sets language apart from many other cognitive domains is the diversity in which language can be externalized (and perceived). Firstly, there are currently approximately 7,000 languages being used around the world (UNESCO, 2021). Each of these languages could be used as an experimental tool to study the cognitive neuroscience of language. Secondly, many of those languages can, in turn, be operationalised in a number of modalities: vocal production and comprehension, or reading and writing. Moreover, the sign languages of the world rely on the visuo-spatial modality for production and are perceived visually (Klima et al., 1979; Pfau et al., 2012), whereas the tacticle languages of the deaf-blind are perceived using touch (Checchetto et al., 2018; Edwards & Brentari, 2020). The methodological choices of the language and its modality come with their own hidden auxiliary assumptions, which, when not considered and made explicit, can weaken the derivation chain.

The discussion on whether language universals exist and to what extent a linguistic process or construct is a property of a particular language (Baker, 2009; Evans & Levinson, 2009; Greenberg, 2005; Longobardi & Roberts, 2010) is beyond the scope of this paper. However, any study on language processing depends on the auxiliary assumptions that the language under investigation is a good representation of language use across languages. For instance, if a study reports on the neural correlates of syntactic processing without specifying the language used, the assumption is made that syntactic processing in language under investigation is representative of syntactic processing as such. Depending on the process, this auxiliary assumption may not be valid given the large cross-linguistic variability in syntactic structure. Indeed, various instances of cross-linguistic differences on the neural level have been found (e.g., Bornkessel-Schlesewsky et al., 2011; Goucha et al., 2022). Ignoring this assumption might have amplified apparently conflicting findings in the literature. For example, investigations of syntactic processing in English have highlighted the role of the posterior temporal lobe (Law & Pylkkänen, 2021; Matar et al., 2021; Matchin et al., 2017, 2019), while studies conducted in German have pointed towards a central role of the left inferior frontal gyrus (Goucha & Friederici, 2015; van der Burght et al., 2019; Zaccarella & Friederici, 2015a). Finally, there seems to be a large difference between language studies on so-called Western, Educated, Industrialized, Rich, and Democratic (WEIRD) populations (Henrich et al., 2010), who often report findings on “language processing” in general while only explicitly mentioning the specific language studied in the methods section. Studies on non-WEIRD populations generally mention the language investigated in a prominent way, e.g., in the title or abstract (e.g., Matar et al., 2021; Ohta et al., 2017; Wu et al., 2019).

Similarly important is to state explicitly which language modality was chosen, and to what extent conclusions on language processing may be shaped by this methodological decision. This debate concerns the question which (neuro-)cognitive processes are shared between comprehending, producing, reading, and signing language, and which processes might be unique to each of these modalities (McQueen & Meyer, 2019). In other words, to what extent can processing mechanisms in a given domain be generalized to the language system as a whole, or are they limited to linguistic processing in the studied domain only. As such, a generic label such as "language processing" almost always represents an underspecification, unless the experimental task used has been established to capture a processing mechanism shared across modalities (e.g., Arana et al., 2020; Gastaldon et al., 2020; Giglio et al., 2021; Heim et al., 2010; Matchin & Wood, 2020). Indeed, studies on auditory language comprehension (Vigneau et al., 2011), sign language comprehension (Trettenbrein, Papitto, et al., 2021), and reading (Arana et al., 2020) have shown large overlap in language-related regions across the cortex. It is therefore crucial for studies in the cognitive neuroscience of language to state explicitly which language was studied and in which modality, and to discuss how these choices may influence any hidden assumptions related to the theory under investigation. Particularly, whether a certain construct or process is unique or shared to the language and modality of choice should be clearly demarcated.

A_{M2}: The Materials Employed are Capable of Testing the Hypothesis Without Introducing Confounds. The experiments conducted in the field of cognitive neuroscience of language usually consist of the presentation of isolated words, phrases or sentences that vary along specific linguistic dimensions of interest. As in other cognitive neuroscientific domains, the materials employed need to be able to test the hypothesis of interest without introducing additional confounds. At present, there are several tools available that allow researchers to control for the perceptual aspects of materials employed in language research (e.g., Boersma, 2001; Trettenbrein & Zaccarella, 2021). However, additional dimensions of interest relate to the psycholinguistic properties of the words employed. For example, a researcher might want to ensure that the difference between two conditions (e.g., abstract and concrete nouns) is not confounded by differences in lexical frequency, or by phonological or orthographic neighborhood (i.e., how many words are pronounced or written very similarly to a given target; Marian, 2017; Marian et al., 2012). This issue can be easily addressed by consulting linguistic corpora when the materials are created. In short, a corpus is a database which allows to easily extract psycholinguistic variables of interests (e.g., length, frequency of occurrence) for the entries (e.g., words or signs) of a specific language, extracted from large datasets such as articles in newspapers, movie subtitles, etc. Accordingly, the use of these corpora before running an experiment ensures that the stimuli (of the distinct categories/variables/factors) are matched along the relevant linguistic dimensions (for a methodological discussion, see also Sassenhagen & Alday, 2016). In cases where corpus data is not available for a certain language, such as, for example, for many sign languages, subjective ratings can be used to establish psycholinguistic paramters (e.g., Caselli et al., 2017; Trettenbrein, Pendzich, et al., 2021).

An additional aspect of stimuli preparation in language studies relevant especially for electroencephalography (EEG) and magnetoencephalography (MEG) concerns pre-target words. In particular, it is common practice to present a target word of interest (underlined in the next example) as part of sentences or phrases that differ along specific linguistic dimensions of interest (e.g., "the driver wears a <u>shirt</u>" vs. "the driver wears a <u>phone</u>" when considering semantic plausibility). These pre-target words need to be matched across relevant perceptual and psycholinguistic variables of interest, or counter-balanced across conditions (for examples, see (Hasting & Kotz, 2008; Maran, Numssen, et al., 2022). If the pre-target words elicit sustained differences across conditions, common E/MEG pre-processing procedures such as baseline correction might artificially create an effect at the target word (for a detailed discussion see Steinhauer & Drury, 2012).

If naturalistic stimuli are employed, these potential confounding variables can (and should) be included as regressors in the statistical model (Hamilton & Huth, 2020). An implicit assumption in psycho- and neurolinguistics research is that the information provided by the corpora is representative of the language as used by the participants in a study. Yet, this assumption can be violated. For example, if the corpus is based on relatively old texts (Brysbaert et al., 2011) that include archaic or disused words. This problem has led to the development of corpora based on more recent movie subtitles, because they seem to more closely match how language is used in everyday life (Boada et al., 2020; Brysbaert et al., 2011; Cai & Brysbaert, 2010; Cuetos et al., 2011; Dimitropoulou et al., 2010; New et al., 2007; Soares et al., 2015; van Heuven et al., 2014). Importantly, the development of accurate corpora for languages (e.g., as for some signed languages, indigenous languages, etc.) that still lack these extensive resources is an important goal for the language sciences at large in the years to come.

In an attempt to overcome some of these issues, studies making use of so-called naturalistic stimuli have recently gained popularity (for reviews, see (Alday, 2019; Hamilton & Huth, 2020; Willems, 2015). Rather than presenting carefully controlled stimuli as part of artificial laboratory tasks, these naturalistic approaches make use of sentences extracted from spoken corpora, short narratives, or audiobooks (Stehwien et al., 2020). Naturalistic stimuli have been successfully used to study phonetic feature encoding (Mesgarani et al., 2014), syntactic representations (Bhattasali et al., 2019; Brennan et al., 2016; Hale et al., 2018), linguistic predictions (Brennan & Hale, 2019; Heilbron et al., 2022; Shain et al., 2020), and semantic processing (Brodbeck et al., 2018; Broderick et al., 2018; Huth et al., 2016). Because of the increased ecological validity of the naturalistic stimuli, the assumption about the method’s ability to capture the process of interest (i.e., naturalistic language processing) should be more readily met. Yet, while the stimulus material may be unconstrained and more natural, these naturalistic approaches pose numerous analysis-based choices that are, in turn, based on their own auxiliary assumptions. More specifically, traditional statistical techniques are often unable to account for the numerous confounding variables present in natural speech (Hamilton & Huth, 2020). Instead, testing for a variable of interest requires complex computational tools with parameters defined by the experimenter. This effectively leads to a shift in auxiliary assumptions related to the experimental design to the analysis phase of the experiment. We will turn to these in a later section.

Data acquisition

A_{M3}: The Neuroscientific Research Technique/Method is Capable of Providing Data Which can be Used to Test the Hypothesis of Interest. Researchers in the field of cognitive neuroscience of language can rely on a large number of neuroimaging, neuro-stimulation and neuro-modulation techniques. For instance, neuroimaging techniques such as fMRI, EEG, and MEG provide *correlational* evidence on the link between brain functioning and cognition, whereas brain stimulation techniques (Bergmann & Hartwigsen, 2021; Hartwigsen & Silvanto, 2022) and lesion studies (Matchin et al., 2022; Vaidya et al., 2019) allow to draw *causal* inferences . Additionally, distinct research techniques show a difference in temporal precision, susceptibility to artifacts (e.g., Abbasi et al., 2021; Luck, 2005; Ouyang et al., 2016) and to cancellation of signals from a particular brain area (e.g., Devlin et al., 2000; Gorno-Tempini et al., 2002; Jezzard & Clare, 1999; Ojemann et al., 1997) and sensitivity to the orientation of the electrical currents (e.g., Ahlfors et al., 2009, 2010; Cohen & Cuffin, 1991). Consequently, each neuroscientific technique poses its own methodological constraints regarding the conclusions that can be drawn based on its data. These distinct methodological constraints might be the reason why, in language research, conclusions drawn based on EEG, MEG, and fMRI data might not always converge (Lau et al., 2013; Wang et al., 2021). Some of these methodological constraints apply to cognitive neuroscience research in general—here we focus on the specific methodological issues that require extra consideration in language research in particular.

Considering that language processing is a complex process that both includes bottom-up and top-down processing (Friederici, 2012; Hagoort, 2019) it is important to note that distinct neuroimaging techniques vary in temporal precision (He & Liu, 2008) and therefore might capture distinct processing stages. For instance, EEG and MEG provide a direct measure of brain activity (Lopes da Silva, 2013) that best captures transient, feedforward neural activity (Kochari et al., 2021; Vartiainen et al., 2011; Wang et al., 2021). In contrast, the sluggish temporal resolution of fMRI is largely insensitive to such transient, feedforward neural activity (Arthurs & Boniface, 2002; Bunge & Kahn, 2009; Furey et al., 2006; Segaert et al., 2013; Vartiainen et al., 2011). Consequently, fMRI might be more suitable to capture downstream, late-stage effects that are sustained or variable in time. Consideration of convergence and divergence across imaging modalities can yield new insights into the contributions of multiple components of the language network. Another important aspect which should be considered, especially for studies on language comprehension, is the fact that the distinct experimental environments in which neuroimaging studies (e.g., noisy MRI environment) take place affects the patterns of neural activation (Andoh et al., 2017; Pellegrino et al., 2022).

Given that neuroimaging methods such as MEG, EEG, and fMRI cannot provide causal evidence for the functional relevance of particular brain regions, an issue that can be overcome using transcranial magnetic stimulation (TMS). In brief, TMS excites neurons noninvasively below the stimulation coil (Polanía et al., 2018). When combined with concurrent EEG measurements (TMS-EEG), TMS can be used to study the causal involvement of different cortical areas in language processing as proposed by neurobiological models of language (e.g., Friederici, 2011, 2012; Hagoort, 2019) with an extremely high temporal resolution. However, there are also several disadvantages to the use of TMS in language research: One obstacle is the uncomfortable stimulation of surface muscle tissue when TMS is applied over frontolateral and temporoparietal language areas. Such discomfort needs to be taken into account when thinking of appropriate control (sham) conditions (e.g., see Duecker & Sack, 2015). When combined with simultaneous EEG measurements, the stimulation of this muscle tissue leads to the generation of long-lasting and distinctly large muscle artifacts that make it hard to interpret the underlying EEG signal (Salo et al., 2020). The development of new real-time visualization tools (Casarotto et al., 2022) can be used in an attempt to minimize muscle twitches before the start of the measurements. This approach is not only important when interested in early components of the EEG response to TMS, but also because these muscle twitches are clearly perceived by the participants and therefore result in unspecific brain responses (Conde et al., 2019). Furthermore, the loud clicking sound also induces an electrophysiological response, which cannot be easily masked (Russo et al., 2022) when interested in auditory language processing. These sensory problems need to be carefully taken into account when designing a TMS-EEG experiment probing language. For instance, one could think of creating good (linguistic) control conditions in terms of stimulus materials, in which the TMS-related conditions (i.e., muscle twitches, clicking sound) are equal (i.e., interaction effects).

It should also be noticed that, when TMS is employed, a large number of parameters need to be set by the experimenter (e.g., the choice of an online or offline stimulation protocol), which might influence whether a neurostimulation study will provide causal evidence or not (Qu et al., 2022), ultimately affecting the derivation chain strength. Another phenomenon when using TMS to study language processing are compensatory effects within the large-scale, distributed language network (Hartwigsen, 2018). For example, during auditory language processing, the listener needs to rapidly analyze the sound, meaning, and structure of spoken words in order to associate the heard sound patterns with meaningful concepts. These complex processes require the interaction of numerous brain regions. Given the large-scale nature of the language network, there is great potential for adaptive plasticity in order to compensate for a focal perturbation of a key region when using TMS (Hartwigsen, 2018). In line with this, evidence shows that unifocal perturbation with (repetitive) TMS is often not sufficient to perturb various language-related processes (see, e.g., Kroczek et al., 2019; Maran, Friederici, et al., 2022), whereas the use of combined perturbation of two brain regions leads to an observable effect (e.g., Hartwigsen et al., 2016; Schroën et al., 2020). Such adaptive plasticity, however, makes it difficult to study the causal dynamics underlying language processing using TMS. The use of condition-and-perturb approaches (Hartwigsen et al., 2012) can increase the perturbation load on the language network, and might be necessary to distinguish between the lack of a causal involvement of a brain region in a process and the initiation of compensatory mechanisms. Consequently, the attribution of a cognitive function to any cortical region may be skewed. The conclusions on structure-function relationships from lesion studies are similarly complex, since they, too, rest on the (implicit) auxiliary assumption that no plastic changes have taken place.

In sum, when using TMS, as with other techniques used in the field, the various auxiliary assumptions involved introduce their own constraints on the inferences that can be drawn. In the case of language research in particular, however, special care needs to be taken because important regions in the language network are posed with imaging modality-specific limitations: TMS over the inferior frontal and frontal temporal lobes is particularly uncomfortable due to the proximity to facial muscle tissues, and as mentioned before, anterior temporal regions are difficult to capture using fMRI due to signal distortions (Devlin et al., 2000; Jezzard & Clare, 1999). Or, as Meehl might have put it, the implicit assumption that a given neuroimaging or neurostimulation technique is equally suitable for different parts of the cortex (e.g., AM: the signal of a given neuroscientific method is uniform throughout the cortex) might lead to weakening of the derivation chain in various language research agendas.

Data analysis

A_{M4}: The Analytic Approach is Capable of Testing the Hypothesis of Interest. Data from EEG and MEG language studies are traditionally analyzed focusing on event-related potentials (ERPs; Luck, 2005) or neural oscillations (Buzsáki & Draguhn, 2004). On the one hand, several ERP components (e.g., ELAN, LAN, N400, P600) have been linked to specific stages of linguistic processing (e.g., phrasal building, morphosyntactic analysis, semantic composition, integration; see (Friederici, 2011; Hernández et al., 2022; Maran, Friederici, et al., 2022; Zaccarella & Friederici, 2015b). On the other, distinct neural oscillations seem to subserve the multiscale property of language, from the processing of phonetic and syllabic units to phrases, and ultimately to more complex aspects of comprehension (e.g., Benítez-Burraco & Murphy, 2019; Giraud & Poeppel, 2012; A. G. Lewis et al., 2015).

Focusing the analysis on either ERPs or neural oscillations reflects an implicit assumption made by the researcher on the nature of the hypothesized effect. ERPs are the result of averaging across trials, and accordingly capture only activity that is both time-locked (i.e., in a constant time-relationship) and phase-locked (i.e., in a constant phase-relationship) to a stimulus or process of interest (Luck, 2005). Accordingly, if a linguistic effect has a variable time-course over trials, or manifests itself via non-phase locked activity, it might not be adequately captured by ERP analyses (Kochari et al., 2021; Maran, Friederici, et al., 2022). Of note, the low temporal resolution of fMRI makes this method less affected by this potential issue (Kochari et al., 2021). Thus, when an ERP-based analysis is employed, researchers are implicitly assuming that their effect of interest manifests itself via time- and phase-locked neural activity. This is not a trivial assumption given that, for example, some jittering across trials might be present when acoustic stimuli are employed.

In contrast to ERPs, analyses based on neural oscillations are better suited to account for the multi-dimensionality of M/EEG data (M. X. Cohen, 2011, 2014), because this type of analysis can capture the power and phase relationships of different frequency bands. This additional information can be used to compute several measures of functional and effective connectivity (e.g., coherence, phase-locking value) at the sensor- or source-level (Bastos & Schoffelen, 2016; Maran et al., 2016; Schoffelen & Gross, 2009). Given the large number of measures that can be extracted via time-frequency analysis, researchers often need to make implicit assumptions on the effect of interest when analyzing their data. The first choice is the frequency (or frequencies) of interest. This decision might not be straightforward when naturalistic stimuli are employed, given the intrinsic variability in time of linguistic information (e.g., syllables; Alday, 2019). Cluster-based permutation tests (Maris & Oostenveld, 2007) have been employed in some studies (e.g., Segaert et al., 2018) in order to address effects that might extend beyond the canonical frequency bands. Critically, this method bears some limitations on the specificity (e.g., in time, space or in frequency bin) of a significant effect (Sassenhagen & Draschkow, 2019), whose definition might require a follow-up investigation tailored to investigate its exact extent. Additional assumptions are made when choosing either to focus on changes in power or in phase of neural oscillations. These additional assumptions should be kept in consideration, because these measures might be differently sensitive to a particular experimental manipulation (Ding & Simon, 2013; Luo & Poeppel, 2007). Overall, these choices should be motivated by the extensively available literature on the ERP (Friederici, 2011; Hernández et al., 2022; Maran, Friederici, et al., 2022) and oscillatory (Benítez-Burraco & Murphy, 2019; Giraud & Poeppel, 2012; A. G. Lewis et al., 2015; Meyer, 2018; Murphy, 2015) correlates of language processing. Furthermore, one should take into account the considerations on how basic neurophysiological mechanisms might subserve linguistic operations (Friederici & Singer, 2015; Fries, 2015; Murphy, 2015).

Aside from electrophysiological methods, analyses of neuroimaging data involve various researcher degrees of freedom that might shape their outcome. In fMRI, researchers can choose between a whole-brain or a region of interest (ROI) based approach. The ROI-based approach is usually used to maximize statistical power. However, this approach is, by definition, more constraining and strongly depends on the validity of the assumption that the process of interest can be (uniquely) attributed to the region of interest. Similarly, novel analysis techniques with respect to the use of TMS have recently used modeling of the TMS-induced electric field in each individual subject. The magnitude of the electric field in a certain ROI can, in turn, be used to explain modulations in behavior in a number of language tasks (Kuhnke et al., 2020; Maran, Numssen, et al., 2022; van der Burght et al., 2022). In this case, the auxiliary assumptions again imply that the process of interest can be localized to one or multiple ROIs, and that stimulation and task performance are linearly related.

A special case of ROI-based analyses in fMRI research is a so-called localiser approach (e.g., Fedorenko et al., 2010). Here, the individual brain is masked using a functional contrast image involving a language processing and a baseline task, e.g., complex sentences vs. strings of unconnected pseudowords. While in some ways regarded as a more powerful approach than whole-brain analyses or group-based ROI-based analyses, the functional localiser approach necessitates researchers to justify the conditions that the localiser contrast is based on, and to carefully explain how the functional localiser constrains the subsequent results.

Finally, a recent development has seen the analysis of fMRI and M/EEG data combined with machine learning techniques, often using naturalistic stimuli. As an alternative to manually annotating the speech stimulus (e.g., Kaufeld et al., 2020), these analyses make use of a computational model that is trained on the stimulus input in order to predict the resulting brain responses. In the subsequent analysis, a comparison is made between the predicted and measured brain responses (Goldstein et al., 2022; Heilbron et al., 2022). Importantly, the results of such neural analyses will strongly depend on the model assumptions. For example, GPT-2 (Radford et al., 2019), one of the most successful models, is designed to form a next-word prediction based upon a sequence of preceding words. Clearly, this type of model relies on the assumption that the defining property of human language processing is to continuously predict the upcoming word. However, this assumption may not be met in every context of natural language use (Ferreira & Qiu, 2021; Huettig, 2015). Here, the auxiliary assumptions involved in computational approaches may be relatively less transparent than those involved in a traditional factorial design, especially for readers with limited computational background. It is therefore particularly important that in such studies the analytical assumptions are carefully justified and that any potential constraints on the results are discussed.

Concluding Remarks

The cognitive neuroscience of language is a wide field with a great diversity of experimental findings and approaches. We have argued that this diversity mostly results from the derivation chains being substantially long, complex, and mostly implicit in our field. Consequently, researchers encounter many degrees of freedom in forming their derivation chain, shaped by the numerous theoretical and methodological choices they make between hypothesis and experimental result. Researchers may choose to start their derivation chain on the computational level of the abstract description of linguistic competence. They may instead choose to start from performance factors relevant on the psychological level of language processing. Finally, they may disregard either for the most part and attempt to cut their derivation chain short by starting from neurobiology. While all these choices can be well-motivated and justified, we have argued here that most derivation chains in our field almost always import hidden assumptions from all three fields: linguistics, psychology, and neuroscience.

We have no illusion that the field should suddenly reach a consensus on key terminology or approaches. Instead, in agreement with Scheel et al. (2021), we argue that a highly interdisciplinary field such as the cognitive neuroscience of language requires researchers to form explicit statements concerning their experiment's derivation chain. The auxiliary theoretical assumptions as well as the constraining factors of the methodology employed in their work should be clearly motivated and discussed. Researchers should be aware of the numerous theoretical auxiliary assumptions attached to a theory when making inferences regarding the neurobiology of language. The same applies to methodological auxiliary assumptions, which constrain the scope of a given experiment, limiting the research questions that can or cannot be addressed. If these (often implicit) theoretical and methodological auxiliary assumptions are not considered, an accurate appraisal of data in the field of cognitive neuroscience of language is compromised.

Lastly, we would like to point out that the prevalence of different definitions and lack of a clear consensus in cognitive neuroscience need not be considered problematic. Instead, we suggest that the observed diversity of viewpoints should be considered complementary. To return to the brickyard metaphor of the *Science* letter quoted in the introduction, we acknowledge that it is not always clear from the outset how the bricks made by different brickmakers could eventually be reassembled into one large edifice. However, we suggest that for the field to move towards the long-term goal of integration, an important first step is to help other researchers to identify how each brick was made. That is, researchers should carefully describe the complete derivation chain involved in their study. A more transparent discussion of the implicit and explicit auxiliary assumptions behind our experiments may therefore significantly improve research into the neurobiology of language.

Acknowledgements

The foundations for this paper were laid at the Leipzig Lectures on Language, a lecture series and symposium held online in 2021 (https://www.cbs.mpg.de/leipzig-lectures-on-language). We are grateful to our lecture series co-organiser Caroline Beese, as well as all speakers and attendants for the inspiring talks and discussions. Also, we wish to thank Antje S. Meyer for comments on this manuscript. This work was funded by the Max Planck Society.

References

Abbasi, O., Steingräber, N., & Gross, J. (2021). Correcting MEG Artifacts Caused by Overt Speech. *Frontiers in Neuroscience*, *15*, 682419. https://doi.org/10.3389/fnins.2021.682419

Ahlfors, S. P., Han, J., Belliveau, J. W., & Hämäläinen, M. S. (2010). Sensitivity of MEG and EEG to Source Orientation. *Brain Topography*, *23*(3), 227–232. https://doi.org/10.1007/s10548-010-0154-x

Ahlfors, S. P., Han, J., Lin, F.-H., Witzel, T., Belliveau, J. W., Hämäläinen, M. S., & Halgren, E. (2009). Cancellation of EEG and MEG signals generated by extended and distributed sources. *Human Brain Mapping*, NA-NA. https://doi.org/10.1002/hbm.20851

Alday, P. M. (2019). M/EEG analysis of naturalistic stories: A review from speech to language processing. *Language, Cognition and Neuroscience*, *34*(4), 457–473. https://doi.org/10.1080/23273798.2018.1546882

Andoh, J., Ferreira, M., Leppert, I. R., Matsushita, R., Pike, B., & Zatorre, R. J. (2017). How restful is it with all that noise? Comparison of Interleaved silent steady state (ISSS) and conventional imaging in resting-state fMRI. *NeuroImage*, *147*, 726–735. https://doi.org/10.1016/j.neuroimage.2016.11.065

Arana, S., Marquand, A., Hultén, A., Hagoort, P., & Schoffelen, J.-M. (2020). Sensory Modality-Independent Activation of the Brain Network for Language. *The Journal of Neuroscience*, *40*(14), 2914–2924. https://doi.org/10.1523/JNEUROSCI.2271-19.2020

Arbib, M. A., Liebal, K., & Pika, S. (2008). Primate vocalization, gesture, and the evolution of human language. *Current Anthropology*, *49*(6), 1053–1076. https://doi.org/10.1086/593015

Arthurs, O. J., & Boniface, S. (2002). How well do we understand the neural origins of the fMRI BOLD signal? *Trends in Neurosciences*, *25*(1), 27–31. https://doi.org/10.1016/S0166-2236(00)01995-0

Baddeley, A. D., Hitch, G. J., & Allen, R. J. (2009). Working memory and binding in sentence recall. *Journal of Memory and Language*, *61*(3), 438–456. https://doi.org/10.1016/j.jml.2009.05.004

Baggio, G. (2020). Epistemic Transfer Between Linguistics and Neuroscience: Problems and Prospects. In R. M. Nefdt, C. Klippi, & B. Karstens (Eds.), *The Philosophy and Science of Language* (pp. 275–308). Springer International Publishing. https://doi.org/10.1007/978-3-030-55438-5_11

Baker, M. C. (2009). Language universals: Abstract but not mythological. *Behavioral and Brain Sciences*, *32*(5), 448–449. https://doi.org/10.1017/S0140525X09990604

Bastos, A. M., & Schoffelen, J.-M. (2016). A Tutorial Review of Functional Connectivity Analysis Methods and Their Interpretational Pitfalls. *Frontiers in Systems Neuroscience*, *9*. https://doi.org/10.3389/fnsys.2015.00175

Beckner, C., Blythe, R., Bybee, J., Christiansen, M. H., Croft, W., Ellis, N. C., Holland, J., Ke, J., Larsen-Freeman, D., & Schoenemann, T. (2009). Language is a complex adaptive system: Position paper. *Language Learning*, *59*, 1–26. https://doi.org/10.1111/j.1467-9922.2009.00533.x

Benítez-Burraco, A., & Murphy, E. (2019). Why Brain Oscillations Are Improving Our Understanding of Language. *Frontiers in Behavioral Neuroscience*, *13*, 190. https://doi.org/10.3389/fnbeh.2019.00190

Bergmann, T. O., & Hartwigsen, G. (2021). Inferring Causality from Noninvasive Brain Stimulation in Cognitive Neuroscience. *Journal of Cognitive Neuroscience*, *33*(2), 195–225. https://doi.org/10.1162/jocn_a_01591

Beste, C. (2021). Disconnected psychology and neuroscience—Implications for scientific progress, replicability and the role of publishing. *Communications Biology*, *4*(1), 1099. https://doi.org/10.1038/s42003-021-02634-9

Bhattasali, S., Fabre, M., Luh, W.-M., Al Saied, H., Constant, M., Pallier, C., Brennan, J. R., Spreng, R. N., & Hale, J. (2019). Localising memory retrieval and syntactic composition: An fMRI study of naturalistic language comprehension. *Language,* *Cognition and Neuroscience*, *34*(4), 491–510. https://doi.org/10.1080/23273798.2018.1518533

Blaubergs, M. S., & Braine, M. D. (1974). Short-term memory limitations on decoding self-embedded sentences. *Journal of Experimental Psychology*, *102*(4), 745–748. https://doi.org/10.1037/h0036091

Boada, R., Guasch, M., Haro, J., Demestre, J., & Ferré, P. (2020). SUBTLEX-CAT: Subtitle word frequencies and contextual diversity for Catalan. *Behavior Research Methods*, *52*(1), 360–375. https://doi.org/10.3758/s13428-019-01233-1

Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot International*, *5*((9/10)), 341–345.

Bornkessel-Schlesewsky, I., Kretzschmar, F., Tune, S., Wang, L., Genç, S., Philipp, M., Roehm, D., & Schlesewsky, M. (2011). Think globally: Cross-linguistic variation in electrophysiological activity during sentence comprehension. *Brain and Language*, *117*(3), 133–152. https://doi.org/10.1016/j.bandl.2010.09.010

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., Stabler, E. P., Van Wagenen, S. E., Luh, W.-M., & Hale, J. T. (2016). Abstract linguistic structure correlates with temporal activity during naturalistic comprehension. *Brain and Language*, *157–158*, 81–94. https://doi.org/10.1016/j.bandl.2016.04.008

Brodbeck, C., Presacco, A., & Simon, J. Z. (2018). Neural source dynamics of brain responses to continuous stimuli: Speech processing from acoustics to comprehension. *NeuroImage*, *172*, 162–174. https://doi.org/10.1016/j.neuroimage.2018.01.042

Broderick, M. P., Anderson, A. J., Di Liberto, G. M., Crosse, M. J., & Lalor, E. C. (2018). Electrophysiological Correlates of Semantic Dissimilarity Reflect the Comprehension of Natural, Narrative Speech. *Current Biology*, *28*(5), 803-809.e3. https://doi.org/10.1016/j.cub.2018.01.080

Brysbaert, M., Buchmeier, M., Conrad, M., Jacobs, A. M., Bölte, J., & Böhl, A. (2011). The Word Frequency Effect: A Review of Recent Developments and Implications for the Choice of Frequency Estimates in German. *Experimental Psychology*, *58*(5), 412–424. https://doi.org/10.1027/1618-3169/a000123

Bunge, S. A., & Kahn, I. (2009). Cognition: An Overview of Neuroimaging Techniques. In *Encyclopedia of Neuroscience* (pp. 1063–1067). Elsevier. https://doi.org/10.1016/B978-008045046-9.00298-9

Buzsáki, G., & Draguhn, A. (2004). Neuronal Oscillations in Cortical Networks. *Science*, *304*(5679), 1926–1929. https://doi.org/10.1126/science.1099745

Cai, Q., & Brysbaert, M. (2010). SUBTLEX-CH: Chinese Word and Character Frequencies Based on Film Subtitles. *PLoS ONE*, *5*(6), e10729. https://doi.org/10.1371/journal.pone.0010729

Casarotto, S., Fecchio, M., Rosanova, M., Varone, G., D'Ambrosio, S., Sarasso, S., Pigorini, A., Russo, S., Comanducci, A., Ilmoniemi, R. J., & Massimini, M. (2022). The rt-TEP tool: Real-time visualization of TMS-Evoked Potentials to maximize cortical activation and minimize artifacts. *Journal of Neuroscience Methods*, *370*, 109486. https://doi.org/10.1016/j.jneumeth.2022.109486

Caselli, N. K., Sehyr, Z. S., Cohen-Goldberg, A. M., & Emmorey, K. (2017). ASL-LEX: A lexical database of American Sign Language. *Behavior Research Methods*, *49*(2), 784–801. https://doi.org/10.3758/s13428-016-0742-0

Cecchetto, C., Giustolisi, B., & Mantovan, L. (2016). Short-Term Memory and Sign Languages. Sign Span and its Linguistic Implications. *Revista de Estudos Linguísticos Da Univerdade Do Porto*, *11*, 59–89.

Checchetto, A., Geraci, C., Cecchetto, C., & Zucchi, S. (2018). The language instinct in extreme circumstances: The transition to tactile Italian Sign Language (LISt) by Deafblind signers. *Glossa: A Journal of General Linguistics*, *3*(1), 65. https://doi.org/10.5334/gjgl.357

Chesi, C., & Canal, P. (2019). Person Features and Lexical Restrictions in Italian Clefts. *Frontiers in Psychology*, *10*, 2105. https://doi.org/10.3389/fpsyg.2019.02105

Chesi, C., & Moro, A. (2015). The subtle dependency between competence and performance. In Á. J. Gallego & D. Ott (Eds.), *50 years later: Reflections on Chomsky's Aspects* (pp. 33–45). MITWPL.

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

Chomsky, N. (1988). *Lectures on Government and Binding: The Pisa lectures* (5th ed.). Foris.

Chomsky, N. (1995). *The minimalist program*. MIT Press.

Chomsky, N. (2002). *Cartesian linguistics: A chapter in the history of rationalist thought* (J. A. McGilvray, Ed.; 2nd ed.). Cybereditions Corp.

Chomsky, N. (2017). The language capacity: Architecture and evolution. *Psychonomic Bulletin & Review*, *24*(1), 200–203. https://doi.org/10.3758/s13423-016-1078-6

Cohen, D., & Cuffin, B. N. (1991). EEG versus MEG localization accuracy: Theory and experiment. *Brain Topography*, *4*(2), 95–103. https://doi.org/10.1007/BF01132766

Cohen, M. X. (2011). It's about Time. *Frontiers in Human Neuroscience*, *5*. https://doi.org/10.3389/fnhum.2011.00002

Cohen, M. X. (2014). *Analyzing neural time series data: Theory and practice*. MIT Press.

Conde, V., Tomasevic, L., Akopian, I., Stanek, K., Saturnino, G. B., Thielscher, A., Bergmann, T. O., & Siebner, H. R. (2019). The non-transcranial TMS-evoked potential is an inherent source of ambiguity in TMS-EEG studies. *NeuroImage*, *185*, 300–312. https://doi.org/10.1016/j.neuroimage.2018.10.052

Cuetos, F., Glez-Nosti, M., Barbón, A., & Brysbaert, M. (2011). SUBTLEX-ESP: Spanish word frequencies based on film subtitles. *Psicológica*, *32*, 133–143.

Devlin, J. T., Russell, R. P., Davis, M. H., Price, C. J., Wilson, J., Moss, H. E., Matthews, P. M., & Tyler, L. K. (2000). Susceptibility-Induced Loss of Signal: Comparing PET and fMRI on a Semantic Task. *NeuroImage*, *11*(6), 589–600. https://doi.org/10.1006/nimg.2000.0595

Dimitropoulou, M., Duñabeitia, J. A., Avilés, A., Corral, J., & Carreiras, M. (2010). Subtitle-Based Word Frequencies as the Best Estimate of Reading Behavior: The Case of Greek. *Frontiers in Psychology*, *1*. https://doi.org/10.3389/fpsyg.2010.00218

Ding, N., & Simon, J. Z. (2013). Power and phase properties of oscillatory neural responses in the presence of background activity. *Journal of Computational Neuroscience*, *34*(2), 337–343. https://doi.org/10.1007/s10827-012-0424-6

Duecker, F., & Sack, A. T. (2015). Rethinking the role of sham TMS. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00210

Edwards, T., & Brentari, D. (2020). Feeling phonology: The conventionalization of phonology in protactile communities in the United States. *Language*. https://doi.org/10.1353/lan.0.0248

Embick, D., & Poeppel, D. (2015). Towards a computational(ist) neurobiology of language: Correlational, integrated and explanatory neurolinguistics. *Language, Cognition and Neuroscience*, *30*(4), 357–366. https://doi.org/10.1080/23273798.2014.980750

Evans, N., & Levinson, S. C. (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Behavioural and Brain Sciences*, *32*, 429–492. https://doi.org/10.1017/S0140525X0999094X

Fedorenko, E., Gibson, E., & Rohde, D. (2006). The nature of working memory capacity in sentence comprehension: Evidence against domain-specific working memory resources☆. *Journal of Memory and Language*, *54*(4), 541–553. https://doi.org/10.1016/j.jml.2005.12.006

Fedorenko, E., Hsieh, P.-J., Nieto-Castañón, A., Whitfield-Gabrieli, S., & Kanwisher, N. (2010). New Method for fMRI Investigations of Language: Defining ROIs Functionally in Individual Subjects. *Journal of Neurophysiology*, *104*(2), 1177–1194. https://doi.org/10.1152/jn.00032.2010

Ferreira, F., & Qiu, Z. (2021). Predicting syntactic structure. *Brain Research*, *1770*, 147632. https://doi.org/10.1016/j.brainres.2021.147632

Forscher, B. K. (1963). Chaos in the brickyard. *Science*, *142*(3590), 339–339. https://doi.org/10.1126/science.142.3590.339

Frank, S. L., Bod, R., & Christiansen, M. H. (2012). How hierarchical is language use? *Proceedings of the Royal Society B: Biological Sciences*, *279*(1747), 4522–4531. https://doi.org/10.1098/rspb.2012.1741

Friederici, A. D. (2011). The brain basis of language processing: From structure to function. *Physiological Reviews*, *91*(4), 1357–1392. https://doi.org/10.1152/physrev.00006.2011

Friederici, A. D. (2012). The cortical language circuit: From auditory perception to sentence comprehension. *Trends in Cognitive Sciences*, *16*(5), 262–268. https://doi.org/10.1016/j.tics.2012.04.001

Friederici, A. D. (2017). *Language in our brain: The origins of a uniquely human capacity*. MIT Press.

Friederici, A. D., Chomsky, N., Berwick, R. C., Moro, A., & Bolhuis, J. J. (2017). Language, mind and brain. *Nature Human Behaviour*. https://doi.org/10.1038/s41562-017-0184-4

Friederici, A. D., & Singer, W. (2015). Grounding language processing on basic neurophysiological principles. *Trends in Cognitive Sciences*, *19*(6), 329–338. https://doi.org/10.1016/j.tics.2015.03.012

Fries, P. (2015). Rhythms for Cognition: Communication through Coherence. *Neuron*, *88*(1), 220–235. https://doi.org/10.1016/j.neuron.2015.09.034

Furey, M. L., Tanskanen, T., Beauchamp, M. S., Avikainen, S., Uutela, K., Hari, R., & Haxby, J. V. (2006). Dissociation of face-selective cortical responses by attention. *Proceedings of the National Academy of Sciences*, *103*(4), 1065–1070. https://doi.org/10.1073/pnas.0510124103

Gallistel, C. R. (2017). The Coding Question. *Trends in Cognitive Sciences*, *21*(7), 498–508. https://doi.org/10.1016/j.tics.2017.04.012

Gallistel, C. R., & King, A. P. (2009). *Memory and the computational brain: Why cognitive science will transform neuroscience*. Wiley-Blackwell.

Gastaldon, S., Arcara, G., Navarrete, E., & Peressotti, F. (2020). Commonalities in alpha and beta neural desynchronizations during prediction in language comprehension and production. *Cortex*, *133*, 328–345. https://doi.org/10.1016/j.cortex.2020.09.026

Giglio, L., Ostarek, M., Weber, K., & Hagoort, P. (2021). Commonalities and Asymmetries in the Neurobiological Infrastructure for Language Production and Comprehension. *Cerebral Cortex*, bhab287. https://doi.org/10.1093/cercor/bhab287

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

Goldberg, A. (1995). *A Construction Grammar Approach to Argument Structure*. University of Chicago Press.

Goldin-Meadow, S., & Brentari, D. (2017). Gesture, sign, and language: The coming of age of sign language and gesture studies. *Behavioral and Brain Sciences*, *40*, e46. https://doi.org/10.1017/S0140525X15001247

Goldstein, A., Zada, Z., Buchnik, E., Schain, M., Price, A., Aubrey, B., Nastase, S. A., Feder, A., Emanuel, D., Cohen, A., Jansen, A., Gazula, H., Choe, G., Rao, A., Kim, C., Casto, C., Fanda, L., Doyle, W., Friedman, D., … Hasson, U. (2022). Shared computational principles for language processing in humans and deep language models. *Nature Neuroscience*, *25*(3), 369–380. https://doi.org/10.1038/s41593-022-01026-4

Gorno-Tempini, M. L., Hutton, C., Josephs, O., Deichmann, R., Price, C., & Turner, R. (2002). Echo Time Dependence of BOLD Contrast and Susceptibility Artifacts. *NeuroImage*, *15*(1), 136–142. https://doi.org/10.1006/nimg.2001.0967

Goucha, T., Anwander, A., Adamson, H., & Friederici, A. D. (2022). *Native language leaves distinctive traces in brain connections* [Preprint]. Neuroscience. https://doi.org/10.1101/2022.07.30.501987

Goucha, T., & Friederici, A. D. (2015). The language skeleton after dissecting meaning: A functional segregation within Broca's Area. *NeuroImage*, *114*, 294–302. https://doi.org/10.1016/j.neuroimage.2015.04.011

Greenberg, J. H. (2005). *Language Universals: With Special Reference to Feature Hierarchies*. DE GRUYTER MOUTON. https://doi.org/10.1515/9783110899771

Hagoort, P. (2013). MUC (Memory, Unification, Control) and beyond. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00416

Hagoort, P. (2017). The core and beyond in the language-ready brain. *Neuroscience & Biobehavioral Reviews*. https://doi.org/10.1016/j.neubiorev.2017.01.048

Hagoort, P. (2019). The neurobiology of language beyond single-word processing. *Science*, *366*(6461), 55–58. https://doi.org/10.1126/science.aax0289

Hagoort, P., & Indefrey, P. (2014). The neurobiology of language beyond single words. *Annual Review of Neuroscience*, *37*(1), 347–362. https://doi.org/10.1146/annurev-neuro-071013-013847

Hale, J., Dyer, C., Kuncoro, A., & Brennan, J. (2018). Finding syntax in human encephalography with beam search. *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 2727–2736. 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*(5), 573–582. https://doi.org/10.1080/23273798.2018.1499946

Harley, T. A. (2013). *The Psychology of Language* (0 ed.). Psychology Press. https://doi.org/10.4324/9781315859019

Hartwigsen, G. (2018). Flexible Redistribution in Cognitive Networks. *Trends in Cognitive Sciences*, *22*(8), 687–698. https://doi.org/10.1016/j.tics.2018.05.008

Hartwigsen, G., Bestmann, S., Ward, N. S., Woerbel, S., Mastroeni, C., Granert, O., & Siebner, H. R. (2012). Left Dorsal Premotor Cortex and Supramarginal Gyrus Complement Each Other during Rapid Action Reprogramming. *Journal of*

Neuroscience, *32*(46), 16162–16171. https://doi.org/10.1523/JNEUROSCI.1010-12.2012

Hartwigsen, G., & Silvanto, J. (2022). Noninvasive Brain Stimulation: Multiple Effects on Cognition. *The Neuroscientist*, 107385842211138. https://doi.org/10.1177/10738584221113806

Hartwigsen, G., Weigel, A., Schuschan, P., Siebner, H. R., Weise, D., Classen, J., & Saur, D. (2016). Dissociating Parieto-Frontal Networks for Phonological and Semantic Word Decisions: A Condition-and-Perturb TMS Study. *Cerebral Cortex*, *26*(6), 2590–2601. https://doi.org/10.1093/cercor/bhv092

Hasting, A. S., & Kotz, S. A. (2008). Speeding Up Syntax: On the Relative Timing and Automaticity of Local Phrase Structure and Morphosyntactic Processing as Reflected in Event-related Brain Potentials. *Journal of Cognitive Neuroscience*, *20*(7), 1207–1219. https://doi.org/10.1162/jocn.2008.20083

He, B., & Liu, Z. (2008). Multimodal Functional Neuroimaging: Integrating Functional MRI and EEG/MEG. *IEEE Reviews in Biomedical Engineering*, *1*, 23–40. https://doi.org/10.1109/RBME.2008.2008233

Heilbron, M., Armeni, K., Schoffelen, J.-M., Hagoort, P., & de Lange, F. P. (2022). A hierarchy of linguistic predictions during natural language comprehension. *Proceedings of the National Academy of Sciences*, *119*(32), e2201968119. https://doi.org/10.1073/pnas.2201968119

Heim, S., van Ermingen, M., Huber, W., & Amunts, K. (2010). Left cytoarchitectonic BA 44 processes syntactic gender violations in determiner phrases. *Human Brain Mapping*, *31*(10), 1532–1541. https://doi.org/10.1002/hbm.20957

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*(2–3), 61–83. https://doi.org/10.1017/S0140525X0999152X

Hernández, D., Puupponen, A., & Jantunen, T. (2022). The Contribution of Event-Related Potentials to the Understanding of Sign Language Processing and Production in the

Brain: Experimental Evidence and Future Directions. *Frontiers in Communication*, *7*, 750256. https://doi.org/10.3389/fcomm.2022.750256

Huettig, F. (2015). Four central questions about prediction in language processing. *Brain Research*, *1626*, 118–135. https://doi.org/10.1016/j.brainres.2015.02.014

Huth, A. G., de Heer, W. A., Griffiths, T. L., Theunissen, F. E., & Gallant, J. L. (2016). Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature*, *532*(7600), 453–458. https://doi.org/10.1038/nature17637

James Langille, J., & Randy Gallistel, C. (2020). The search for the engram: Should we look for plastic synapses or information-storing molecules? *Neurobiology of Learning and Memory*, 107164. https://doi.org/10.1016/j.nlm.2020.107164

Jezzard, P., & Clare, S. (1999). Sources of distortion in functional MRI data. *Human Brain Mapping*, *8*(2–3), 80–85. https://doi.org/10.1002/(SICI)1097-0193(1999)8:2/3<80::AID-HBM2>3.0.CO;2-C

Joshi, A. K., Levy, L. S., & Takahashi, M. (1975). Tree adjunct grammars. *Journal of Computer and System Sciences*, *10*(1), 136–163. https://doi.org/10.1016/S0022-0000(75)80019-5

Kaufeld, G., Bosker, H. R., ten Oever, S., Alday, P. M., Meyer, A. S., & Martin, A. E. (2020). Linguistic Structure and Meaning Organize Neural Oscillations into a Content-Specific Hierarchy. *The Journal of Neuroscience*, *40*(49), 9467–9475. https://doi.org/10.1523/JNEUROSCI.0302-20.2020

Klima, E. S., Bellugi, U., Battison, R., Boyes-Braem, P., Fischer, S., Frishberg, N., Lane, H., Lentz, E. M., Newkirk, D., Newport, E. L., Pedersen, C. C., & Siple, P. (1979). *The signs of language*. Harvard UP.

Kochari, A. R., Lewis, A. G., Schoffelen, J.-M., & Schriefers, H. (2021). Semantic and syntactic composition of minimal adjective-noun phrases in Dutch: An MEG study. *Neuropsychologia*, *155*, 107754. https://doi.org/10.1016/j.neuropsychologia.2021.107754

Krakauer, J. W., Ghazanfar, A. A., Gomez-Marin, A., MacIver, M. A., & Poeppel, D. (2017). Neuroscience needs behavior: Correcting a reductionist bias. *Neuron*, *93*(3), 480–490. https://doi.org/10.1016/j.neuron.2016.12.041

Kroczek, L. O. H., Gunter, T. C., Rysop, A. U., Friederici, A. D., & Hartwigsen, G. (2019). Contributions of left frontal and temporal cortex to sentence comprehension: Evidence from simultaneous TMS-EEG. *Cortex*, *115*, 86–98. https://doi.org/10.1016/j.cortex.2019.01.010

Kuhnke, P., Beaupain, M. C., Cheung, V. K. M., Weise, K., Kiefer, M., & Hartwigsen, G. (2020). Left posterior inferior parietal cortex causally supports the retrieval of action knowledge. *NeuroImage*, *219*, 117041. https://doi.org/10.1016/j.neuroimage.2020.117041

Lau, E. F., Gramfort, A., Hämäläinen, M. S., & Kuperberg, G. R. (2013). Automatic Semantic Facilitation in Anterior Temporal Cortex Revealed through Multimodal Neuroimaging. *The Journal of Neuroscience*, *33*(43), 17174–17181. https://doi.org/10.1523/JNEUROSCI.1018-13.2013

Law, R., & Pylkkänen, L. (2021). Lists with and without Syntax: A New Approach to Measuring the Neural Processing of Syntax. *The Journal of Neuroscience*, *41*(10), 2186–2196. https://doi.org/10.1523/JNEUROSCI.1179-20.2021

Lewis, A. G., Wang, L., & Bastiaansen, M. (2015). Fast oscillatory dynamics during language comprehension: Unification versus maintenance and prediction? *Brain and Language*, *148*, 51–63. https://doi.org/10.1016/j.bandl.2015.01.003

Lewis, R. L. (1993). *An architecturally-based theory of human sentence comprehension*. Carnegie Mellon University.

Lewis, R. L. (1996). Interference in short-term memory: The magical number two (or three) in sentence processing. *Journal of Psycholinguistic Research*, *25*(1), 93–115. https://doi.org/10.1007/BF01708421

Longobardi, G., & Roberts, I. (2010). Universals, diversity and change in the science of language: Reaction to “The Myth of Language Universals and Cognitive Science.” *Lingua*, *120*(12), 2699–2703. https://doi.org/10.1016/j.lingua.2010.04.006

Lopes da Silva, F. (2013). EEG and MEG: Relevance to Neuroscience. *Neuron*, *80*(5), 1112–1128. https://doi.org/10.1016/j.neuron.2013.10.017

Luck, S. J. (2005). *An Introduction to the Event-Related Potential Technique*. MIT Press.

Luo, H., & Poeppel, D. (2007). Phase Patterns of Neuronal Responses Reliably Discriminate Speech in Human Auditory Cortex. *Neuron*, *54*(6), 1001–1010. https://doi.org/10.1016/j.neuron.2007.06.004

Maran, M., Friederici, A. D., & Zaccarella, E. (2022). Syntax through the looking glass: A review on two-word linguistic processing across behavioral, neuroimaging and neurostimulation studies. *Neuroscience & Biobehavioral Reviews*, *142*, 104881. https://doi.org/10.1016/j.neubiorev.2022.104881

Maran, M., Grent-‘t-Jong, T., & Uhlhaas, P. J. (2016). Electrophysiological insights into connectivity anomalies in schizophrenia: A systematic review. *Neuropsychiatric Electrophysiology*, *2*(1), 6. https://doi.org/10.1186/s40810-016-0020-5

Maran, M., Numssen, O., Hartwigsen, G., & Zaccarella, E. (2022). Online neurostimulation of Broca’s area does not interfere with syntactic predictions: A combined TMS-EEG approach to basic linguistic combination. *Frontiers in Psychology*, *13*, 968836. https://doi.org/10.3389/fpsyg.2022.968836

Marian, V. (2017). Orthographic and phonological neighborhood databases across multiple languages. *Written Language & Literacy*, *20*(1), 6–26. https://doi.org/10.1075/wll.20.1.02mar

Marian, V., Bartolotti, J., Chabal, S., & Shook, A. (2012). CLEARPOND: Cross-Linguistic Easy-Access Resource for Phonological and Orthographic Neighborhood Densities. *PLoS ONE*, *7*(8), e43230. https://doi.org/10.1371/journal.pone.0043230

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, *164*(1), 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024

Marr, D. (2010). *Vision: A computational investigation into the human representation and processing of visual information*. MIT Press.

Martin, A. E. (2020). A Compositional Neural Architecture for Language. *Journal of Cognitive Neuroscience*, 1–20. https://doi.org/10.1162/jocn_a_01552

Martorell, J. (2018). Merging Generative Linguistics and Psycholinguistics. *Frontiers in Psychology*, *9*, 2283. https://doi.org/10.3389/fpsyg.2018.02283

Matar, S., Dirani, J., Marantz, A., & Pylkkänen, L. (2021). Left posterior temporal cortex is sensitive to syntax within conceptually matched Arabic expressions. *Scientific Reports*, *11*(1), 7181. https://doi.org/10.1038/s41598-021-86474-x

Matchin, W., Basilakos, A., Ouden, D.-B. den, Stark, B. C., Hickok, G., & Fridriksson, J. (2022). Functional differentiation in the language network revealed by lesion-symptom mapping. *NeuroImage*, *247*, 118778. https://doi.org/10.1016/j.neuroimage.2021.118778

Matchin, W., Brodbeck, C., Hammerly, C., & Lau, E. (2019). The temporal dynamics of structure and content in sentence comprehension: Evidence from fMRI-constrained MEG. *Human Brain Mapping*, *40*(2), 663–678. https://doi.org/10.1002/hbm.24403

Matchin, W., Hammerly, C., & Lau, E. (2017). The role of the IFG and pSTS in syntactic prediction: Evidence from a parametric study of hierarchical structure in fMRI. *Cortex*, *88*, 106–123. https://doi.org/10.1016/j.cortex.2016.12.010

Matchin, W., & Hickok, G. (2020). The Cortical Organization of Syntax. *Cerebral Cortex*, *30*(3), 1481–1498. https://doi.org/10.1093/cercor/bhz180

Matchin, W., & Wood, E. (2020). Syntax-Sensitive Regions of the Posterior Inferior Frontal Gyrus and the Posterior Temporal Lobe Are Differentially Recruited by Production and Perception. *Cerebral Cortex Communications*, *1*(1), tgaa029. https://doi.org/10.1093/texcom/tgaa029

McQueen, J. M., & Meyer, A. S. (2019). Key issues and future directions: Towards a comprehensive cognitive architecture for language use. In *Human language: From genes and brain to behavior* (pp. 85–96). MIT Press.

Meehl, P. E. (1990). Why Summaries of Research on Psychological Theories are Often Uninterpretable. *Psychological Reports*, *66*(1), 195–244. https://doi.org/10.2466/pr0.1990.66.1.195

Mesgarani, N., Cheung, C., Johnson, K., & Chang, E. F. (2014). Phonetic Feature Encoding in Human Superior Temporal Gyrus. *Science*, *343*(6174), 1006–1010. https://doi.org/10.1126/science.1245994

Meyer, L. (2018). The neural oscillations of speech processing and language comprehension: State of the art and emerging mechanisms. *European Journal of Neuroscience*, *48*(7), 2609–2621. https://doi.org/10.1111/ejn.13748

Miller, G. A., & Chomsky, N. (1963). Finitary models of language users. In R. Duncan Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. 2, pp. 419–491). Wiley.

Murphy, E. (2015). The brain dynamics of linguistic computation. *Frontiers in Psychology*, *6*(1515). https://doi.org/10.3389/fpsyg.2015.01515

New, B., Brysbaert, M., Veronis, J., & Pallier, C. (2007). The use of film subtitles to estimate word frequencies. *Applied Psycholinguistics*, *28*(4), 661–677. https://doi.org/10.1017/S014271640707035X

Ohta, S., Koizumi, M., & Sakai, K. L. (2017). Dissociating Effects of Scrambling and Topicalization within the Left Frontal and Temporal Language Areas: An fMRI Study in Kaqchikel Maya. *Frontiers in Psychology*, *8*, 748. https://doi.org/10.3389/fpsyg.2017.00748

Ojemann, J. G., Akbudak, E., Snyder, A. Z., McKinstry, R. C., Raichle, M. E., & Conturo, T. E. (1997). Anatomic Localization and Quantitative Analysis of Gradient Refocused Echo-Planar fMRI Susceptibility Artifacts. *NeuroImage*, *6*(3), 156–167. https://doi.org/10.1006/nimg.1997.0289

Ouyang, G., Sommer, W., Zhou, C., Aristei, S., Pinkpank, T., & Abdel Rahman, R. (2016). Articulation Artifacts During Overt Language Production in Event-Related Brain Potentials: Description and Correction. *Brain Topography*, *29*(6), 791–813. https://doi.org/10.1007/s10548-016-0515-1

Özyürek, A. (2014). Hearing and seeing meaning in speech and gesture: Insights from brain and behaviour. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1651), 20130296. https://doi.org/10.1098/rstb.2013.0296

Pellegrino, G., Schuler, A.-L., Arcara, G., Di Pino, G., Piccione, F., & Kobayashi, E. (2022). Resting state network connectivity is attenuated by fMRI acoustic noise. *NeuroImage*, *247*, 118791. https://doi.org/10.1016/j.neuroimage.2021.118791

Pfau, R., Steinbach, M., & Woll, B. (Eds.). (2012). *Sign language: An international handbook*. de Gruyter.

Poeppel, D., & Embick, D. (2013). Defining the relation between linguistics and neuroscience. In A. Cutler (Ed.), *Twenty-first century psycholinguistics: Four cornerstones* (pp. 103–118). Psychology Press.

Poeppel, D., & Idsardi, W. (2022). We don't know how the brain stores anything, let alone words. *Trends in Cognitive Sciences*, *26*(12), 1054–1055. https://doi.org/10.1016/j.tics.2022.08.010

Polanía, R., Nitsche, M. A., & Ruff, C. C. (2018). Studying and modifying brain function with non-invasive brain stimulation. *Nature Neuroscience*, *21*(2), 174–187. https://doi.org/10.1038/s41593-017-0054-4

Pollard, C., & Sag, I. A. (1994). *Head-driven phrase structure grammar*. University of Chicago Press.

Pyatigorskaya, E., Maran, M., & Zaccarella, E. (2022). *Testing the automaticity of syntax using masked visual priming* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/auy69

Pylkkänen, L. (2019). The neural basis of combinatory syntax and semantics. *Science*, *366*(6461), 62–66. https://doi.org/10.1126/science.aax0050

Qu, X., Wang, Z., Cheng, Y., Xue, Q., Li, Z., Li, L., Feng, L., Hartwigsen, G., & Chen, L. (2022). Neuromodulatory effects of transcranial magnetic stimulation on language performance in healthy participants: Systematic review and meta-analysis. *Frontiers in Human Neuroscience*, *16*, 1027446. https://doi.org/10.3389/fnhum.2022.1027446

Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). *Language Models are Unsupervised Multitask Learners*. https://life-extension.github.io/2020/05/27/GPT技术初探/language-models.pdf

Russo, S., Sarasso, S., Puglisi, G. E., Dal Palù, D., Pigorini, A., Casarotto, S., D'Ambrosio, S., Astolfi, A., Massimini, M., Rosanova, M., & Fecchio, M. (2022). TAAC - TMS Adaptable Auditory Control: A universal tool to mask TMS clicks. *Journal of Neuroscience Methods*, *370*, 109491. https://doi.org/10.1016/j.jneumeth.2022.109491

Salo, K. S.-T., Mutanen, T. P., Vaalto, S. M. I., & Ilmoniemi, R. J. (2020). EEG Artifact Removal in TMS Studies of Cortical Speech Areas. *Brain Topography*, *33*(1), 1–9. https://doi.org/10.1007/s10548-019-00724-w

Sassenhagen, J., & Alday, P. M. (2016). A common misapplication of statistical inference: Nuisance control with null-hypothesis significance tests. *Brain and Language*, *162*, 42–45. https://doi.org/10.1016/j.bandl.2016.08.001

Sassenhagen, J., & Draschkow, D. (2019). Cluster-based permutation tests of MEG/EEG data do not establish significance of effect latency or location. *Psychophysiology*, *56*(6), e13335. https://doi.org/10.1111/psyp.13335

Scheel, A. M., Tiokhin, L., Isager, P. M., & Lakens, D. (2021). Why Hypothesis Testers Should Spend Less Time Testing Hypotheses. *Perspectives on Psychological Science*, *16*(4), 744–755. https://doi.org/10.1177/1745691620966795

Schoffelen, J.-M., & Gross, J. (2009). Source connectivity analysis with MEG and EEG. *Human Brain Mapping*, *30*(6), 1857–1865. https://doi.org/10.1002/hbm.20745

Schroën, J. A. M., Gunter, T. C., & Hartwigsen, G. (2020). The functional relevance of left frontal and temporal cortex for semantic processing: A condition-and-perturb TMS-EEG study. *Clinical Neurophysiology*, *131*(4), e73.

Scott-Phillips, T. C., Kirby, S., & Ritchie, G. R. S. (2009). Signalling signalhood and the emergence of communication. *Cognition*, *113*(2), 226–233. https://doi.org/10.1016/j.cognition.2009.08.009

Segaert, K., Mazaheri, A., & Hagoort, P. (2018). Binding language: Structuring sentences through precisely timed oscillatory mechanisms. *European Journal of Neuroscience*, *48*(7), 2651–2662. https://doi.org/10.1111/ejn.13816

Segaert, K., Weber, K., de Lange, F. P., Petersson, K. M., & Hagoort, P. (2013). The suppression of repetition enhancement: A review of fMRI studies. *Neuropsychologia*, *51*(1), 59–66. https://doi.org/10.1016/j.neuropsychologia.2012.11.006

Shain, C., Blank, I. A., van Schijndel, M., Schuler, W., & Fedorenko, E. (2020). FMRI reveals language-specific predictive coding during naturalistic sentence comprehension. *Neuropsychologia*, *138*, 107307. https://doi.org/10.1016/j.neuropsychologia.2019.107307

Shao, Z., Janse, E., Visser, K., & Meyer, A. S. (2014). What do verbal fluency tasks measure? Predictors of verbal fluency performance in older adults. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00772

Soares, A. P., Machado, J., Costa, A., Iriarte, Á., Simões, A., de Almeida, J. J., Comesaña, M., & Perea, M. (2015). On the Advantages of Word Frequency and Contextual Diversity Measures Extracted from Subtitles: The Case of Portuguese. *Quarterly Journal of Experimental Psychology*, *68*(4), 680–696. https://doi.org/10.1080/17470218.2014.964271

Stehwien, S., Henke, L., Hale, J., Brennan, J., & Meyer, L. (2020). The Little Prince in 26 Languages: Towards a Multilingual Neuro-Cognitive Corpus. In *Proceedings of the Second Workshop on Linguistic and Neurocognitive Resources* (pp. 43–49). European Language Resources Association.

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

Steinhauer, K., & Drury, J. E. (2012). On the early left-anterior negativity (ELAN) in syntax studies. *Brain and Language*, *120*(2), 135–162. https://doi.org/10.1016/j.bandl.2011.07.001

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*(1), 64–81. https://doi.org/10.1037/0096-3445.136.1.64

Trettenbrein, P. C. (2016). The demise of the synapse as the locus of memory: A looming paradigm shift? *Frontiers in Systems Neuroscience*, *10*(88). https://doi.org/10.3389/fnsys.2016.00088

Trettenbrein, P. C., Papitto, G., Friederici, A. D., & Zaccarella, E. (2021). Functional neuroanatomy of language without speech: An ALE meta-analysis of sign language. *Human Brain Mapping*, *42*(3), 699–712. https://doi.org/10.1002/hbm.25254

Trettenbrein, P. C., Pendzich, N.-K., Cramer, J.-M., Steinbach, M., & Zaccarella, E. (2021). Psycholinguistic norms for more than 300 lexical signs in German Sign Language (DGS). *Behavior Research Methods*, *53*, 1817–1832. https://doi.org/10.3758/s13428-020-01524-y

Trettenbrein, P. C., & Zaccarella, E. (2021). Controlling Video Stimuli in Sign Language and Gesture Research: The OpenPoseR Package for Analyzing OpenPose Motion-Tracking Data in R. *Frontiers in Psychology*, *12*, 628728. https://doi.org/10.3389/fpsyg.2021.628728

UNESCO. (2021). *UNESCO World Atlas of Languages: Summary document*. https://unesdoc.unesco.org/ark:/48223/pf0000380132

Vaidya, A. R., Pujara, M. S., Petrides, M., Murray, E. A., & Fellows, L. K. (2019). Lesion Studies in Contemporary Neuroscience. *Trends in Cognitive Sciences*, *23*(8), 653–671. https://doi.org/10.1016/j.tics.2019.05.009

van der Burght, C. L., Friederici, A. D., Goucha, T., & Hartwigsen, G. (2021). Pitch accents create dissociable syntactic and semantic expectations during sentence processing. *Cognition*, *212*, 104702. https://doi.org/10.1016/j.cognition.2021.104702

van der Burght, C. L., Goucha, T., Friederici, A. D., Kreitewolf, J., & Hartwigsen, G. (2019). Intonation guides sentence processing in the left inferior frontal gyrus. *Cortex*, *117*, 122–134. https://doi.org/10.1016/j.cortex.2019.02.011

van der Burght, C. L., Numssen, O., Schlaak, B., Goucha, T., & Hartwigsen, G. (2022). Differential contributions of inferior frontal gyrus subregions to sentence processing guided by intonation. *Human Brain Mapping*, hbm.26086. https://doi.org/10.1002/hbm.26086

van Heuven, W. J. B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). Subtlex-UK: A New and Improved Word Frequency Database for British English. *Quarterly Journal of Experimental Psychology*, *67*(6), 1176–1190. https://doi.org/10.1080/17470218.2013.850521

Vartiainen, J., Liljestrom, M., Koskinen, M., Renvall, H., & Salmelin, R. (2011). Functional Magnetic Resonance Imaging Blood Oxygenation Level-Dependent Signal and Magnetoencephalography Evoked Responses Yield Different Neural Functionality in Reading. *Journal of Neuroscience*, *31*(3), 1048–1058. https://doi.org/10.1523/JNEUROSCI.3113-10.2011

Vigneau, M., Beaucousin, V., Hervé, P.-Y., Jobard, G., Petit, L., Crivello, F., Mellet, E., Zago, L., Mazoyer, B., & Tzourio-Mazoyer, N. (2011). What is right-hemisphere contribution to phonological, lexico-semantic, and sentence processing? *NeuroImage*, *54*(1), 577–593. https://doi.org/10.1016/j.neuroimage.2010.07.036

Wang, L., Schoot, L., Brothers, T., Alexander, E., Warnke, L., Kim, M., Khan, S., Hämäläinen, M., & Kuperberg, G. R. (2021). *Dynamic predictive coding across the left fronto-temporal language hierarchy: Evidence from MEG, EEG and fMRI* [Preprint]. Neuroscience. https://doi.org/10.1101/2021.02.17.431452

Weber, A., Grice, M., & Crocker, M. (2006). The role of prosody in the interpretation of structural ambiguities: A study of anticipatory eye movements. *Cognition*, *99*(2), B63–B72. https://doi.org/10.1016/j.cognition.2005.07.001

Willems, R. M. (Ed.). (2015). *Cognitive Neuroscience of Natural Language Use* (1st ed.). Cambridge University Press. https://doi.org/10.1017/CBO9781107323667

Wu, C., Zaccarella, E., & Friederici, A. D. (2019). Universal neural basis of structure building evidenced by network modulations emerging from Broca's area: The case of Chinese. *Human Brain Mapping*, *40*(6), 1705–1717. https://doi.org/10.1002/hbm.24482

Zaccarella, E., & Friederici, A. D. (2015a). Merge in the human brain: A sub-region based functional investigation in the left pars opercularis. *Frontiers in Psychology*, 1818. https://doi.org/10.3389/fpsyg.2015.01818

Zaccarella, E., & Friederici, A. D. (2015b). Syntax in the brain. In A. W. Toga, P. Bandettini, P. Thompson, & K. Friston (Eds.), *Brain mapping: An encyclopedic reference* (Vol. 3, pp. 461–468). Academic Press. http://linkinghub.elsevier.com/retrieve/pii/B9780123970251002116

Zaccarella, E., Meyer, L., Makuuchi, M., & Friederici, A. D. (2015). Building by syntax: The neural basis of minimal linguistic structures. *Cerebral Cortex*, bhv234. https://doi.org/10.1093/cercor/bhv234

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. https://doi.org/10.1016/j.neubiorev.2017.06.011