Neuroscience and Biobehavioral Reviews 142 (2022) 104881

Contents lists available at ScienceDirect

Neuroscience and Biobehavioral Reviews

journal homepage: www.elsevier.com/locate/neubiorev

Syntax through the looking glass: A review on two-word linguistic processing across behavioral, neuroimaging and neurostimulation studies

Matteo Maran [a,b,*], Angela D. Friederici [a], Emiliano Zaccarella [a]

[a] Max Planck Institute for Human Cognitive and Brain Sciences, Department of Neuropsychology, Stephanstraße 1a, 04103 Leipzig, Germany
[b] International Max Planck Research School on Neuroscience of Communication: Function, Structure, and Plasticity, Stephanstraße 1a, 04103 Leipzig, Germany

ARTICLE INFO

Keywords:
Syntax
Composition
Language
Two-word
Merge

ABSTRACT

In recent years a growing number of studies on syntactic processing has employed basic two-word constructions (e.g., "the tree") to characterize the fundamental aspects of linguistic composition. This large body of evidence allows, for the first time, to closely examine which cognitive processes and neural substrates support the combination of two syntactic units into a more complex one, mirroring the nature of combinatory operations described in theoretical linguistics. The present review comprehensively examines behavioral, neuroimaging and neurostimulation studies investigating basic syntactic composition, covering more than forty years of psycho- and neuro-linguistic research. Across several paradigms, four key features of syntactic composition have emerged: (1) the rule-based and (2) automatic nature of the combinatorial process, (3) a central role of Broca's area and the posterior temporal lobe in representing and combining syntactic features, and (4) the reliance on efficient bottom-up integration rather than top-down prediction.

1. Introduction

Scrolling through a dictionary, we easily notice that the list of entries corresponds to a large set of different meanings (tree, eat, green, etc.), but to a few grammatical categories (e.g., noun, verb, adjective, etc.). It is suggested that what enables us to produce and comprehend a virtually infinite number of linguistic expressions is based upon the capacity to combine the few available grammatical categories into abstract hierarchical structures, according to grammar (Berwick et al., 2013; Everaert et al., 2015). As an example, the adjective *big* and the noun *tree* can be combined to form the noun phrase *big tree*, in accordance with the syntactic rules of the English language. Conceptual-semantic information fills the abstract hierarchical structure with meanings from the lexicon, integrating the semantic information of individual words into more complex ones—e.g., the entity size denoted by *big* and the entity itself denoted by *tree*.

Characterizing the neural basis of language composition is a central focus of the most recent research programs in neurolinguistics (Fedorenko et al., 2016; Friederici et al., 2017; Hagoort, 2016; Martin and Baggio, 2020; Matchin and Hickok, 2020; Pylkkänen, 2019; Zaccarella and Friederici, 2017). Functional studies have provided evidence for a left-lateralised fronto-temporal network involved in language comprehension, composed of the inferior frontal gyrus (IFG), the ventromedial portion of the prefrontal cortex (vmPFC), the anterior, middle, and posterior sections of the temporal lobe (ATL, MTL and PTL), and the angular gyrus (AG; Chen et al., 2021; Fedorenko et al., 2016, 2020; Goucha and Friederici, 2015; Hagoort, 2016; Hagoort and Indefrey, 2014; Matchin et al., 2017; Matchin and Hickok, 2020; Pallier et al., 2011; Schell et al., 2017; Zaccarella et al., 2017a; Zaccarella et al., 2017b). In recent years, a few functional studies have begun to employ simple two-word combinations to more closely focus on the fundamental question of how linguistic features are combined at the most basic level within the language network (Bemis and Pylkkänen, 2011; Zaccarella and Friederici, 2015b). There is general agreement that linguistic processes driven by conceptual-semantic information particularly involve the ATL, the vmPFC, and the AG (Graessner et al., 2021a; Graessner

Abbreviations: BA, Brodmann area; EEG, Electroencephalography; ELAN, Early left anterior negativity; ERP, Event-related potential; ESN, Early syntactic negativity; fMRI, Functional magnetic resonance imaging; FOP, Frontal operculum; IFG, Inferior frontal gyrus; ISI, Interstimulus interval; LAN, Left anterior negativity; MEG, Magnetoencephalography; MTG, Middle temporal gyrus; PTL, Posterior temporal lobe; sMMN, Syntactic mismatch negativity; SOA, Stimulus onset asynchrony; STG, Superior temporal gyrus; STS, Superior temporal sulcus; TMS, Transcranial magnetic stimulation.

* Corresponding author at: Max Planck Institute for Human Cognitive and Brain Sciences, Department of Neuropsychology, Stephanstraße 1a, 04103 Leipzig, Germany.

E-mail address: maran@cbs.mpg.de (M. Maran).

https://doi.org/10.1016/j.neubiorev.2022.104881
Received 16 February 2022; Received in revised form 13 September 2022; Accepted 15 September 2022
Available online 19 September 2022

et al., 2021b), as also discussed in a recent comprehensive review (Pylkkänen, 2020) specifically addressing the so-called "red boat" paradigm (Bemis and Pylkkänen, 2011).

The way syntactic information controls linguistic processing during basic combination has received increasing attention in recent years, through a growing number of behavioral and neuroimaging studies. This growing body of evidence, which is set out to highlight the central features of the syntactic combinatorial system and which can provide fundamental guidance to the interpretation of data from longer structures, still awaits systematic examination. Moreover, the only existing review on simple syntactic combination has been published more that twenty-five years ago (Münte and Heinze, 1994), and cannot include any more recent work. In order to provide a comprehensive overview of what we know about simple syntactic combinations, the purpose of the present paper is to systematically examine all the available evidence from studies that employed two-word paradigms to test how syntactic information is used to form more complex linguistic representations. The choice to look at the very basic two-word level is motivated by a number of reasons:

i. *Detailed level of observation:* The two-word level allows to isolate the neural correlates of a single application of a compositional operation (Bemis and Pylkkänen, 2011; Zaccarella and Friederici, 2015b), which are otherwise observed only indirectly when employing more complex sentence materials. Crucially, two words are indeed sufficient to appreciate the syntactic compositional engine. For example, a determiner phrase (e.g., "The cat") already reflects the application of a syntactic rule ({Determiner}, {Noun} → Determiner Phrase).
ii. *Fitting the level of observation to linguistic theory:* At the syntactic level, the binding mechanism which combines words into phrases (called "Merge" in theoretical linguistics, Chomsky, 1995) is a binary one (i.e., operates on two elements). Therefore, from a theoretical point of view, the two-word level closely matches the properties of the basic compositional mechanism. Furthermore, as complex structures are the result of the recursive application of Merge (Chomsky, 1995), the two-word level might be sufficient to capture the key aspects of syntactic composition.
iii. *Cross-linguistic and cross-population comparison:* The use of two-word constructions greatly reduces the need for additional processes required for the analysis of long structures (e.g., working memory), which otherwise must be subtracted by careful experimental manipulations (Iwabuchi et al., 2019; Makuuchi et al., 2009). Beyond the two-word level, separating pure compositional operations from these processes without introducing additional confounds might be possible in some languages (e.g., languages with free word order) but not in others. This could hinder cross-language comparisons and, ultimately, the search for generalizable combinatorial effects. In addition, the use of paradigms with limited extra-linguistic demands makes it possible to study composition in populations whose cognitive reserve might differ from healthy adults. For example, language composition can be studied and compared in young and older populations (Poulisse et al., 2019, 2020), distinguishing its developmental trajectory from the acquisition and decline of domain general processes. Similarly, brain damage can be linked to specific linguistic deficits (Graessner et al., 2021a).
iv. *Cross-study comparison:* Despite the minimal structures involved, a wide range of detailed linguistic hypotheses can be tested by employing only two words, with the results from different experiments becoming easily comparable. An example is offered by the research program described by Pylkkänen (2020), where the use of two-word stimuli was kept constant and the dimension of composition was manipulated in a step-wise fashion across studies. This structured series of studies provided great insights into the conceptual compositional processes occurring in the ATL, portraying precisely which types of representation are combined by this region. A similar approach can be adopted to study syntactic composition.
v. *Isolating top-down and bottom-up processes:* The two-word level allows to isolate the relative contribution of top-down and bottom-up processes in language comprehension. In particular, it is possible to employ experimental conditions which restrict prediction to the first word and integration to the second one, manipulating selectively one of the two key processes (Maran et al., 2021; Matar et al., 2021). This line of research could provide crucial insights into the neural basis of parsing (Abney and Johnson, 1991; Resnik, 1992), which at present is tested by comparing top-down, left-corner and bottom-up models of language comprehension in the context of narratives or longer stimuli (Bhattasali et al., 2019; Brennan and Hale, 2019; Brennan and Pylkkänen, 2017; Hale et al., 2018; Nelson et al., 2017).
vi. *Filling a gap in the literature:* Systematic reviews on the neurocognitive correlates of basic syntactic combination are almost completely missing (Münte and Heinze, 1994), and mostly focus on studies in which the two words were presented in the visual modality.

2. Syntactic composition at the two-word level

The reviewed studies are the result of a literature search process which proceeded along three complementary lines:

i. A literature search on the "PubMed" database, including as keywords "two-word", "two words", "phrase", "Merge", in combination with "behavioral", "EEG" (Electroencephalography), "MEG" (Magnetoencephalography), "TMS" (Transcranial Magnetic Stimulation), and "fMRI" (Functional Magnetic Resonance Imaging).
ii. The additional inclusion of studies which, given their relevance in the field, were already known to the authors.
iii. A final inclusion of further relevant studies which were cited in the works retrieved in the previous two steps.

This multi-step search process resulted in the selection of approximately sixty studies, ensuring a comprehensive review of the literature on two-word syntactic processing, which is rather scattered along multiple keywords often not denoting the basic nature of the constructions employed. Accordingly, we recommend the addition of the keyword "two-word" in future studies employing basic constructions to study linguistic processing.

At the two-word level, three main paradigms have been employed to investigate the representations and cognitive processes underlying basic syntactic composition (Fig. 1): (1) categorical and agreement violations, (2) phrase versus list contrasts and (3) noun versus verb contrasts (as part of phrases and sentences, respectively). These paradigms have employed the following eleven languages: Chinese, English, Dutch, Finnish, French, German, Italian, Japanese, Serbo-Croatian, Spanish, and Standard Arabic. Note that, despite the growing interest in the neural basis of signed language processing (Blanco-Elorrieta et al., 2018; Matchin et al., 2022; Moreno et al., 2018; Stroh et al., 2019; Trettenbrein et al., 2020), we are not aware of studies addressing syntactic composition at the "two-word" level in this modality. Therefore, only studies in the visual or auditory modality are included in the present review. Furthermore, we have limited the focus of this review on studies investigating comprehension rather than production.

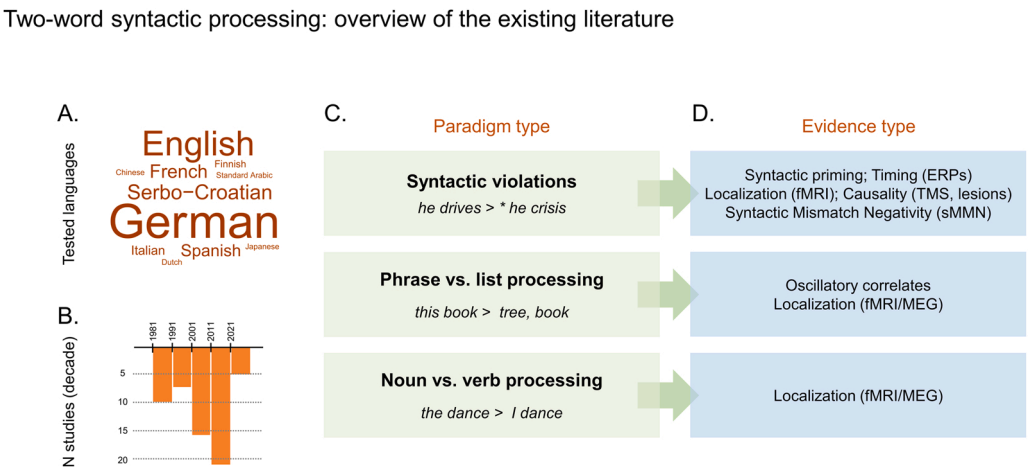

Fig. 1. : Overview of the existing literature. (A) Wordcloud representing the relative proportion of the languages tested in the reviewed studies; (B) Temporal evolution of the use of two-word paradigms at the syntactic level; (C) Summary of the paradigms, including examples; (D) Effects of interest highlighted by each of the paradigms.

2.1. Syntactic violations

Syntactic violations consist of comparing well-formed phrases and sentences with utterances that are ungrammatical, either because of the inability to combine two grammatical categories into a constituent (e.g., "*[1]he boat") or due to an agreement error (e.g., "*a boats"). Under the broad definition of syntactic violations, three lines of research can be found:

i. *Syntactic priming:* A first line of behavioral studies investigated whether a minimal syntactic context (e.g., a determiner) can affect the recognition and extraction of linguistic features from an upcoming word (e.g., a noun).[2] These studies are reviewed in Section 2.1.1.
ii. *Syntactic violations (EEG, fMRI, TMS studies*): A second line of research employed violations in combination with neuroimaging and/or neurostimulation techniques to characterise the time-course of syntactic analysis and the corresponding functional localization. These studies are reviewed in Section 2.1.2, Section 2.1.3 and Section 2.1.4.
iii. *Syntactic Mismatch Negativity*: Finally, a related line of research investigated the degree of automaticity in syntactic processing, focusing on a specific Event-Related Potential (ERP) component, namely the syntactic mismatch negativity (sMMN). These studies are reviewed in Section 2.1.5.

2.1.1. Syntactic priming

Rationale. The syntactic priming paradigm consists in the sequential presentation of two words (Fig. 2), the first one being the prime (e.g., "the", establishing a syntactic context) and the second one being the target (e.g., "book"), on which the participants perform a linguistic task (e.g., lexical decision). This paradigm allows to investigate how the certain features introduced by the syntactic context affect the processing of the incoming word, by manipulating the grammatical relationship between the prime and the target. Compared to the use of long sentences, two-word priming allows to keep working memory demands to the minimum. Secondly, confounds at the semantic level are also minimized, especially if function words are used as primes (Goodman et al., 1981). For example, a determiner presented as prime (e.g., "the") does not point to any specific semantic dimension of the target.

A first application of this paradigm is found in the study by Goodman and colleagues (1981). The authors presented nouns and verbs as target words, preceded by primes which could be determiners (e.g., "the"), possessive adjectives (e.g., "my"), personal pronouns (e.g., "you") or general nouns (e.g., "kids"). Therefore, the relationship between primes and targets could be grammatical (e.g., "you slept", "the flower") or ungrammatical (e.g., "*the slept", "*you flower"). The authors observed faster response times (RTs) in a lexical decision task for targets embedded in grammatical constructions than in ungrammatical ones, when the only contextual effect present in the block was syntactic and when the subjects were instructed regarding the prime-target relationship. Overall, this study provided initial evidence for the influence of syntactic context in word processing, while leaving open the degree to which this effect was automatic. A limitation of this study was that some of the target words employed might not have been optimal to induce syntactic violations. With one exception, target verbs were presented in the past simple form, to avoid categorical ambiguity (e.g., "slept" instead of "sleep"). However, some of the chosen verbs shared the same form for the past simple and past participle, either as an irregular (e.g., "paid") or regular (e.g., "agreed") form. Since the past participle can function as an adjectival form, some of the ungrammatical conditions (e.g., "*no agreed") could in principle still be part of a larger grammatical construction (e.g., "no agreed policy"), in which the past participle serves as a left-side modifier of a noun. This might have diminished the observed syntactic priming effect.

Morphosyntactic features. The results of Goodman and colleagues (1981) could not be easily explained in terms of spread of activation within the lexicon, contrary to semantic and associative priming effects. Under such an account, a prime (e.g., "he") would lead to the paradoxical diffuse activation of all the members of a category (e.g., verb) in the lexicon (Lukatela, Kostić et al., 1987; Seidenberg et al., 1984). The results of Goodman and colleagues (1981) inspired a line of studies employing the syntactic priming paradigm, aimed at understanding the computational basis of syntactic contextual effects. Most of these studies were conducted in Serbo-Croatian language, which has a richer inflectional system than English, therefore providing an ideal background to investigate the role of morphology in syntactic priming. A common

[1] Following the convention adopted in theoretical linguistics, we use the asterisk symbol * to denote an ungrammatical construction.

[2] We here use the general term "syntactic priming", as in the first (Goodman et al., 1981) and most recent (Berkovitch and Dehaene, 2019; Pyatigorskaya et al., 2021) applications of the paradigm. To avoid confusion, we make it explicit here that the effect we describe is different from the structural priming effect (Bock, 1986), sometimes also called "syntactic priming", which refers to the facilitation observed in processing or producing two sentences with the same grammatical structure.

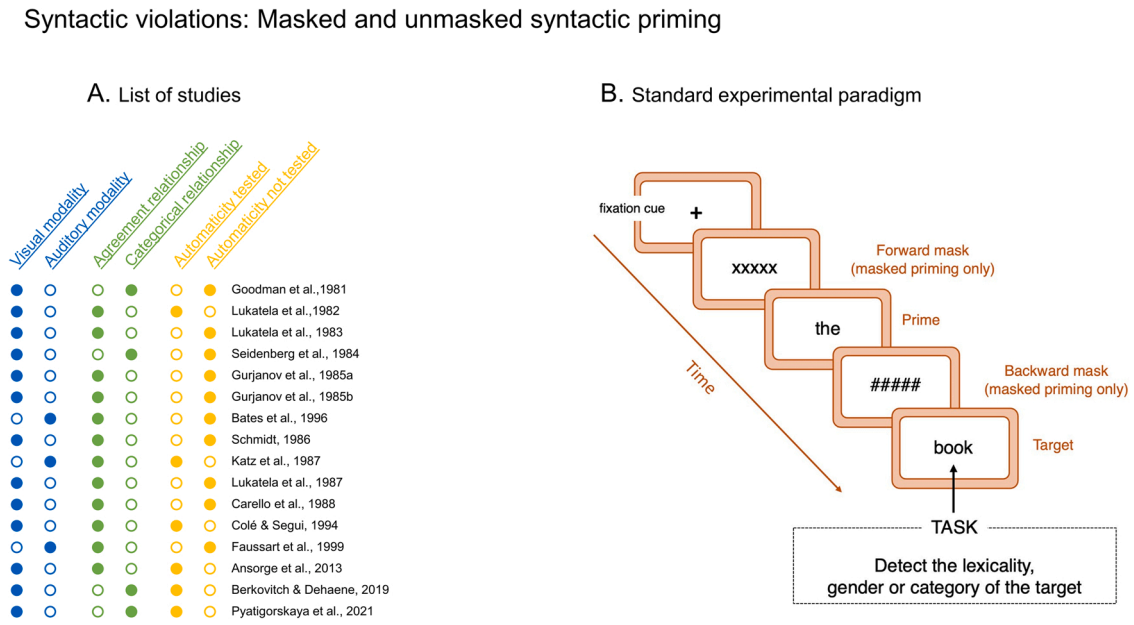

Fig. 2. : Syntactic priming studies. (A) Overview of the manipulations; (B) Example of syntactic priming.

working hypothesis behind these studies is that syntactic priming might reflect the spread of activation from a prime to a limited number of inflectional morphemes (e.g., *he* and *-s*), agreeing in syntactic features, rather than to a large number of full lexical entries (Lukatela et al., 1983).

First evidence in support of a crucial role of agreeing morphemes during syntactic priming (e.g., case, person, number) comes from a series of lexical decision tasks showing faster response times for case-marked target nouns agreeing with prepositions (Lukatela et al., 1983) and adjectives (Gurjanov, Lukatela, Moskovljević et al., 1985) presented as primes, compared to ungrammatical prime-target relationships. Similar effects were observed for target verbs agreeing with pronouns presented as primes (Lukatela et al., 1982). Notably, agreement effects were also observed in the nominative case along the dimension of grammatical gender (Gurjanov, Lukatela, Lukatela et al., 1985). Converging evidence comes from a study conducted in Italian with auditory materials (Bates et al., 1996), in which priming of grammatical gender was observed independently of phonetic transparency. These two studies support the existence of syntactic priming effects also in the absence of case-marking suffixes, converging on results observed in English at the categorical level (Goodman et al., 1981; Seidenberg et al., 1984).

Importantly, as these studies employed a lexical decision task, they also included pseudo-words, both as primes and targets. When presented as targets, pseudo-words included suffixes agreeing or not with the prime (Gurjanov et al., 1985; Lukatela et al., 1982, 1983), therefore allowing to establish grammatical or ungrammatical relationships. As pseudo-words are not elements of the lexicon, any contextual effect observed on these items must depend on grammatical and rule-based prime-suffix relationship in word processing. In line with a role of prime-suffix relationships, longer RTs were observed for rejecting pseudo-nouns and pseudo-verbs when they syntactically agreed with prepositions (Lukatela et al., 1983) or pronouns (Lukatela et al., 1982). This pattern of results suggests that the participants were influenced by the presence or absence of a grammatical relationship with the prime when evaluating the lexicality of the target, supporting a combinatorial and rule-based nature of syntactic priming effects.

All-or-none nature of syntactic priming. The studies of Gurjanov, Lukatela, Lukatela and colleagues (1985) and Bates and colleagues (1996) draw attention to the relevance of the different linguistic dimensions along which syntactic agreement takes place (e.g., case, person features, number and grammatical gender). A small number of studies further addressed this aspect, testing whether syntactic priming varies according to the type (e.g., number, gender) and number (single or double) of violations (Faussart et al., 1999; Lukatela et al., 1987), or whether it differs between closed-class and open-class elements (Colé and Segui, 1994). While initial evidence supported an all-or-none nature of syntactic priming (Colé and Segui, 1994; Lukatela et al., 1987), a subsequent study in the auditory modality reported larger interference induced by violation of grammatical gender compared to number (Faussart et al., 1999), both in French[3] and Spanish. As the authors discussed, these results might stem from the fact that gender but not number is an inherent feature of lexical stems, and that different projections exist between Number and Gender Phrases. Accordingly, violations of gender agreement might result in a more costly re-analysis process. Indeed, this notion is supported by the results of subsequent ERP studies in this domain (Barber and Carreiras, 2003, 2005).

A theoretically relevant aspect, which was addressed in only one study (Colé and Segui, 1994), is whether syntactic priming can be modulated by the type of class (function and content words) of the prime. Given the merely syntactic nature of closed-class elements, it is reasonable to expect larger effects for primes belonging to this category. Indeed, closed-class primes elicited a larger syntactic priming effect, even when controlling for frequency of occurrence. However, it remains unclear whether this effect merely reflects length differences between the open and closed-class elements.[4] Overall, further research might be needed to shed light on the all-or-none or graded nature of syntactic priming, as methodological differences (e.g., orthogonality of the manipulations, see footnote 3) or different presentation modality (visual or auditory) might have contributed to the diverging results (Colé and Segui, 1994; Faussart et al., 1999; Lukatela et al., 1987).

[3] Note that in the study of Colé and Segui (1994) and in the Experiment 1 of Faussart and colleagues (1999) only masculine target noun were used, which might be problematic for testing the influence of the gender of the prime on syntactic processing. This issue is addressed in the Experiment 2 of Faussart and colleagues (1999) with Spanish materials.

[4] Experiment 3: an average of 6 vs. 3 letters; Experiment 4: an average of 5.2 vs. 4.1 letters.

Pre-lexical vs. *post-lexical processing.* A line of studies tested whether syntactic priming arises at the pre- or post-lexical level of word processing (Carello et al., 1988; Seidenberg et al., 1984). The distinction between pre-lexical and post-lexical levels refers to whether the prime affected the target word's recognition or a later stage of its processing (e.g., integration). These studies are part of a larger discussion on whether lexical access is an encapsulated process (Forster, 1979, 1981) or can be influenced by systems outside of the lexicon (Marslen-Wilson and Tyler, 1980). To address this research question, researchers built on the properties of two different tasks, i.e. naming and lexical decision, which have been related to manipulations at the pre- and post- lexical stage respectively (West and Stanovich, 1982). In particular, as lexical decisions occur more slowly than naming, there is sufficient time to check the congruity of incoming linguistic features with the context, and bias the decision towards a negative outcome in case of a violation (Carello et al., 1988; Seidenberg et al., 1984; West and Stanovich, 1982). Furthermore, lexical decision can be characterised as a decision making process which can be biased in light of congruity checks (Seidenberg et al., 1984), contrary to naming. Both studies conducted in English (Seidenberg et al., 1984) and Serbo-Croatian (Carello et al., 1988) languages provided evidence for a *post-lexical nature* of syntactic priming. This aspect converges on the notion that, during language comprehension, the goal is an efficient integration of incoming linguistic information (Seidenberg et al., 1984), which at least at the syntactic level might be achieved without predictive processes at the pre-lexical level.

Automaticity. The post-lexical nature of syntactic priming does not necessarily imply that it stems from a non-automatic process. Note that a grammatical relationship between prime and target is not necessary to provide a correct response (i.e., indicating whether the target is a word or not) in the studies described above, but still the participants seemed unable to ignore this information (Faussart et al., 1999; Seidenberg et al., 1984), further pointing towards a strongly automatic process. Studies manipulating the stimulus-onset asynchrony (SOA) and the inter-stimulus interval (ISI) between prime and target (Colé and Segui, 1994; Katz et al., 1987; Lukatela et al., 1982) provided further evidence for the automaticity of syntactic priming. The rationale behind these manipulations is that when the target quickly follows the prime there is insufficient time for attentional and strategic processes to influence the response (Lukatela et al., 1982). Interestingly, despite the very short SOAs employed (e.g., 300, 150 and 130 ms), syntactic priming has been reliably observed in the literature (Colé and Segui, 1994; Lukatela et al., 1982).

An additional approach adopted to address the automaticity of syntactic priming is to employ a non-visual modality. As Katz and colleagues (1987) pointed out, syntactic priming was extensively studied in experiments in which primes and targets were presented visually and with no incrementality (i.e., the entire wordform was presented simultaneously and not letter by letter). In these experiments the subjects might have focused on the suffixes and ignored the stems, therefore inflating the role of morphological markers in word processing. Presenting the stimuli auditorily would therefore allow to control for this potential confound, as the subjects would hear the stem first and the suffix later. Supporting the notion of automaticity, syntactic priming has been observed in the auditory modality in different languages (Faussart et al., 1999; Katz et al., 1987).

More recent studies tested the automaticity of syntactic priming via the subliminal presentation of the prime word (Ansorge et al., 2013; Berkovitch and Dehaene, 2019; Pyatigorskaya et al., 2021), using the masking technique to suppress its conscious perception. In a first series of experiments in German, Ansorge and colleagues (2013) showed subliminal syntactic priming of grammatical gender, which was task-dependent (i.e., observed in a syntactic but not a semantic categorization task). However, as the authors discuss, the primes used in this study were ambiguous regarding the grammatical gender feature, once not only the nominative and singular forms are considered.[5,6] Subsequent subliminal syntactic priming studies focused mostly on categorical relationships between prime and target (Berkovitch and Dehaene, 2019; Pyatigorskaya et al., 2021). Syntactic priming was observed under both masked and unmasked conditions, with the target's category being recognized faster when it formed a constituent with the prime, both in French (Berkovitch and Dehaene, 2019) and German (Pyatigorskaya et al., 2021). Interestingly, subliminal syntactic priming appears to be to a certain extent independent of transition probabilities between prime and target. In particular, under specific task conditions, syntactic priming at the categorical level can be observed also when primes and targets do not agree in grammatical number (Berkovitch and Dehaene, 2019), and consequently have no probability to co-occur in natural language.

Inhibition vs. *facilitation.* A final theoretical question of great importance is whether syntactic priming reflects a facilitatory (i.e., faster RTs driven by grammatical prime-target relationship) or inhibitory (i.e., slowing of performance driven by the violation) contextual effect. Addressing this question requires the inclusion of neutral baseline conditions, whose design is not straightforward[7] (see Friederici and Jacobsen, 1999 for a detailed discussion).

Carello and colleagues (1988) used a series of "X" as a neutral prime condition, which served as a baseline condition for a separate test of associative/semantic priming (e.g., "hospital doctor" vs "XXX doctor"). While in this study the baseline condition was not directly contrasted with grammatical or ungrammatical two-word items (i.e., no statistical test was conducted to compare the two conditions), at least at the descriptive level inhibition was observed, with the ungrammatical constructions being slower than the baseline. In studies directly testing grammatical and ungrammatical conditions against a neutral baseline, inhibition has been reliably observed in lexical decision (Schmidt, 1986) and categorical/gender categorization (Bates et al., 1996; Pyatigorskaya et al., 2021) tasks. Facilitatory effects, when present, might be limited to tasks and conditions relying on surface-based processing (Bates et al., 1996; Schmidt, 1986, see also Friederici and Jacobsen, 1999 for a related discussion).

Interestingly, facilitation compared to a baseline has been observed at the two-word level when focusing on the semantic dimension (Lukatela, Carello et al., 1987). This might stem from a substantial difference in how incremental syntactic and semantic composition take place. At the semantic level, facilitation might reflect the pre-activation of a small set of lexical candidates (e.g., "sky", "sea") after a given word (e.g., "blue"). On the contrary, the syntactic engine might find little use in pre-activating a large set of candidates, for example all the nouns and adjectives, following the presentation of a determiner. Accordingly, the syntactic system might proceed with an efficient checking of incoming words and integration of their features into the previously establish context.

Critical summary. Syntactic priming studies show that the processing of a target word is affected by categorical (Berkovitch and Dehaene, 2019; Goodman et al., 1981; Pyatigorskaya et al., 2021) and agreement

[5] "Der" is both a masculine, singular, nominative form and a feminine, singular, genitive form. "Die" is both a feminine, singular, nominative form and a masculine, plural, nominative form.

[6] Note also that some of the female target words used by the authors share the same form for the nominative and genitive case (e.g., nominative: "die Gabel", *the fork*, genitive: "der Gabel", *of the fork*), and some of the masculine nouns do not change the wordform between singular and plural (e.g., "der Tiger", *the tiger*, "die Tiger", *the tigers*). The authors addressed this issue in some additional analyses on a subset of items (Experiment 2), focusing on masculine nouns with different singular and plural forms.

[7] For example, the use of inflected pseudo-words as baseline condition (Gurjanov et al., 1985) might be problematic for testing agreement relationships, given that they don't establish a neutral context along the main dimension of interest.

(Gurjanov, Lukatela, Moskovljević et al., 1985; Lukatela et al., 1982, 1983) relationships with the preceding word. The post-lexical (Carello et al., 1988; Seidenberg et al., 1984) and inhibitory (Friederici and Jacobsen, 1999; Pyatigorskaya et al., 2021) nature of this effect might reflect the impossibility of combining a given word with the established syntactic context, rather than facilitation in processing grammatical structures. This process appears to be automatic, as shown in particular by studies employing subliminal presentation of the prime (Berkovitch and Dehaene, 2019; Pyatigorskaya et al., 2021). Further studies are needed to examine whether subliminal syntactic priming might be modulated by the type of task (e.g., analysis of categorical, agreement or semantic information) performed by participants (Ansorge et al., 2013; Berkovitch and Dehaene, 2019). This research question remains still open, as it is unclear how syntactic relationships between prime and target (e.g., agreement) could affect the processing of information in a different linguistic domain (e.g., semantic categorization). Another aspect of syntactic priming which should be addressed by future studies is related to its *all-or-none* nature. In this regard, conflicting findings (Colé and Segui, 1994; Faussart et al., 1999; Lukatela et al., 1987) might be grounded on differences along linguistic dimensions which should be examined in detail (e.g., features expressed in the stem or in the suffix, see Faussart et al., 1999).

2.1.2. Neural indexes of syntactic violation processing: timing

Rationale. Unless specific analysis approaches are employed (e.g., Drift Diffusion Model, Ratcliff, 1978; Ratcliff et al., 2004; Voss et al., 2004), behavioral studies provide measures such as RTs and accuracy which can vary only along a single dimension (i.e., magnitude). On the contrary, neuroimaging techniques such Electroencephalography (EEG) and Magnetoencephalography (MEG) provide multiple dimensions along which syntactic effects can be observed, such as amplitude of the responses, latency and differences in the topographical distribution (Münte and Heinze, 1994). Furthermore, the high temporal resolution of E/MEG allows to characterize the different stages of syntactic processing, for example with a distinction between early and late components.

To the best of our knowledge, Münte, Heinze and colleagues were the first to analyse the Event-Related Potentials (ERPs, see Luck, 2005 for an introduction) elicited by syntactic violations at the two-word level (Münte et al., 1993; Münte and Heinze, 1994). The authors were particularly interested in these basic constructions to highlight syntactic violation detection, in the absence of repairing processes which might be engaged when longer sentences are employed (see also Herrmann et al., 2012 for a similar discussion). In a first experiment conducted in German, participants read determiners or pronouns as first words, followed by a noun or an inflected verb. Compared to grammatical structures, both categorical and agreement violations resulted in an increased negativity with an anterior distribution, starting approximately 300 ms after the second word presentation. A similar negativity for categorical violations at the two-word level was reported in English by the same group (Münte et al., 1993), albeit with a slightly different topography possibly related to the EEG reference site (Münte and Heinze, 1994). As we will discuss below, similar effects have been observed in the auditory modality (Hasting and Kotz, 2008) with an earlier latency onset. Notably, at least at the agreement level,[8] there is support for the independence of syntactic negativities from semantic information and co-occurrence probabilities. In particular, increased negativities were observed for case-marked violations in adjectival phrases in Finnish using pseudo-adjectives and pseudo-nouns (Münte and Heinze, 1994).

Morphosyntactic features. Barber and Carreiras (2003, 2005) investigated whether different types of morphosyntactic features, namely gender and number, are processed similarly by the brain. This line of research was based on conflicting findings from studies testing the *all-or-none* nature of syntactic priming (Colé and Segui, 1994; Faussart et al., 1999; Lukatela et al., 1987), and on the notion that gender but not number is a feature of the stem (Faussart et al., 1999). In a first experiment with Spanish materials (Barber and Carreiras, 2003), participants read two-word phrases composed of a noun and an adjective. Four conditions were included: grammatical, gender disagreement, number disagreement and double violation. Compared to grammatical constructions, syntactic violations elicited an increased negativity between 300 and 500 ms, with a central and centro-parietal topography (N400). Notably, the amplitude of the N400 was similar for agreement violations of gender, number and for double violations. Following this time-window, both grammatical and ungrammatical constructions elicited a positivity (P3), whose amplitude did not differ across conditions. Differences were however appreciated in the latency of this component, which was the fastest for the grammatical condition, followed by the double violation, the number violation and finally by the gender disagreement. Both the increased N400 for syntactic violations and the modulation of the P3 latency by the type of violation were replicated in a follow-up study (Barber and Carreiras, 2005), which included gender and number agreement both in adjective-noun pairs and determiner phrases. Overall, these two studies showed that at least the first phase of agreement analysis does not distinguish between violations of gender or number features, which are however relevant for subsequent re-analysis and repairing processes (P3). Accordingly, the larger interference created by gender violations in syntactic priming previously observed by Faussart and colleagues (1999) might map onto the re-analysis phase of two-word constructions. This follow-up study is of particular interest as it reported an additional Left Anterior Negativity (LAN) for agreement violations in determiner phrases, which was absent in adjective-noun pairs presented in isolation. However, when the same adjective-noun pairs were included in a sentence, violations resulted in a LAN followed by a P600 (Barber and Carreiras, 2005). Accordingly, the left-lateralisation of ERP agreement effects might depend on the presence of strong syntactic cues, either in the form of function words or a sentential context (Barber and Carreiras, 2005).

Automaticity. As in the case of syntactic priming, a central question on the reported ERP effects concerns their automaticity. The automatic nature of these effects is supported by two lines of evidence. First, frontal negativities for categorical and agreement violations were reported even when the participants were performing a task not related to grammatical relationships (Münte, 1992; Münte and Heinze, 1994). Second, early negativities for syntactic violations in the auditory domain were observed under conditions of visual distraction (Hasting and Kotz, 2008). These findings converge on the automatic nature of syntactic processing, previously shown behaviorally by subliminal syntactic priming studies (Berkovitch and Dehaene, 2019; Pyatigorskaya et al., 2021).

Auditory modality. Compared to the reading studies reviewed above, earlier latencies have been reported for syntactic violations at the two-word level in the auditory modality. Hasting and Kotz (2008) reported an Early Syntactic Negativity (ESN) with a latency onset of approximately 100 ms for categorical and agreement violations. The ESN was elicited both when participants performed a grammaticality judgement task and when their attention was diverted away from the stimuli, supporting the automaticity of syntactic analysis. The presence of the ESN was replicated in two experiments with agreement (Jakuszeit et al., 2013) and categorical violations (Maran et al., 2021). Of note, Jakuszeit and colleagues (2013) failed to replicate the ESN effect for categorical

[8] We want to point out that it might be particularly challenging (if not virtually impossible) to induce categorical violations using pseudo-words (e.g., "flirk"), as they can take any grammatical function (e.g., "they flirk", "the flirk"). Orthographic and phonological differences between grammatical categories might be exploited to construct, for example, pseudo-nouns and pseudo-verbs. Studies employing this strategy require careful designs, ensuring that grammaticality is orthogonal to low-level properties of the pseudo-words (see for example the design of Hasting and Kotz, 2008, where grammaticality is orthogonal to the category of both the first and second word. See also Steinhauer and Drury, 2012 for similar methodological considerations).

violations, possibly due to reduced statistical power.

From a methodological point of view, the earlier latency of the ESN compared to the negativities elicited in the visual modality (N400 and LAN) might reflect a more accurate time-locking of the ERP analysis (see also Hasting and Kotz, 2008 for a similar argument). In the visual modality it is not possible to establish at which timepoint the offending morphemes are processed, therefore the ERP analysis is usually time-locked to the word onset. Conversely, with auditory materials, Hasting and Kotz (2008) could time-lock the ERP analysis precisely at the onset of the phonemes eliciting the agreement violations (e.g., grammatical: "Du fal|test", *you fold*, "Er fal|tet", *he folds*; ungrammatical: "*Du fal|tet", *you folds*, "*Er fal|test", *he fold,* with | corresponding to the point of time-locking). A similar precision was obtained when testing categorical violations, since nouns and verbs with an ambiguous stem were used (e.g., grammatical: "Ein Fal|ter", *a butterfly*, "Er fal|tet", *he folds*; ungrammatical: "*Ein Fal|tet", *a he folds*, "*Er fal|ter", *he butterfly*). Techniques such as the cross-splicing further allowed to control for phonetic and acoustic features up to the time-locking point (Hasting et al., 2007; Hasting and Kotz, 2008; Shtyrov et al., 2003). An interesting aspect of ESN studies (Hasting and Kotz, 2008; Jakuszeit et al., 2013; Maran et al., 2021) is that their design, in combination with the cross-splicing technique, allows to investigate grammaticality effects which are orthogonal to the identity and the category of both words and to low-level acoustic features of the employed stimuli (see Steinhauer and Drury, 2012 for a related discussion).

The ESN is not the only ERP component elicited by two-word auditory categorical or agreement violations. First, while the effect was first reported in the 100–300 ms time-window under attentive conditions (Hasting and Kotz, 2008), follow-up studies showed a prolonged negativity lasting approximately until 500 ms (Jakuszeit et al., 2013; Maran et al., 2021), possibly reflecting the overlap of an ESN and a subsequent N400. This second negativity, reminiscent of the effect reported by Barber and Carreiras (2003, 2005), might reflected additional processes related to processing a mismatching suffix, as it was observed in conditions in which syntactic violations were realised with overt marking (Jakuszeit et al., 2013; Maran et al., 2021). Second, an increased late positivity starting approximately at 500 ms has been reported for agreement (Hasting and Kotz, 2008) and categorical (Jakuszeit et al., 2013; Maran et al., 2021) violations when the participants judge the grammaticality of items (Hasting and Kotz, 2008; Maran et al., 2021) or are not actively distracted (Jakuszeit et al., 2013). However, when care is taken to ensure that participants are not attending the auditory stimuli, only the ESN is observed (Hasting and Kotz, 2008). In this regard, the functional profiles of the ESN and the late positivity mirror the ones of the Early Left Anterior Negativity (ELAN) and the P600 observed with longer stimuli (Friederici, 2011; Hahne and Friederici, 1999), suggesting that the existence of a first automatic and a second late controlled stage of analysis is a core aspect of syntactic composition.

Extra-linguistic information. Preliminary evidence suggests that grammatical two-word phrases can elicit negativities, closely resembling those elicited by agreement violations, when their information does not match extra-linguistic context (Arcara et al., 2019). In particular, when determiner phrases (e.g., "some apples") denote a quantity not matching the previously depicted one (a single apple) a broad negativity is observed. The existence of potential similarities between the neural systems processing grammatical number and quantities should be addressed in future studies. This represents an interesting research question, given that brain regions processing numerical information such as the right parietal lobe have also been shown to process grammatical number agreement (Carreiras et al., 2010).

Additional aspects. We would like to highlight some methodological aspects of the ERPs studies discussed above. As pointed out in the context of studies employing longer constructions (Steinhauer and Drury, 2012), sustained differences between grammatical and ungrammatical conditions before the critical word (i.e., before the second word) might be problematic, especially if combined with pre-processing steps as baseline correction. To this end, orthogonal designs such as those employed by ESN studies and baseline correction procedures based on an interval preceding the first word onset might be preferred. Furthermore, the appropriate application of common pre-processing steps such as high-pass filtering, re-referencing and baseline correction to language studies is not trivial (Maess et al., 2016; Molinaro et al., 2015; Tanner, 2015; Tanner et al., 2015; Widmann et al., 2015), and should be guided by careful methodological considerations.

Critical summary. ERP studies employing syntactic violations provide evidence for two stages of analysis. Syntactic negativities (Hasting and Kotz, 2008; Maran et al., 2021; Münte and Heinze, 1994) seem to reflect an all-or-none detection of ungrammaticality, given that they are not modulated by the type (Hasting and Kotz, 2008) or number (Barber and Carreiras, 2003) of syntactic violations. These negativities are observed also under distraction conditions (Hasting and Kotz, 2008), pointing towards an early automatic stage of syntactic analysis. Late positivities reflect slower task-dependent repairing processes (Barber and Carreiras, 2003, 2005; Jakuszeit et al., 2013; Maran et al., 2021), in which fine-grained distinctions between the types of violations emerge (Barber and Carreiras, 2003, 2005). Modulation of this component's latency seems to reflect the fact that repairing linguistic features expressed by the stem might be more costly than those overtly marked. This finding supports and complements the result of a previous syntactic priming study addressing a similar research question (Faussart et al., 1999), described in the previous section.

2.1.3. Neural indexes of syntactic violation processing: localization

Rationale. Two-word syntactic violations have been used in combination with techniques with high spatial resolution (e.g., Functional Magnetic Resonance Imaging, fMRI) to highlight brain regions involved in evaluating whether the grammatical information of incoming words can be integrated into a constituent. A first fMRI reading study (Kang et al., 1999) employed well-formed verb phrases and two-word constructions being ill-formed, either due to a categorical error or to an unlikely meaning (e.g., "heard shirts"). Both syntactic and semantic violations resulted in activation of Broca's area, albeit with different patterns in its cytoarchitectonic subregions. On the one hand, Broca's area pars opercularis (Brodmann area, BA, 44) showed left-lateralised activity for both types of violations, but larger for the syntactic type. On the other hand, Broca's area pars triangularis (BA45) and the more anterior portions of the left prefrontal cortex (BA46 and BA10) showed left- and right-lateralised activity for syntactic and semantic violations respectively. This study provided first evidence for a dissociation between brain regions processing syntactic and semantic information, observable even at the most minimal level of composition and within Broca's area. As Kang and colleagues (1999) pointed out, in this study grammaticality was confounded with the category of the second word, which was always a noun in well-formed structures (e.g., "blew bubbles") and a verb in violations (e.g., "*grew heard"). However, as described below, the involvement of the left IFG in processing syntactic violations received large support by studies overcoming this limitation.

Categorical violations. Adapting the original ESN paradigm (Hasting and Kotz, 2008), Herrmann and colleagues (2012) contrasted auditory well-formed two-word phrases and violations at the categorical level. A German preposition or a pronoun was presented as the first word, followed by an uninflected noun or an inflected verb. The four combinations of first and second words allowed to orthogonally manipulate grammaticality and perceptual markedness of the second word (i.e., the presence of the inflectional suffix "-t"). The left BA44 and temporal lobe were linked to syntactic processing, while the bilateral temporal cortices to the detection of perceptual marking. The involvement of Broca's area in categorical analysis converges on previous fMRI data at the two-word level and points towards modality-independent computations (see also Heim et al., 2010 for converging evidence in the agreement domain), given that it was previously observed with visual presentation of phrases (Kang et al., 1999). On the contrary, at least at the two-word level the

involvement of temporal regions in syntactic analysis seems to be restricted to the auditory modality. Due to the poor temporal resolution of fMRI, it remains underspecified which of these activations reflect the early (ESN) and the late (P600) stages of analysis discussed above.

Agreement violations. To the best of our knowledge, only two fMRI studies have employed agreement violations at the two-word level (Carreiras et al., 2010; Heim et al., 2010). These two studies converge in highlighting the left IFG and premotor area as key regions in processing agreement violations. Using the same paradigm of a previous EEG study (Barber and Carreiras, 2005), Carreiras et al. (2010) presented their participants with two-word expressions (adjective and determiner phrases) which could be well-formed or contain gender or number agreement violations. Both types of agreement violations resulted in increased activation of the left IFG and the left premotor cortex, mirroring the pattern of LAN and N400 observed in EEG studies (Barber and Carreiras, 2003, 2005). Additionally, number violations in determiner phrases resulted in increased activity of the right parietal lobe, a region involved in numerical processing (Arcara et al., 2021; Dehaene et al., 2003; Zorzi et al., 2011). While this study was conducted in the visual modality, Heim et al. (2010) showed that activations triggered by agreement violations in the left IFG and premotor cortex are modality-independent. This study, which focused on violations of grammatical gender agreement, reported an additional involvement of the left posterior parietal and right supplementary motor areas.

Critical summary. The left IFG appears as a key region in analysing and detecting the violation of grammatical rules, both at the categorical (Herrmann et al., 2012) and agreement (Carreiras et al., 2010; Heim et al., 2010) level. In this regard, the profile of activation of the left IFG seems to mirror the presence of equivalent early negativities for categorical and agreement violations at the two-word level (Hasting and Kotz, 2008). Outside of the left IFG, these two domains differentially engage additional brain regions: superior and posterior middle temporal regions are activated by categorical errors (Herrmann et al., 2012), while agreement violations are processed by premotor and parietal regions (Carreiras et al., 2010; Heim et al., 2010; Fig. 3B). Differences between the types of agreement violations, which were previously observed at the behavioral level (Faussart et al., 1999) and linked to late ERP components (Barber and Carreiras, 2003, 2005), map outside of the left IFG. In particular, the right parietal lobe has been linked to processing number violations (Carreiras et al., 2010), while grammatical gender processing has been linked to the left parietal lobe (Heim et al., 2010). Accordingly, these two regions might be linked to late repairing processes, indexed by previous behavioral (Faussart et al., 1999) and ERP studies (Barber and Carreiras, 2003, 2005).

2.1.4. Neural indexes of syntactic violation processing: causality

Rationale. A limitation of neuroimaging techniques is that they allow to test only a correlational link between brain activity and a given linguistic process of interest. To overcome this limitation, two complementary approaches have been employed. First, neuroimaging and behavioral data have been acquired from patients with specific brain lesions (Vaidya et al., 2019). Second, neurostimulation techniques such as Transcranial Magnetic Stimulation (TMS) have been employed to temporarily disrupt brain functioning in healthy participants engaged in specific tasks of interest (Hartwigsen, 2015), possibly in combinations with neuroimaging techniques. The rationale of both approaches is that, if a brain region is necessary for a given process of interest, its permanent damage (lesion studies) or transient perturbation (TMS studies) should affect the behavioral or neural correlates of the operation of interest.

Cortical lesions. Patient and neurostimulations studies employing two-word syntactic violations have focused on Broca's area, given the fMRI studies described above. In a follow-up study employing the ESN paradigm, Jakuszeit and colleagues (2013) compared healthy controls and patients with left IFG lesion to test the causal role of this region in generating top-down categorical and morphological predictions. The presence of top-down syntactic predictions (Lau et al., 2006) could in principle explain the earliness of violations effects observed in ERP studies (ELAN, Friederici et al., 1993; ESN, Hasting and Kotz, 2008). Contrary to healthy controls, patients with left IFG lesion did not show an ESN for agreement violations. While this finding has been linked to a role of the left IFG in generating top-down predictions (Jakuszeit et al., 2013), it is also compatible with a role of the left IFG in detecting an agreement violation, as by definition this region in patients is damaged both during the prediction and integration phase (see also Maran et al., 2021 for a similar discussion).

The study of Jakuszeit et al. (2013) also revealed that categorical violations elicited a late positivity[9] in healthy controls but not in patients. This is in principle compatible with a causal role of the left IFG in repairing or re-analysis processes when a minimal syntactic context is available. Note that, as an intact P600 was previously observed in patients with Broca's aphasia (Friederici et al., 1999), it remains an open question under which circumstances repairing and re-analysis processes can be initiated despite damage of the left IFG.

Neurostimulation. The causal role of the left IFG in generating categorical predictions has been recently tested in a combined TMS-EEG study (Maran et al., 2021). In this study, the authors employed an ESN paradigm with categorical violations, with determiners and pronouns being followed by nouns or verbs. Using online TMS,[10] Broca's area functioning was disrupted specifically at the predictive stage of a two-word phrase (e.g., during a determiner predicting a noun). Compared to lesion studies (Jakuszeit et al., 2013), this approach has the advantage of inducing a temporally specific interference, targeting a specific cognitive process. Interestingly, the authors reported a prolonged early negativity effect, possibly reflecting the overlap of an ESN and a second posterior effect, followed by a late positivity. However, these components were not affected by Broca's area stimulation at the predictive stage. Even at a more fine-grained level of analysis, with an accurate modelling of the TMS-induced electrical field (Numssen et al., 2021; Weise et al., 2020), no effect of Broca's area disruption during the predictive phase could be observed. As the authors suggested, these findings do not support a causal role of Broca's area in generating syntactic predictions at the categorical level, but are compatible with a role of this region in the bottom-up integration of words into constituents. This remains a testable hypothesis for future studies. Indeed, a previous study found an effect of TMS over the left IFG during the second word (Sakai et al., 2002), when focusing on violations of verb transitivity in basic two-word construction. Note however that this study was conducted on a very limited sample size (N = 6).

The causal involvement of Broca's area in syntactic processing, and specifically in agreement analysis, has been further tested by an online TMS behavioral study (Carreiras et al., 2012). In this study, participants performed a grammaticality judgement task on determiner phrases, which contained agreement violations or were well-formed. Replicating previous findings (Carreiras et al., 2010), faster RTs were observed for grammatical than ungrammatical items. Crucially, this difference was selectively reduced by TMS over Broca's area. This study therefore provides initial evidence for the causal involvement of Broca's area in morphosyntactic processing, and specifically in the syntactic checking and integration of features, given that TMS was delivered during the second word of a constituent. In light of this study, the absence of an ESN

[9] In this study, categorical violations failed to elicit a significant ESN effect in both controls and patients. Importantly, the small number of patients and healthy controls included in the study of Jakuszeit and colleagues (10 per group) might have made the study underpowered, especially for detecting early categorical effects, where two types of constituents (determiner phrases and sentences) are present in each level of grammaticality.

[10] The term online TMS is used to refer to the delivery of TMS pulses during the task, as opposed to offline TMS where after-effects of the stimulation are used to induce a virtual lesion (for a review, see Hartwigsen, 2015).

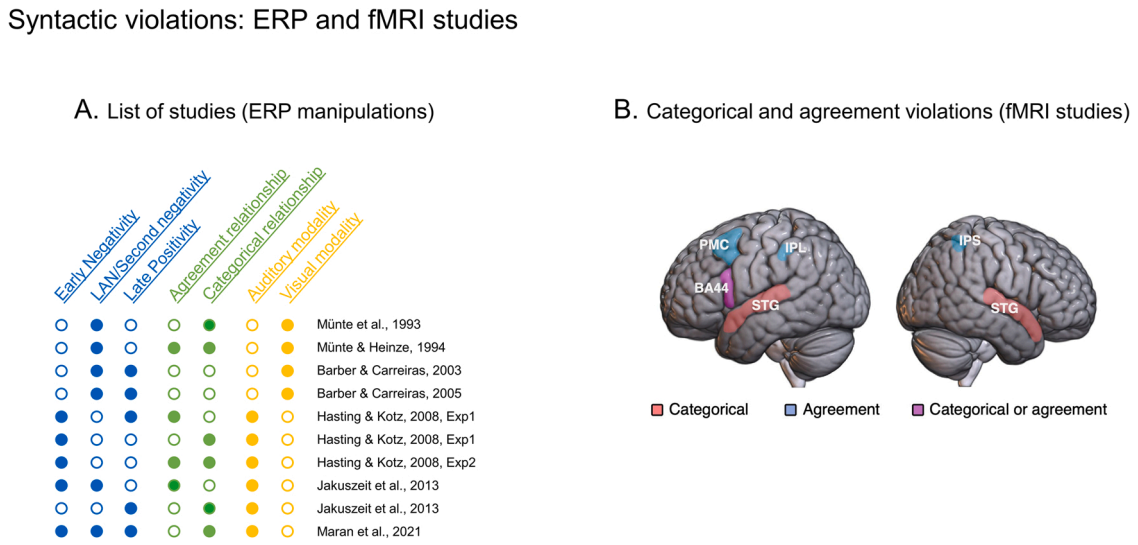

Fig. 3. : Summary of neuroimaging studies employing two-word syntactic violations. (A) Overview of the ERP studies; (B) Key brain regions highlighted by fMRI studies. BA44: Brodmann area 44; STG: superior temporal gyrus; PMC: premotor cortex; IPL: inferior parietal lobe; IPS: intraparietal sulcus.

for agreement violations in patients with left IFG lesions (Jakuszeit et al., 2013) could also reflect a deficit in bottom-up integration rather than top-down prediction.

Critical summary. The available patient (Jakuszeit et al., 2013) and TMS (Carreiras et al., 2012) studies support a causal involvement of Broca's area in processing syntactic violations in the agreement domain. Further studies are needed to draw stronger conclusions on the causal role of this region in categorical processing. On the one hand, patient studies in this domain (Jakuszeit et al., 2013) might have been affected by reduced statistical power (see footnote 9). On the other hand, Broca's area might have been stimulated during an earlier phase compared to the critical one (i.e., the prediction rather than the integration phase) in the study of Maran and colleagues (2021). In this regard, the available evidence from the agreement domain supports a role of this region in bottom-up syntactic integration rather than prediction, given that its stimulation during the second word causally affected the participants' performance (Carreiras et al., 2012).

On a final note, recent methodological advances allow to precisely estimate the degree of stimulation of target regions when using neurostimulation techniques (Numssen et al., 2021; Weise et al., 2020), revealing graded effects at the individual level (e.g., Kuhnke et al., 2020). This method represents a promising tool for future TMS studies on syntactic processing.

2.1.5. *Syntactic violations: syntactic Mismatch Negativity (sMMN)*

Rationale. Syntactic Mismatch Negativity (sMMN) studies employ an adapted version of the classic oddball paradigm, in which a rarely presented deviant stimulus occurs among repeated standard stimuli. In its traditional form, deviant and standard stimuli are presented acoustically to the subjects, who are distracted by means of a silent movie. A classic ERP finding is the observation of an increased response (MMN) for the deviant stimuli compared to the standard ones (Näätänen et al., 1978). Traditional MMN paradigms have focused on central auditory processing, as standard and deviant stimuli usually differ along a dimension of interest such as frequency and intensity (for a review see Näätänen et al., 2007). In its sMMN form, a grammatical (e.g., "we come") and an ungrammatical (e.g., "*we comes") construction are used as stimuli, being employed both as a standard and as a deviant in the oddball (Fig. 4). For example, in a first block "we come" and "*we comes" are used as standard and deviant, respectively. In a second block, the opposite assignment is employed, taking care to counterbalance the order of the blocks across subjects (Pulvermüller and Shtyrov, 2003). The sMMN effect consists in a larger MMN response evoked by ungrammatical deviants compared to grammatical ones (Fig. 4), usually observed with a latency smaller than 200 ms (Pulvermüller and Shtyrov, 2003; Shtyrov et al., 2003). This effect stems from automatic linguistic processing, as similarly to the classic MMN studies participants are usually actively distracted by means of silent movies.

In sMMN paradigms, great care is taken to control for acoustic features of the stimuli. Procedures such as the cross-splicing technique make the different items acoustically identical up to the information which makes trials grammatical or ungrammatical (Pulvermüller and Shtyrov, 2003). Furthermore, control conditions are included to ensure that the sMMN depends on grammaticality and is not confounded by acoustic features (e.g., the presence of the "-s" suffix in the examples above). In this regard two strategies are usually employed. Additional blocks can be included, in which the two second words are presented out of context (e.g., "[noise] come", "[noise] comes"), both as deviants and standards. The MMN response to these stimuli can be subtracted from the respective two-word utterances (e.g. "we come" – "[noise] come" and "*we comes" – "[noise] comes"), therefore allowing to control for acoustic differences between grammatical and ungrammatical items (Shtyrov et al., 2003). A second strategy consists in using two-word constructions which make grammaticality orthogonal to the suffix and first word (e.g., "Mä tuon", *I bring*$_{[1ST\ PERSON\ SINGULAR]}$, "*Mä tuot", **you bring*$_{[2ND\ PERSON\ SINGULAR]}$, "*Sä tuon", *you bring*$_{[1ST\ PERSON\ SINGULAR]}$, "Sä tuot", *you bring*$_{[2ND\ PERSON\ SINGULAR]}$, from the Finnish materials in Shtyrov et al., 2003). From a methodological point of view, the use of only few two-word items in each study (e.g., one correct and one incorrect) might allow to characterise very early stages of linguistic analysis, which might be lost in the ERP averaging process if stimulus materials variance results in jittered evoked activity (Pulvermüller and Shtyrov, 2003). At the same time, generalizability of the results needs to be ensured by varying the stimulus lists across studies.

Automaticity. As in the case of attended syntactic violations, sMMN paradigms have mostly focused on agreement and categorical features. At the agreement level, sMMN effects have been reported in several languages: English (Hanna et al., 2016; Pulvermüller and Shtyrov, 2003; Pulvermüller et al., 2008), Finnish (Shtyrov et al., 2003), French (Brunellière et al., 2007; Hanna et al., 2014) and German (Hasting et al., 2007; Lucchese et al., 2017a; Lucchese et al., 2017b; Pulvermüller and Assadollahi, 2007). These studies converge in reporting an automatic effect of grammaticality, with ungrammatical two-word utterances used

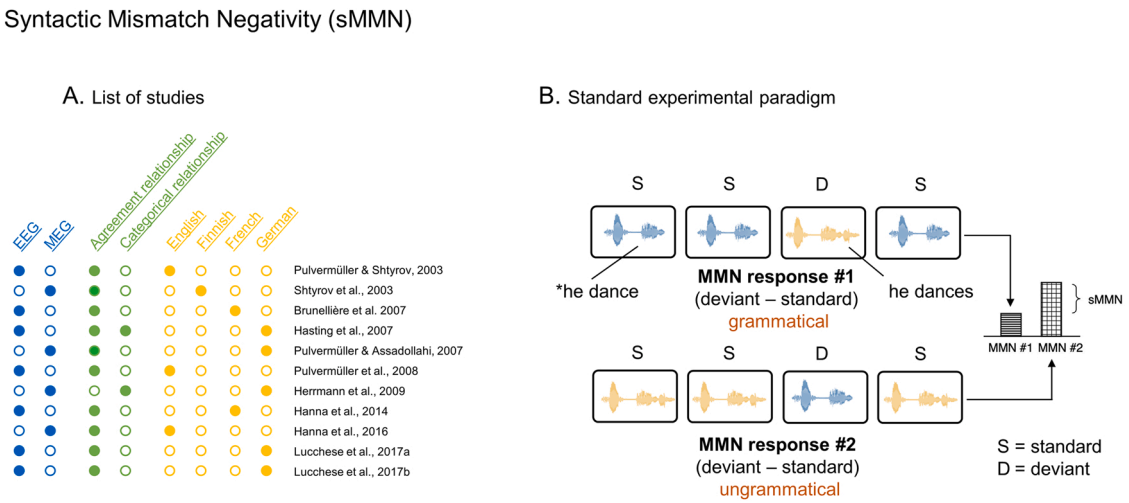

Fig. 4. : Summary of the sMMN studies. (A) List of studies, with an overview of the included manipulations; (B) Example of a standard sMMN paradigm.

as deviants eliciting a more negative[11] MMN that the correct counterparts, generally occurring before 200 ms. Similar effects have been reported in sMMN studies employing violations at the categorical level (Hasting et al., 2007; Herrmann et al., 2009).

From a cognitive point of view, the sMMN effect has been described as a neurophysiological trace of syntactic priming (Hasting et al., 2007; Pulvermüller and Shtyrov, 2003). At least at the agreement level, this effect has been linked to facilitation for grammatical structures rather than the result of violation detection, as no difference was observed between the MMN evoked by ungrammatical deviants and words presented following noise (Pulvermüller and Shtyrov, 2003, 2006). It is unclear whether this explanation applies to categorical sMMN effects, as a recent syntactic priming study including a baseline condition reported only inhibition effects for ungrammatical constructions (Pyatigorskaya et al., 2021). An interesting hypothesis is that these diverging findings might reflect a difference between the processing of agreement features (Pulvermüller and Shtyrov, 2003) and categorical information (Pyatigorskaya et al., 2021). In this regard, it has been pointed out that the sMMN components for agreement and categorical violations have different topographies (Hasting et al., 2007), possibly reflecting separate underlying processes.

Rule-based analysis. One potential issue which arises when comparing grammatical and ungrammatical expressions is that these conditions differ not only at the syntactic level, but also in the frequency of occurrence, which is virtually zero for violations. Accordingly, the sMMN and the results from the studies employing syntactic violations reviewed above might reflect either a grammatical process or purely differences in transition probabilities. Two carefully designed sMMN experiments addressed this issue, supporting a purely grammatical nature of the processes reflected in this ERP component. The rationale of these two sMMN studies (Herrmann et al., 2009; Pulvermüller and Assadollahi, 2007), focusing on categorical and agreement features respectively, is very similar. In both studies, together with grammatical (frequent) and ungrammatical two-word expressions, rare grammatical constructions were used. The authors hypothesised that, if the sMMN reflects the application of grammar, the MMN evoked by rare and common grammatical constructions would be similar and smaller than syntactic violations. This is exactly what was observed in both studies, providing evidence for a purely grammatical explanation of the reported effects.

Seriality vs. *interactivity.* Studies employing basic sMMN paradigms support the notion that syntactic analysis occurs in an automatic fashion. Building on these findings, a recent multi-feature sMMN paradigm (see Näätänen et al., 2004 for a description) sought to investigate whether lexico-semantic and syntactic processing would occur strictly in a serial fashion or would interact in early stages of processing (Lucchese et al., 2017b). In this study, two dimensions of the utterances were manipulated, namely the lexicality of the second word (real verb or pseudo-verb) and the grammaticality of the agreement relationship between pronoun and second word. This study reported an early interaction between lexicality and agreement analysis, challenging serial processing of linguistic information. While a clear sMMN effect was observed for the double violation, the fact that the MMN of grammatical items and agreement violations realised with real verbs were not statistically different is somehow difficult to interpret. One possibility is that this pattern of results depends on the multi-feature nature of the sMMN paradigm employed by the authors. Given the relevance of this research question for psycholinguistic models of language comprehension, future replications of this study might provide clear insights into the early interaction between lexical access and agreement feature analysis. The syntactic priming studies employing inflected pseudo-words reviewed above can further provide a behavioral background to this line of research.

Neural generators. With respect to the neural generators of the sMMN, the available evidence points towards an involvement of the left superior temporal lobe (Hanna et al., 2016; Herrmann et al., 2009; Shtyrov et al., 2003) and frontal (Pulvermüller and Shtyrov, 2003) regions, with one study highlighting Broca's area in particular (Hanna et al., 2014). Additionally, sources in the primary auditory cortex have also been reported in one study (Herrmann et al., 2009), in which the sMMN effect was observed only in the presence of overt marking of the category. At present it is still unclear why no consistent source localization of the sMMN is present in the literature. Some authors have suggested that the different results might depend on the neuroimaging techniques used (Herrmann et al., 2009; Pulvermüller and Shtyrov, 2006), as the sMMN was localised in the left frontal cortex in EEG studies (Hanna et al., 2014; Pulvermüller and Shtyrov, 2003), while the involvement of the temporal lobe was shown with MEG recordings (Hanna et al., 2016; Herrmann et al., 2009; Shtyrov et al., 2003). The diverging findings might be grounded in the different sensitivity of EEG and MEG to radial and tangential dipoles, or to distortions inherent to the source localization procedure (Baillet, 2017). Thus, at present further evidence is needed to draw more precise conclusions on the neural generators of the sMMN. Studies acquiring simultaneous EEG and MEG recordings might provide crucial insights on the localization debate (Baillet, 2017).

Additional aspects. While the reviewed studies highlighted the

[11] In one study (Lucchese et al., 2017a) an effect of the opposite polarity was observed, possibly due to the short SOA used by the authors.

relevance of the sMMN for characterising syntactic processes, this paradigm might also have important clinical applications. In particular, the sMMN might allow to avoid both an under- and an over-estimation of linguistic functioning, which would otherwise interfere both with the diagnosis and treatment evaluation. On the one hand, this paradigm ensures that non-specific effects of lesions or conditions (e.g., fatigue or lack of sustained attention) lead to an underestimation of patients' functioning, given that no task is employed. On the other, the lack of a task might allow to evaluate the treatment outcome, linking it specifically to linguistic abilities. This is of particular importance as at later stages of recovery task-related and strategic processes might become available to patients, leading to an overestimation of their language abilities (Lucchese et al., 2017a). Recently, a sMMN multi-feature paradigm has been used to assess the outcome of speech language therapy in patients with post-stroke aphasia (Lucchese et al., 2017a). In this study, the MMN was recorded employing a grammatical two-word expression, a grammatical but not lexical sentence (pronoun and correctly inflected pseudo-verb) and an agreement violation. Patients were tested before and after four weeks of intensive training. Treatment selectively affected the MMN evoked by grammatical constructions and correctly inflected pseudo-word sentences. This finding nicely fits with the observation that ungrammatical constructions and words out of context elicit comparable MMN (Pulvermüller and Shtyrov, 2003), possibly suggesting that speech therapy began to restore the functionality of the syntactic combinatorial mechanism. However, considering the small sample size (N = 10) and the lack of a significant correlation between the changes in the MMN amplitude and standardised clinical inventories of language abilities, further studies are needed to draw stronger conclusions. Furthermore, the lack of a healthy control group and of an analysis based on the grammaticality of the deviant does not allow to estimate whether a proper sMMN effect was restored following speech therapy.

A further application of the sMMN paradigm is the study of the neural basis of second language acquisition. In this context, the sMMN is an idea tool to quantify rapid and automatic syntactic processes, as the ones supporting everyday language use (Hanna et al., 2016). Furthermore, considering that only few stimuli are used as standards and deviants, it is possible to test grammatical knowledge even in learners who might have a small vocabulary of the second language. In a two-word study focusing on agreement features, Hanna and colleagues (2016) compared the sMMN response of English native speakers and non-native learners with different levels of proficiency. Interestingly, comparable sMMN were observed for native speakers and learners with high-level of proficiency, in both cases larger than the non-proficient group. Furthermore, sMMN effects were localised in the bilateral superior temporal lobes, and significant correlations were found between source-level activity and behavioral measures of syntactic abilities. Thus, this finding supports the notion that second language analysis might be supported by early and automatic mechanisms, similar to those of the first language.

We conclude this section by pointing out that, as in some studies employing attended syntactic violations (Barber and Carreiras, 2003, 2005; Jakuszeit et al., 2013; Maran et al., 2021), some sMMN studies reported a second negativity following the early effect of grammaticality (Hanna et al., 2014; Hasting et al., 2007; Herrmann et al., 2009). As we have previously proposed (Maran et al., 2021), this negativity could reflect to the detection of a violation between an expected suffix and the current one being heard. The presence of this effect in sMMN paradigms provides initial evidence for an automatic nature of its underlying cognitive process.

Critical summary. Studies employing the sMMN paradigm have provided evidence for automatic rule-based syntactic processes, independent of transition probabilities (Herrmann et al., 2009; Pulvermüller and Assadollahi, 2007). This notion converges on the subliminal syntactic priming studies discussed above. While the available evidence points towards neural generators in the fronto-temporal cortices (Hanna et al., 2014, 2016; Herrmann et al., 2009; Pulvermüller and Shtyrov, 2003), further studies are still needed for a precise definition of the sMMN sources. In this regard, the combination of EEG and MEG might provide further insights into the role of the neuroimaging technique employed in mapping sMMN onto frontal or temporal regions (Herrmann et al., 2009; Pulvermüller and Shtyrov, 2006). At present, only few studies have adopted the sMMN paradigm to investigate the acquisition (Hanna et al., 2016) or repairing (Lucchese et al., 2017a) of syntactic abilities, which await further replication. Given the advantages offered by the sMMN to address these aspects (i.e., absence of an overt task and the requirement of a limited vocabulary knowledge), the replication of these two studies might provide a new and impactful tool to evaluate linguistic abilities in healthy and clinical conditions.

2.2. Phrases and sentences versus lists

Rationale. The rationale of the phrase/sentence-versus-list paradigm is simple: phrases (e.g., a determiner phrase) are characterised by hierarchical dependencies between a head (e.g., a determiner) and a dependent (e.g., a noun), while lists (e.g., a noun followed by a noun) are not. Therefore, this paradigm can be used to highlight the neural correlates of hierarchical processing in language. In principle, hierarchical processing can be tested comparing sentences and lists beyond the two-word level, an approach employed by several fMRI (Goucha and Friederici, 2015; Matchin et al., 2017; Snijders et al., 2010; Zaccarella et al., 2017a) and MEG studies (Hultén et al., 2019; Lam et al., 2016; Schoffelen et al., 2017). However, unless careful control conditions are included (Goucha and Friederici, 2015; Pallier et al., 2011), contrasting sentences and lists does not highlight specifically syntactic composition, since in the former type of stimuli both hierarchical structure and sentential meaning are built. Furthermore, long sentences differ from lists not only because of the presence of a hierarchical structure, but also because of "progression effects". In particular, words which are part of a sentence, especially in late positions, benefit from contextual effects which are absent in the list conditions and have been linked to specific neurophysiological correlates (Hultén et al., 2019; Lam et al., 2016). Thus, by highlighting the single analysis of the head-dependent relationship, two-word constructions provide an ideal level of observation of hierarchical processing.

We discuss in this section also a recent MEG study conducted in Standard Arabic language (Matar et al., 2021), which tested hierarchical processing by manipulating syntactic complexity in semantically matched two-word expressions. This study did not include a list condition, as three different structures were compared: indefinite phrase, definite phrase, full sentence. However, the comparison of syntactic trees of increased complexity is related to contrasting structures with and without a hierarchical structure.

2.2.1. Phrases and sentences versus lists: neurophysiological correlates

EEG and MEG studies employing the phrase-versus-list paradigm mostly focused on the oscillatory dynamics of syntactic processing, rather than its ERP correlates. This allowed researchers to interpret their findings in light of a rich literature, addressing how specific rhythms can be well-suited for distinct linguistic computations (for reviews see: Benítez-Burraco and Murphy, 2019; Lewis et al., 2015; Meyer, 2018; Murphy, 2015; Weiss and Mueller, 2012). This approach complements and extends the ERP literature on syntactic processing, mostly based on the use of violations (Friederici, 2011). Furthermore, it provides the possibility to understand the neural basis of syntactic composition under primary principles governing brain functioning, such as the coordination of neural activity within networks (Friederici and Singer, 2015; Fries, 2015) and the processing and integration of information which might occur at different time-scales (Meyer, 2018).

Alpha and beta oscillatons. Alpha and beta oscillations have been linked to syntactic binding by recent EEG (Poulisse et al., 2020; Segaert et al., 2018) and MEG (Hardy et al., 2022; Matar et al., 2021) studies,

albeit with their results not always converging. For example, while Segaert and colleagues (2018) showed increased synchronization of alpha and beta oscillations for pseudo-sentences (e.g., "He grushes") compared to lists of two pseudo-verbs, the same effect did not reach significance in a follow-up study (Poulisse et al., 2020). Furthermore, an MEG study employing the same paradigm reported an effect in the opposite direction in the alpha band (Hardy et al., 2022), together with changes in connectivity in this rhythm between the left IFG and temporal lobe following the presentation of the second word. Potential differences in the language of investigation (Dutch and English) and in the neuroimaging techniques employed (MEG and EEG) in these two studies might need to be addressed in future replication studies. On the one hand, MEG is less sensitive than EEG to dipoles oriented radially to the skull (Ahlfors et al., 2010; Baillet, 2017) and to deeper sources of activity (Baillet, 2017; Goldenholz et al., 2009). On the other, EEG signal is more affected by spatial distortions stemming from the different electrical conductivities of the biological tissues interposed between the brain and the electrodes (Baillet, 2017; Da Silva, 2013; van den Broek et al., 1998). Finally, another recent MEG study (Matar et al., 2021) comparing two-word constituents of different syntactic complexity linked hierarchical processing to beta-band[12] oscillations.

Age-related changes. A recent study conducted in English, adapting the paradigm of Segaert and colleagues (2018), tested whether differences in the neural correlates of syntactic binding could be observed when comparing older and young adults (Poulisse et al., 2020). This study builds on previous evidence showing that, especially in the absence of semantic information, the older population shows impairment in processing basic two-word constructions (Poulisse et al., 2019). Older adults showed increased synchronization in the theta, alpha and beta bands for the list condition compared to the pseudo-sentence. Significant differences between groups in the syntactic binding effect (pseudo-sentence versus list) were observed in the alpha band, reflecting the trend[13] present in the young adults and the effect in the opposite direction observed in the older adults. It must be noted that in this study significant differences between pseudo-sentences and lists were also observed in the time- and phase-locked ERPs, with increased positivity in the P1, P3 and P600 components.

Critical summary. The available EEG and MEG studies employing the phrase-versus-list paradigm point towards an involvement of alpha and beta oscillations in syntactic binding (Fig. 5), albeit sometimes with effects in opposite directions (Hardy et al., 2022; Segaert et al., 2018). The role of the specific neuroimaging techniques (EEG and MEG) and tested languages (e.g., Dutch and English) in driving conflicting findings should be examined by future studies. Importantly, differences between the complexity of constituents (e.g., definite and indefinite phrases, sentences) described in the field of theoretical linguistics affect the neurophysiological correlates of syntactic binding (Matar et al., 2021), and should be carefully considered when contrasting them against lists.

As shown in Fig. 5, some of these studies reported differences between constituents and lists extending across the first and second word of the stimuli. While these findings are in principle compatible with the co-existence of syntactic expectation and integration, some of the statistical tests employed in these studies (e.g., cluster-based permutation test) bear limitations on the claims that can be made regarding the latency of a given effect (see Sassenhagen and Draschkow, 2019 for a detailed discussion). Furthermore, in some of these studies the first word was a real pronoun in the constituent condition and a pseudo-verb in the list, therefore confounding the generation of syntactic predictions with lexicality. Further studies, building on the present methodological considerations, might provide a solid link between alpha and beta oscillations and incremental parsing operations.

2.2.2. Phrases and sentences versus lists: functional localization

The cortical syntactic network. The comparison of phrases versus lists has highlighted the left IFG and the PTL as key regions involved in hierarchical processing. A first fMRI study by Zaccarella and Friederici (2015b) contrasted determiner phrases including a German pseudo-noun (e.g., "Diese Flirk", *this flirk*) and noun-pseudo-noun lists (e.g., "Apfel Flirk", *apple flirk*). Additional one-word conditions were presented, including a determiner or a noun followed by a non-linguistic character string. Increased activity was found in the ventral-anterior portion of BA44 for two-word phrases, which was the only condition with hierarchical dependency between the linguistic units. Conversely, evolutionary older brain regions, such as the frontal operculum (FOP) and anterior insula (aINS), reflected mere accumulation of words, with increased activation for both two-word phrases and lists compared to single words (Zaccarella and Friederici, 2015a, 2015b).

With the exception of one study (Bozic et al., 2015), which observed a bilateral temporal pattern of activation, the other studies employing this paradigm support a role of a left-lateralized network in building basic two-word constituents. At the two-word level, activity of the left IFG might be amplified by merging operations involving closed-class elements. In a follow-up fMRI study (Schell et al., 2017), participants were presented with a German noun in isolation or in two phrasal contexts, following either a determiner or an adjective. Compared to single words, both types of phrases resulted in increased activation of the left Broca's area, albeit with different patterns. Phrases including a closed-class element activated specifically Broca's area pars opercularis (BA44), while increased activation in the pars triangularis (BA45) was found for the combination of an adjective and a noun (see also Schell et al., 2022 for a dissociation between BA 44 and 45 in representing determiner phrases and noun phrases). Additional activations for the determiner phrase were observed in the posterior middle temporal gyrus (pMTG) and superior temporal sulcus (STS). Direct comparison of the determiner phrase against the adjective-noun pair revealed increased activity in the ventral portion of the left IFG and posterior temporal regions. A similar study conducted in Chinese further supports the role of Broca's area in opening determiner phrases (Wu et al., 2019). This study also highlighted directed connectivity from this region to the posterior middle temporal lobe as a key neural process for building constituents. This finding converges on the recent observation of stronger interactions between the left IFG and the temporal lobe during the second word of a basic two-word construction (Hardy et al., 2022). An involvement of the temporal lobe in hierarchical processing is also supported by the recent MEG study with Standard Arabic noun-adjective pairs (Matar et al., 2021), briefly introduced at the beginning of this section. Notably, in this study effects were observed both during the first and second word, pointing towards a role of this region in both syntactic prediction and integration. However, the same first words presented in isolation in a separate block did not elicit any reliable effect. Therefore, potential predictive effects in the PTL might be to a certain extent strategic and not automatic.

Differences with the red-boat studies. We have not discussed in the present section the so-called "red boat" studies (e.g., Bemis and Pylkkänen, 2011), given that a well-structured research program has highlighted the semantic and conceptual nature of this task (Pylkkänen, 2019, 2020). Studies employing the "red boat" paradigm contrast two-word and one-word expressions, in compositional (e.g., "red boat" vs "xtp boat") and list (e.g., "cup boat" vs "xtp boat") contexts. Across multiple studies, activation of the left ATL and ventromedial prefrontal cortex (vmPFC) for two-word expressions was observed only in the compositional context (and the respective task). However, given that an adjective and a noun form a constituent, the absence of effects in the left IFG and PTL is somehow unexpected and deserves careful consideration.

A first possible reason for the null effects in these regions is the lack

[12] Note that Matar and colleagues (2021) reported two clusters (8–32 Hz and 8–20 Hz) which extend beyond the canonical beta-band (12–30 Hz).

[13] The lack of statistical significance might be due to the reduced power of the follow-up study, which included approximately 65% of the trials per condition compared to the first study by Segaert and colleagues (2018).

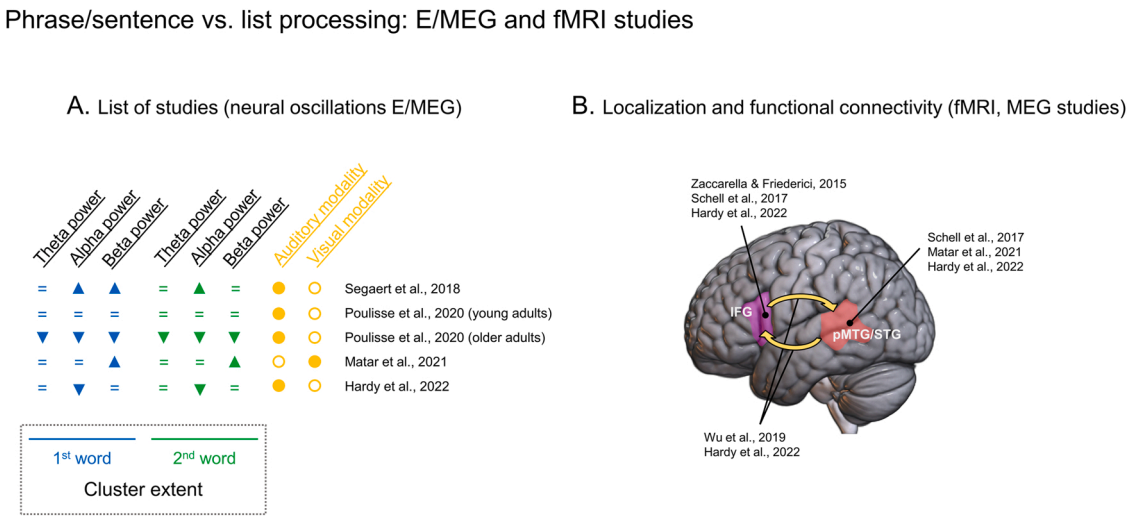

Fig. 5. : Summary of studies comparing phrases and sentences against lists. (A) List of E/MEG studies. The symbol ▲ denotes an increase in synchronization driven by hierarchical processing, ▼ denotes a decrease in synchronization, = denotes no significant change reported; (B) Key brain regions highlighted by fMRI and MEG studies. IFG: Inferior frontal gyrus; pMTG: posterior middle temporal gyrus; pSTG: posterior superior temporal gyrus.

of function words in the adjective-noun pairs, as reliable activation of the left IFG (Wu et al., 2019; Schell et al., 2017; Zaccarella and Friederici, 2015b) and posterior temporal regions (Matar et al., 2021) is found when constituents include closed-class elements. Secondly, variability in the time-course of syntactic composition might make this process more elusive to techniques with high temporal resolution such as MEG and EEG (Kochari et al., 2021), especially in the absence of an overt violation. Similarly, activity in left IFG might be more difficult to capture with MEG, due to its anatomical position and the reduced coverage from the helmet sensors. As previously pointed out (Kochari et al., 2021), the use of structural templates rather than individual MRI images might also have resulted in a minor sensitivity to changes of activity in this region in some of the "red boat" studies, especially given the precise localization of composition effects in BA44's clusters (Zaccarella and Friederici, 2015b). Thirdly, it is important to consider that building minimal syntactic structures might rely on synchronized oscillatory activity (see previous section) which is not necessary phase-locked, contrary to the one highlighted in the "red boat" studies (see Murphy et al., 2022 for an exception). Fourth, it is possible that no effects were observed in these two regions because in the red-boat English studies the adjective (e.g., "red") could also be interpreted as a noun (e.g., the colour red), resulting in a noun-noun list. Indeed, in languages where the presence of inflectional morphology removes ambiguity regarding the category of the first word, increased activation is observed in the left IFG for phrases composed of an adjective and a noun (e.g., "Blaues Schiff", *blue ship*) compared to single words (Schell et al., 2017). Finally, two studies employing the "red boat" paradigm (Fló et al., 2020; Neufeld et al., 2016) have also shown that the neural correlates highlighted might, under certain circumstances, reflect task-progression or expectancy effects rather than compositional processes. Experimental designs dissociating explicit and implicit semantic processing at the phrasal level (Graessner et al., 2021a) might provide useful insights into the task-dependency of the conceptual-semantic network activation.

Critical summary. The left IFG and PTL (Fig. 5B) appear as key regions in combining categorical information into hierarchical structures (Matar et al., 2021; Schell et al., 2017; Wu et al., 2019; Zaccarella and Friederici, 2015b). Within the left IFG and the temporal lobe (Matar et al., 2021; Schell et al., 2017), differences in brain activations appear to be driven by the type of constituents built (e.g., determiner phrases and noun phrases; see also Schell et al., 2022). Interestingly, not only changes in the level of activation, but also in the transmission of information between the left inferior frontal and temporal regions might subserve the formation of syntactic constituents (Hardy et al., 2022; Wu et al., 2019). A central research question for future studies is whether activations of these regions vary cross-linguistically, possibly according to the role of serial order in different languages (Bozic et al., 2015).

2.3. Comparison of nouns and verbs in two-word constructions

Rationale. In the previous sections we have addressed the neural basis of syntactic composition, either with the use of violations or with the comparison of constituents and lists. Especially when focusing on phrasal building effects, an underlying assumption is the existence of neuro-cognitive operations to extract categorical information from incoming words. However, despite the fact that nouns and verbs represent one of the most basic distinctions found in human languages, isolating their neural substrates is a challenging task (Crepaldi et al., 2011, 2013; Vigliocco et al., 2011), given that these two categories differ along several syntactic and semantic dimensions.

A small number of studies have employed basic two-word constructions to highlight the neural basis of nouns' and verbs' categorical representations. Most of these studies share the assumption that, in order to observe traces of categorical differences, nouns and verbs need to be part of constructions where their syntactic features are necessarily extracted (see Vigliocco et al., 2011, for a similar discussion). At present these studies allow to draw only preliminary conclusions, given their limited number and conflicting findings, possibly stemming from the limited sample sizes tested and variability in the employed design and materials. To improve the readability of the present work, these studies are summarized in Table 1 and in the section "Critical summary" below. We limit the in-text discussion only to key studies which allow to formulate methodological considerations on which further studies can build.

Paradigms and methodological considerations. A promising approach to investigate categorical differences between nouns and verbs is the use of homonyms and homophones in minimal syntactic contexts (e.g., "the dance", "they dance"). This manipulation offers two major advantages. First, orthographic and acoustic differences are by definition controlled for, as nouns and verbs share the same form. Notably, this can be achieved not only in English (Burton et al., 2009; Lee and Federmeier, 2006; Pulvermüller et al., 2012; Tyler et al., 2008), but also in some languages with a more complex morphological system (e.g., Italian, see Tsigka et al., 2014). Second, semantic differences between nouns and verbs can

Table 1
Summary of neuroimaging studies comparing nouns and verbs in minimal syntactic contexts.

Study	Technique	Sample	Language	Design	Key results
Burton et al. (2009)	fMRI	12	English	Presentation of ambiguous (i.e., homonyms) and unambiguous nouns and verbs in a minimal syntactic context.	- Independently of ambiguity, nouns elicited increased activation of the left IFG. - When ambiguous, nouns elicited a larger response of the left pSTS compared to verbs. - The opposite effect was observed in this region for unambiguous conditions.
Fiebach et al. (2002)	MEG	6	German	Presentation of unambiguous nouns and verbs, in isolation or in a two-word context.	- No differences between nouns and verbs emerged when they were presented in isolation. - When part of a basic syntactic context, nouns elicited increased activation of the PTL compared to verbs.
Lee and Federmeier (2006)	EEG	26	English	Presentation of unambiguous nouns and verbs in a minimal context. Additional semantically ambiguous items were presented, whose meaning could be similar (e.g., "to vote" and "the vote") or vary (e.g., "to duck" and "the duck") according to the assigned category.	- Independently of semantic ambiguity, nouns elicited a larger N400. - Unambiguous verbs elicited an increased frontal positivity. - Words characterized by both class and semantic ambiguity (e.g., "duck") elicited a sustained fronto-central positivity.
Pulvermüller et al. (2012)	fMRI	23	English	Presentation of noun-verb homonyms, which could be inflected or in the stem form, as part of two-word syntactic contexts. Homonyms varied along the semantic dimension, referring to the arm, face, or leg.	- A trend towards increased activation of BA44 and the left premotor cortex for determiner phrases compared to sentences was observed. - Activation of the motor system driven by the semantic content of the homonyms was observed only in stems.
Strijkers et al. (2019)	MEG	22	French	Presentation of unambiguous nouns and verbs, in isolation or in a two-word context.	- No differences between nouns and verbs emerged when they were presented in isolation. - When part of a basic syntactic context, nouns elicited increased activation of the left and right IFG. A similar effect was observed when contrasting determiners and pronouns.
Tsigka et al. (2014)	MEG	12	Italian	Presentation of homonyms (e.g., "dance") in a minimal syntactic context (e.g., "I dance", "the dance").	- Verbs elicited stronger activation than nouns in the left inferior frontal and right parietal lobes, together with a larger involvement of central areas. - Compared to determiners, pronouns elicited stronger activation of the left prefrontal and right parietal lobes.
Tyler et al. (2008)	fMRI	15	English	Presentation of noun-verbs homophones, either in isolation or in a minimal syntactic context.	- Verbs resulted in increased activation of the left pMTG, but only when presented in a basic syntactic context.

be minimized (e.g., "the dance", "they dance"; Tsigka et al., 2014; Tyler et al., 2008), or included as a factor in the experimental design once the two categories have different meanings (e.g., "the duck" versus "to duck", Lee and Federmeier, 2006).

An important methodological consideration, especially for EEG and MEG studies, concerns categorical differences between determiners and pronouns presented as first words before nouns and verbs in this paradigm. It is important to ensure that any effect observed during the presentation of nouns and verbs does not merely reflect a prolonged difference stemming from the previous word (see Steinhauer and Drury, 2012 for similar methodological considerations), since determiners and pronouns have been linked to different brain responses (Strijkers et al., 2019; Tsigka et al., 2014). A first approach to overcome this issue is the inclusion of control conditions in which determiners and pronouns are presented in isolation, in order to subtract the relative brain responses from the two-word conditions. Alternatively, specific analyses need to be employed to dissociate pre-target and post-target syntactic effects (see Strijkers et al., 2019 for an example).

Critical summary. At the neural level, categorical differences between nouns and verbs might emerge when they are part of constituents, rather than presented in isolation (Fiebach et al., 2002; Strijkers et al., 2019; Tyler et al., 2008). In most of the reviewed studies, effects have been reported in the left IFG (Burton et al., 2009; Pulvermüller et al., 2012; Strijkers et al., 2019; Tsigka et al., 2014) and/or PTL (Burton et al., 2009; Fiebach et al., 2002; Tyler et al., 2008), albeit sometimes in opposite directions. For example, increased activation for nouns in the left IFG compared to verbs has been reported by some authors (Burton et al., 2009; Pulvermüller et al., 2012; Strijkers et al., 2019), while Tsigka and colleagues (2014) reported the opposite pattern. Similarly, increased activation of the PTL for verbs than nouns has been reported by Tyler and colleagues (2008), but the opposite pattern was observed in a previous MEG study (Fiebach et al., 2002).

The lack of convergence across studies should be considered in light of important methodological and theoretical considerations. Some of the reviewed studies tested a relatively small number of participants (i.e., less than 20 participants, see Table 1), which might have resulted in reduced power to detect categorical effects. The semantic content of the nouns and verbs tested constitutes another factor which might have driven conflicting findings across studies. As shown by EEG and fMRI studies specifically designed to address this aspect (Burton et al., 2009; Lee and Federmeier, 2006; Pulvermüller et al., 2012), semantic differences between nouns and verbs can modulate the activation of the language network and surrounding brain regions, and should be carefully controlled by future studies.

On a final note, two-word constructions formed with nouns and verbs are, in most of the cases, two different types of constituents, namely phrases and sentences. Thematic role assignment might be differently engaged by these constructions, unless specific types of phrases are employed (Zaccarella et al., 2017a). The extent to which thematic role assignment is responsible for the effects observed in the left IFG and PTL in the reviewed studies remains, at present, an important and testable research question for future studies.

3. Key aspects of basic syntactic composition

We have reviewed more than fifty studies, conducted in eleven languages, which have employed basic two-word constructions to investigate the behavioral and neural basis of syntactic processing. Across these studies, four key features of the syntactic combinatorial system have emerged, which are summarized in the following sections (Fig. 6).

3.1. Combining two words into a constituent is a rule-based process

A first central feature of the syntactic combinatorial system is that it builds structures according to abstract grammatical rules, operating at the categorical and agreement level. Syntactic priming studies (e.g., Goodman et al., 1981; Lukatela et al., 1983, see Section 2.1.1) demonstrated that, when the categorical or agreement features of two words

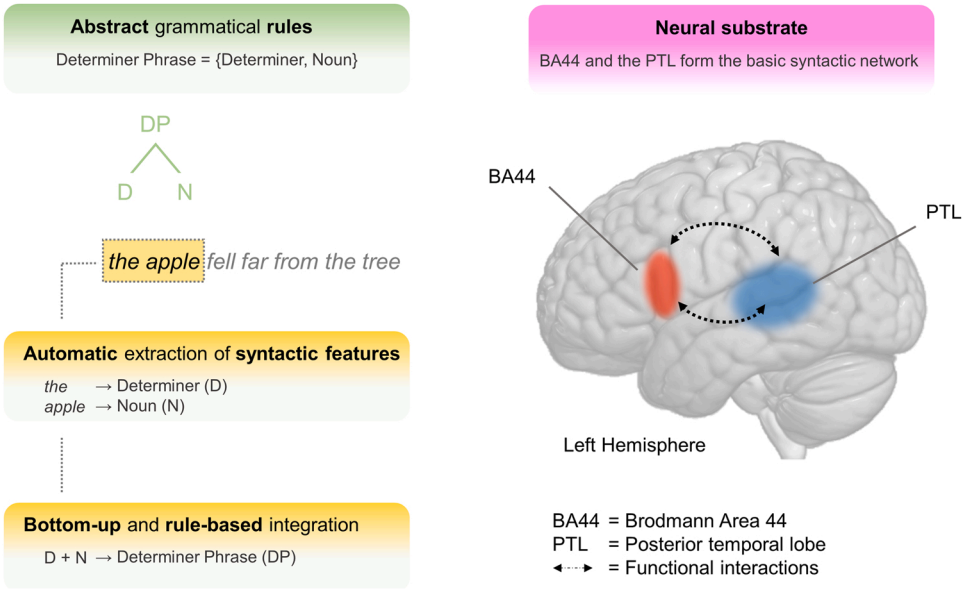

Fig. 6. : The three levels of description of basic syntactic composition. At the formal level, grammar (in green) provides a set of abstract rules which govern the combination of two words into a constituent. At the computational level, grammatical rules are implemented with the automatic extraction and integration of syntactic features (e.g., category) of incoming words. At the neural level, basic syntactic composition relies on a syntactic network composed of BA44 and the PTL, and on the functional interaction between these two regions.

are not compatible, linguistic processing is disrupted. Similarly, E/MEG studies showed that categorical and agreement violations elicit syntactic negativities (e.g., Barber and Carreiras, 2003; Hasting and Kotz, 2008; Pulvermüller and Shtyrov, 2003), whose early latency is compatible with a fast and efficient application of grammatical rules. The rule-based nature of syntactic composition is further supported by two sMMN studies (Herrmann et al., 2009; Pulvermüller and Assadollahi, 2007), comparing grammatical structures of different frequency and ungrammatical ones. In particular, the early stage of syntactic analysis reflects a binary decision: either an utterance follows the rules of a given grammar or it does not, independently from its frequency of occurrence in natural language. This claim is further reinforced by the observation of syntactic priming effects (e.g., Gurjanov et al., 1985; Katz et al., 1987; Lukatela et al., 1982, 1983) and negativities (Münte and Heinze, 1994) when grammatical rules are violated in two-word constructions with pseudo-words, which by definition do not occur in natural language. Converging evidence comes from studies which have compared constituents and lists built with pseudo-words (Hardy et al., 2022; Poulisse et al., 2019; Segaert et al., 2018; Zaccarella and Friederici, 2015b).

The rule-based combinatorial process strongly relies on the purely syntactic nature of function words (Barber and Carreiras, 2005; Colé and Segui, 1994; Schell et al., 2017) and inflectional morphemes (Lukatela et al., 1982, 1983). Indeed, a recent study showed that different function words seem to initiate specific pathways to extract the categorical features of upcoming words (Pyatigorskaya et al., 2021), on which the combinatorial system can operate. Interestingly, when both function words and inflectional particles are missing in a two-word construction (e.g., "red boat"), a different compositional system seems to be engaged, operating on the conceptual features of the two content words (Lukatela, Carello et al., 1987; Pylkkänen, 2020). This might reflect an early recognition of the impossibility of building a grammatical structure starting from an *adjective*$_{[\text{SINGULAR}]}$ and a *noun*$_{[\text{SINGULAR}]}$ ("*Red boat sails down the river"), which would require a function word preceding them ("A red boat sails down the river") or inflectional plural marking present (e.g., "Red boats sail down the river").

3.2. *Early syntactic analysis occurs automatically, followed by task-dependent processes*

Both behavioral and neurophysiological data at the two-word level support an automatic nature of syntactic analysis. In particular, syntactic priming effects have been observed when SOA manipulations reduced the time available for strategic processes (Colé and Segui, 1994; Katz et al., 1987; Lukatela et al., 1982) and, at least at the categorical level, with subliminal presentation of the prime (Berkovitch and Dehaene, 2019; Pyatigorskaya et al., 2021). Notably, in this paradigm participants seem unable to avoid analysing the grammatical relationship between prime and target, despite not being necessary to provide a correct answer (Faussart et al., 1999; Seidenberg et al., 1984). Converging evidence for automaticity in syntactic analysis comes from the observation of the ESN during visual distraction conditions (Hasting and Kotz, 2008) and by several sMMN studies (Hanna et al., 2014; Hasting et al., 2007; Herrmann et al., 2009; Pulvermüller et al., 2008; Pulvermüller and Assadollahi, 2007; Shtyrov et al., 2003). Both ERP components are characterized by early latency onsets, in line with a fast and efficient application of syntactic analysis.

This first and automatic step of analysis is followed by late controlled processes. At the behavioral level, they might contribute to the larger syntactic priming effects which are observed when the prime is not masked (Berkovitch and Dehaene, 2019; Pyatigorskaya et al., 2021). At the neural level, a late positivity (P600) follows the early syntactic negativities when participants are performing a task or are not actively distracted (Barber and Carreiras, 2005; Hasting and Kotz, 2008; Jakuszeit et al., 2013; Maran et al., 2021). Interestingly, differences between types of agreement violations (i.e., gender vs number) have been described in this time-window (Barber and Carreiras, 2003, 2005), in line with the notion of a longer re-analysis step when the violated syntactic feature (i.e., gender) is expressed by the stem (Faussart et al., 1999).

Overall, the reviewed studies support a two-fold model of syntactic analysis. A first step, occurring in an automatic fashion, provides a binary outcome: either something is grammatical or it is not (Herrmann et al., 2009; Pulvermüller and Assadollahi, 2007), independently of frequency effects. This diagnostic process is characterised by a fast and efficient analysis of information, as reflected in the early latencies of the ESN (Hasting and Kotz, 2008) and sMMN (Hasting et al., 2007; Herrmann et al., 2009; Pulvermüller and Assadollahi, 2007) components. As a second step, repairing and re-analysis processes can be subsequently engaged in a task-dependent manner, reflected in the P600 component. Note that in this time-window fine-grained differences along a continuum might emerge within ungrammatical structures, possibly driven by the depth of the re-analysis and repair processes invoked (Barber and Carreiras, 2005; Faussart et al., 1999).

3.3. The left IFG and posterior temporal lobe are key regions of the syntactic combinatorial system

Data from all the reviewed paradigm (syntactic violations, phrases/sentences versus lists, comparison of nouns and verbs) support a central role of Broca's area and the posterior temporal lobe in analysing and combining grammatical information in two-word constituents. Several fMRI and MEG studies showed an involvement of Broca's area in detecting an incompatibility between the syntactic features of an incoming word and the preceding one, both at the categorical level (Herrmann et al., 2012) and agreement level (Carreiras et al., 2010; Hanna et al., 2014; Heim et al., 2010). Additional activations vary according to the type of violation, as categorical errors engage the bilateral STS (Herrmann et al., 2012), while the premotor cortex and the parietal lobe are activated by agreement errors (Carreiras et al., 2010; Heim et al., 2010). Even in the absence of violations, an involvement of the left IFG (Hardy et al., 2022; Schell et al., 2017; Wu et al., 2019; Zaccarella and Friederici, 2015b) and PTL (Matar et al., 2021; Wu et al., 2019) in syntactic composition is supported by the phrase-versus-list comparisons, with stronger activations and connectivity between these regions driven by hierarchical syntactic processing. Importantly, the involvement of these two regions in syntactic composition has been observed both in the auditory (Hardy et al., 2022; Heim et al., 2010; Herrmann et al., 2012; Schell et al., 2017) and visual (Carreiras et al., 2010; Heim et al., 2010; Matar et al., 2021; Wu et al., 2019; Zaccarella and Friederici, 2015b) modality, pointing towards modality-independent operations in these regions. This is further supported by recent studies showing that the left IFG (Trettenbrein et al., 2020) and the posterior temporal lobe (Matchin et al., 2022) are central regions in processing signed languages, in line with the abstract nature of their linguistic operations. Similarly, recent evidence from artificial grammar paradigms supports the involvement of the left IFG and PTL in abstract categorical processes (Chen et al., 2019, 2021), putting forward the working hypothesis that dissociations between the left IFG and PTL in syntactic composition can be traced—the left IFG being specialized in building up hierarchies on the basis of categorical features, the PTL integrating hierarchies with other sources of linguistic information, including meaning. Under this account, the information exchange between the left IFG and PTL would allow to reconstruct the hierarchical dependencies characterizing human language, providing an analysis that will also interface with the semantic system.

Interestingly, the same regions activated by the detection of syntactic violations and hierarchical processing (i.e., left IFG and pMTG/STG) seem to be involved in processing differences between constituents formed with nouns and verbs (Burton et al., 2009; Pulvermüller et al., 2012; Strijkers et al., 2019; Tsigka et al., 2014; Tyler et al., 2008). An interesting hypothesis is that the effects observed in the left fronto-temporal syntactic network reflect the different load on the combinatorial mechanism when building constituents with nouns or verbs, rather than categorical effects per se. While, at least in languages with a rich inflectional system, both nouns and verbs might engage to a similar extent morphological analysis, only the latter category requires the analysis of tense features (which can be represented as an additional node, *Tense Phrase*, in the derivation) and assigns thematic roles. Indeed, two studies have observed categorical effects in the left IFG (Strijkers et al., 2019) or the pMTG (Tyler et al., 2008) when two-word constructions are employed, but not for nouns, verbs or homonyms presented in isolation. Further studies are needed to shed light on this hypothesis, in line with the theoretical (Vigliocco et al., 2011) and methodological (Crepaldi et al., 2011, 2013) considerations previously pointed out.

Overall, the reviewed studies support the notion of a neuroanatomical dissociation between the syntactic and semantic combinatorial processes when building basic two-word structures. In particular, the posterior portion of the left IFG (BA44) and temporal lobe combine abstract syntactic information into constituents, while the ATL, vmPFC (Pylkkänen, 2020), angular gyrus and more anterior portion of the left IFG (Graessner et al., 2021a, 2021b) are involved in semantic and conceptual composition. Initial evidence exists also for a modality-independent nature of this dissociation (Blanco-Elorrieta et al., 2018; Matchin et al., 2022; Moreno et al., 2018), which should be addressed in future studies.

3.4. Syntactic composition as an efficient bottom-up process

In recent years interest has grown in understanding the contribution of top-down prediction and bottom-up integration in incremental syntactic composition (e.g., left-corner, bottom-up and top-down parsing, see Abney and Johnson, 1991; Resnik, 1992). In particular, several studies have correlated metrics based on computational linguistics parsers with neural signals recorded while participants listen to narratives, looking for a convergence between parsing operations and brain functioning (Bhattasali et al., 2019; Brennan et al., 2016, 2020; Brennan and Hale, 2019; Brennan and Martin, 2020; Brennan and Pylkkänen, 2017; Hale et al., 2018). A limitation of this approach is that syntactic and semantic composition become difficult to disentangle, as each incoming word of the narratives builds not only grammatical information but also semantic one (Bhattasali et al., 2019). Two-word constructions might give complementary insights, as they allow to isolate predictive and integration processes, occurring respectively during the first and second word of the utterance.

At present, predictive effects in two-word constructions have been observed in a rather limited number of studies (Hardy et al., 2022; Matar et al., 2021; Segaert et al., 2018), sometimes reporting conflicting findings. Conversely, the reviewed studies provide initial evidence for fast and efficient bottom-up operations in basic syntactic composition. At the behavioral level, syntactic priming effects have been related to a post-lexical stage of processing (Carello et al., 1988; Seidenberg et al., 1984), rather than pre-activation of lexical entries, which would be inefficient as a very large cohort of candidates (e.g., all the nouns and adjectives) would be activated following a function word (e.g., a determiner). This suggests the existence of operations efficiently integrating the syntactic features of incoming words into the preceding context, in line with the notion of bottom-up parsing (Hale, 2014). Converging evidence comes from the inhibitory nature of syntactic priming (Friederici and Jacobsen, 1999; Pyatigorskaya et al., 2021). If top-down predictions were at work, faster RTs should be observed for a grammatical two-word phrase compared to a matched baseline. This however is not the case, and syntactic priming rather stems from longer RTs of the ungrammatical condition compared to a neutral baseline (Friederici and Jacobsen, 1999; Pyatigorskaya et al., 2021). In this regard, some sMMN studies have shown an opposite pattern, with the response to a grammatical construction differing from the ones to words presented in isolation (Pulvermüller and Shtyrov, 2003, 2006), possibly reflecting facilitation process. A possible explanation is that, since only two constructions are used in sMMN paradigms, function words could be used to predict a specific lexical item, therefore facilitation can be observed. In paradigms which use a larger number of lexical items (e.g., syntactic priming) that might not be the case, and only reliable inhibition for ungrammatical structures is observed. As discussed above and previously in the literature (Seidenberg et al., 1984), the syntactic features of a given word are compatible with a large number of candidates in the lexicon that would be inefficient to pre-activate. Indeed, even if many candidates (e.g., all the nouns and adjectives) were pre-activated, mechanisms such as lateral inhibition and lexical competition would remove any facilitatory effect. This might constitute a key difference with the semantic system, for which top-down prediction might be more efficient, given that few specific lexical items can be pre-activated, therefore giving rise to facilitation compared to neutral baselines (Lukatela, Carello et al., 1987).

Neuroimaging studies employing syntactic violations in which grammaticality is orthogonal to the identity of the first word and the

syntactic features of the second one (Carreiras et al., 2010; Hasting and Kotz, 2008; Heim et al., 2010; Herrmann et al., 2012; Maran et al., 2021) allow to highlight the neural correlates of bottom-up integration, as by definition any virtual predictive processes is subtracted out.[14] These studies have shown the presence of syntactic negativities (Barber and Carreiras, 2005; Hasting and Kotz, 2008; Maran et al., 2021), suggesting the fast analysis of two words' syntactic feature compatibility. In parallel, they have provided evidence for a key role of Broca's area in this process, both with respect to agreement (Carreiras et al., 2010; Heim et al., 2010) and categorical (Herrmann et al., 2012) features. Converging evidence comes also from two TMS studies, showing that disruption of Broca's area during the integration stage (i.e., the second word of a two-word construction) affects syntactic processing (Carreiras et al., 2012), while stimulation of this region at the predictive phase does not interfere with it (Maran et al., 2021). A role of Broca's area in bottom-up syntactic composition is also supported by a phrase-versus-list study, which included a one-word condition controlling for potential predictive effects (Zaccarella and Friederici, 2015b). Finally, stronger interactions between the left IFG and the temporal lobe during the second word of a two-word constituent (Hardy et al., 2022) might represent an additional neural trace of syntactic integration within the language network.

Overall, the reviewed studies support a crucial role of bottom-up processes in integrating the syntactic features of two words into a constituent. This reliance on basic bottom-up operations, with a limited role of top-down predictions, might represent a critical distinction between the syntactic combinatorial system and the semantic one, possibly grounded in differences at the neuro-anatomical level (Graessner et al., 2021a, 2021b; Pylkkänen, 2020; Zaccarella et al., 2017b; Zaccarella and Friederici, 2015b). Similarly, this aspect might differentiate syntactic composition, characterised by hierarchical processing, from non-linguistic domains which might strongly rely on serial top-down predictions (Zaccarella et al., 2021).

4. Conclusions and future directions

Two-word studies help characterizing key aspects of the neurocognitive basis of syntactic composition. The automatic extraction and integration of syntactic features into constituents emerged as a central aspect of linguistic composition. At the neural level, these operations appear to be supported by Broca's area, the PTL and by functional interactions between these two brain regions.

The reviewed studies allowed also to formulate important methodological considerations. The use of orthogonal designs, cross-splicing techniques and difference waves for ERP studies might allow addressing important methodological issues, previously pointed out in syntactic studies (Steinhauer and Drury, 2012). While at present basic syntactic composition has not been tested in signed languages, (but see Blanco-Elorrieta et al., 2018, for an application of the "red boat" paradigm), recent methodological advances and norms' development allow to test this process directly, while achieving great control over perceptual and psycholinguistic variables of interest (Bungeroth et al., 2008; Caselli et al., 2017; Sehyr et al., 2021; Trettenbrein et al., 2021; Trettenbrein and Zaccarella, 2021).

Based on the evidence reviewed in this article, methodological advances now make it possible to progress towards a mechanistic understanding of syntactic composition at the neural and cognitive levels (Fig. 6). Fundamental research questions for future studies include how the information flows within the syntactic network, how labels are applied to recursively build hierarchical structures, and which are the key phylogenetic and ontogenetic changes which made the human brain an efficient integrator of syntactic features. As outlined in Section 2.3, an additional area which awaits further studies relates to the neural basis of noun and verb categorical representation. We thus believe that a coherent research program that integrates hypotheses from linguistic theory can help reveal how exactly composition occurs at this fundamental level, thus providing the basis for understanding the uniquely human ability to use language.

[14] This is ensured by the presence of the same first words both in grammatical and ungrammatical structures (e.g., "Ein Falter", *a butterfly*, "Er faltet", *he folds*, "*Ein faltet", **a folds*, "*Er Falter", **he butterfly*). In other words, the same syntactic features would be in principle predicted both in grammatical and ungrammatical structures. See Hasting & Kotz (2007) for examples both at the agreement and categorical level.

Funding

Matteo Maran was supported by direct funding from the Department of Neuropsychology (Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany).

Acknowledgments

The authors would like to thank Patrick Trettenbrein and Elena Pyatigorskaya for the valuable discussions on the present work.

References

Abney, S.P., Johnson, M., 1991. Memory requirements and local ambiguities of parsing strategies. J. Psycholinguist. Res. 20 (3), 233–250. https://doi.org/10.1007/BF01067217.

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

Ansorge, U., Reynvoet, B., Hendler, J., Oettl, L., Evert, S., 2013. Conditional automaticity in subliminal morphosyntactic priming. Psychol. Res. 77 (4), 399–421. https://doi.org/10.1007/s00426-012-0442-z.

Arcara, G., Franzon, F., Gastaldon, S., Brotto, S., Semenza, C., Peressotti, F., Zanini, C., 2019. One can be some but some cannot be one: ERP correlates of numerosity incongruence are different for singular and plural. Cortex 116, 104–121. https://doi.org/10.1016/j.cortex.2018.10.022.

Arcara, G., Pezzetta, R., Benavides-Varela, S., Rizzi, G., Formica, S., Turco, C., Piccione, F., Semenza, C., 2021. Magnetoencephalography reveals differences in brain activations for fast and slow responses to simple multiplications. Sci. Rep. 11 (1), 1–13. https://doi.org/10.1038/s41598-021-97927-8.

Baillet, S., 2017. Magnetoencephalography for brain electrophysiology and imaging. Nat. Neurosci. 20, 327–339. https://doi.org/10.1038/nn.4504.

Barber, H., Carreiras, M., 2003. Integrating gender and number information in Spanish word pairs: an ERP study. Cortex 39 (3), 465–482. https://doi.org/10.1016/S0010-9452(08)70259-4.

Barber, H., Carreiras, M., 2005. Grammatical gender and number agreement in Spanish: an ERP comparison. J. Cogn. Neurosci. 17 (1), 137–153. https://doi.org/10.1162/0898929052880101.

Bates, E., Devescovi, A., Hernandez, A., Pizzamiglio, L., 1996. Gender priming in Italian. Percept. Psychophys. 58 (7), 992–1004. https://doi.org/10.3758/BF03206827.

Bemis, D.K., Pylkkänen, L., 2011. Simple composition: a magnetoencephalography investigation into the comprehension of minimal linguistic phrases. J. Neurosci. 31 (8), 2801–2814. https://doi.org/10.1523/JNEUROSCI.5003-10.2011.

Benítez-Burraco, A., Murphy, E., 2019. Why brain oscillations are improving our understanding of language. Front. Behav. Neurosci. 13, 1–10. https://doi.org/10.3389/fnbeh.2019.00190.

Berkovitch, L., Dehaene, S., 2019. Subliminal syntactic priming. Cogn. Psychol. 109, 26–46. https://doi.org/10.1016/j.cogpsych.2018.12.001.

Berwick, R.C., Friederici, A.D., Chomsky, N., Bolhuis, J.J., 2013. Evolution, brain, and the nature of language. Trends Cogn. Sci. 17 (2), 98. https://doi.org/10.1016/j.tics.2012.12.002.

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. Lang. Cogn. Neurosci. 34 (4), 491–510. https://doi.org/10.1080/23273798.2018.1518533.

Blanco-Elorrieta, E., Kastner, I., Emmorey, K., Pylkkänen, L., 2018. Shared neural correlates for building phrases in signed and spoken language. Sci. Rep. 8 (1), 1–10. https://doi.org/10.1038/s41598-018-23915-0.

Bock, J.K., 1986. Syntactic persistence in language production. Cogn. Psychol. 18, 355–387, 10.1016/0010-0285(86)90004-6.

Bozic, M., Fonteneau, E., Su, L., Marslen-Wilson, W.D., 2015. Grammatical analysis as a distributed neurobiological function. Hum. Brain Mapp. 36 (3), 1190–1201. https://doi.org/10.1002/hbm.22696.

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

Brennan, J.R., Martin, A.E., 2020. Phase synchronization varies systematically with linguistic structure composition. Philos. Trans. R. Soc. B: Biol. Sci. 375 (1791), 20190305. https://doi.org/10.1098/rstb.2019.0305.
Brennan, J.R., Pylkkänen, L., 2017. MEG Evidence for incremental sentence composition in the anterior temporal lobe. Cogn. Sci. 41, 1515–1531. https://doi.org/10.1111/cogs.12445.
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 Lang. 157–158, 81–94. https://doi.org/10.1016/j.bandl.2016.04.008.
Brennan, J.R., Dyer, C., Kuncoro, A., Hale, J.T., 2020. Localizing syntactic predictions using recurrent neural network grammars. Neuropsychologia 146, 107479. https://doi.org/10.1016/j.neuropsychologia.2020.107479.
Brunellière, A., Franck, J., Ludwig, C., Frauenfelder, U.H., 2007. Early and automatic syntactic processing of person agreement. NeuroReport 18 (6), 537–541. https://doi.org/10.1097/WNR.0b013e3280b07ba1.
Bungeroth, J., Stein, D., Dreuw, P., Ney, H., Morrissey, S., Way, A., & Van Zijl, L. (2008). The ATIS sign language corpus. Proceedings of the 6th International Conference on Language Resources and Evaluation, LREC 2008, 2943–2946.
Burton, M.W., Krebs-Noble, D., Gullapalli, R., Berndt, R.S., 2009. Functional neuroimaging of grammatical class: ambiguous and unambiguous nouns and verbs. Cogn. Neuropsychol. 26 (2), 148–171. https://doi.org/10.1080/02643290802536090.
Carello, C., Lukatela, G., Turvey, M.T., 1988. Rapid naming is affected by association but not by syntax. Mem. Cogn. 16 (3), 187–195. https://doi.org/10.3758/BF03197751.
Carreiras, M., Carr, L., Barber, H.A., Hernandez, A., 2010. Where syntax meets math: right intraparietal sulcus activation in response to grammatical number agreement violations. NeuroImage 49 (2), 1741–1749. https://doi.org/10.1016/j.neuroimage.2009.09.058.
Carreiras, M., Pattamadilok, C., Meseguer, E., Barber, H., Devlin, J.T., 2012. Broca's area plays a causal role in morphosyntactic processing. Neuropsychologia 50 (5), 816–820. https://doi.org/10.1016/j.neuropsychologia.2012.01.016.
Caselli, N.K., Sehyr, Z.S., Cohen-Goldberg, A.M., Emmorey, K., 2017. ASL-LEX: a lexical database of American sign language. Behav. Res. Methods 49 (2), 784–801. https://doi.org/10.3758/s13428-016-0742-0.
Chen, L., Wu, J., Fu, Y., Kang, H., Feng, L., 2019. Neural substrates of word category information as the basis of syntactic processing. Hum. Brain Mapp. 40 (2), 451–464. https://doi.org/10.1002/hbm.24386.
Chen, L., Goucha, T., Männel, C., Friederici, A.D., Zaccarella, E., 2021. Hierarchical syntactic processing is beyond mere associating: Functional magnetic resonance imaging evidence from a novel artificial grammar. *Hum. Brain Mapp.* 25432. https://doi.org/10.1002/hbm.25432.
Chomsky, N. (1995). The Minimalist Program. Cambridge MA: MIT Press.
Colé, P., Segui, J., 1994. Grammatical incongruency and vocabulary types. Mem. Cogn. 22 (4), 387–394. https://doi.org/10.3758/BF03200865.
Crepaldi, D., Berlingeri, M., Paulesu, E., Luzzatti, C., 2011. A place for nouns and a place for verbs? A critical review of neurocognitive data on grammatical-class effects. Brain Lang. 116 (1), 33–49. https://doi.org/10.1016/j.bandl.2010.09.005.
Crepaldi, D., Berlingeri, M., Cattinelli, I., Borghese, N.A., Luzzatti, C., Paulesu, E., 2013. Clustering the lexicon in the brain: a meta-analysis of the neurofunctional evidence on noun and verb processing. Front. Hum. Neurosci. 7, 1–15. https://doi.org/10.3389/fnhum.2013.00303.
Da Silva, F.L., 2013. EEG and MEG: relevance to neuroscience. Neuron 80 (5), 1112–1128. https://doi.org/10.1016/j.neuron.2013.10.017.
Dehaene, S., Piazza, M., Pinel, P., Cohen, L., 2003. Three parietal circuits for number processing. Cogn. Neuropsychol. 20 (3–6), 487–506. https://doi.org/10.1080/02643290244000239.
Everaert, M.B.H., Huybregts, M.A.C., Chomsky, N., Berwick, R.C., Bolhuis, J.J., 2015. Structures, not strings: linguistics as part of the cognitive sciences. Trends Cogn. Sci. 19 (12), 729–743. https://doi.org/10.1016/j.tics.2015.09.008.
Faussart, C., Jakubowicz, C., Costes, M., 1999. Gender and number processing in spoken French and Spanish. Ital. J. Linguist. 11 (1), 75–102.
Fedorenko, E., Scott, T.L., Brunner, P., Coon, W.G., Pritchett, B., Schalk, G., Kanwisher, N., 2016. Neural correlate of the construction of sentence meaning. Proc. Natl. Acad. Sci. U.S.A. 113 (41), E6256–E6262. https://doi.org/10.1073/pnas.1612132113.
Fedorenko, E., Blank, I.A., Siegelman, M., Mineroff, Z., 2020. Lack of selectivity for syntax relative to word meanings throughout the language network. Cognition 203. https://doi.org/10.1016/j.cognition.2020.104348.
Fiebach, C.J., Maess, B., & Friederici, A.D. (2002). Neuromagnetic evidence that differences in noun and verb processing are modulated by the presence of a syntactic context. Proceedings of the 13th International Conference on Biomagnetism, 5, 339–341.
Fló, E., Cabana, Á., Valle-Lisboa, J.C., 2020. EEG signatures of elementary composition: Disentangling genuine composition and expectancy processes. Brain Lang. 209, 104837 https://doi.org/10.1016/j.bandl.2020.104837.
Forster, K.I. (1979). Levels of processing and the structure of the language processor. In W. E. Cooper & E. C. T. Walker (Eds.), Sentence Processing: Psycholinguistic Studies Presented to Merrill Garrett (pp. 27–85). Lawrence Erlbaum Associates, Publishers.
Forster, K.I., 1981. Priming and the effects of sentence and lexical contexts on naming time: evidence for autonomous lexical processing. Q. J. Exp. Psychol. Sect. A 33 (4), 465–495. https://doi.org/10.1080/14640748108400804.
Friederici, A.D., 2011. The brain basis of language processing: from structure to function. Physiol. Rev. 91 (4), 1357–1392. https://doi.org/10.1152/physrev.00006.2011.
Friederici, A.D., Jacobsen, T., 1999. Processing grammatical gender during language comprehension. J. Psycholinguist. Res. 28 (5), 467–484. https://doi.org/10.1023/A:1023264209610.
Friederici, A.D., Singer, W., 2015. Grounding language processing on basic neurophysiological principles. Trends Cogn. Sci. 19 (6), 329–338. https://doi.org/10.1016/j.tics.2015.03.012.
Friederici, A.D., Pfeifer, E., Hahne, A., 1993. Event-related brain potentials during natural speech processing: effects of semantic, morphological and syntactic violations. Cogn. Brain Res. 1 (3), 183–192. https://doi.org/10.1016/0926-6410(93)90026-2.
Friederici, A.D., Von Cramon, D.Y., Kotz, S.A., 1999. Language related brain potentials in patients with cortical and subcortical left hemisphere lesions. Brain 122 (6), 1033–1047. https://doi.org/10.1093/brain/122.6.1033.
Friederici, A.D., Chomsky, N., Berwick, R.C., Moro, A., Bolhuis, J.J., 2017. Language, mind and brain. Nat. Hum. Behav. 1 https://doi.org/10.1038/s41562-017-0184-4.
Fries, P., 2015. Rhythms for cognition: communication through coherence. Neuron 88 (1), 220–235. https://doi.org/10.1016/j.neuron.2015.09.034.
Goldenholz, D.M., Ahlfors, S.P., Hämäläinen, M.S., Sharon, D., Ishitobi, M., Vaina, L.M., Stufflebeam, S.M., 2009. Mapping the signal-to-noise-ratios of cortical sources in magnetoencephalography and electroencephalography. Hum. Brain Mapp. 30 (4), 1077–1086. https://doi.org/10.1002/hbm.20571.
Goodman, G.O., McClelland, J.L., Gibbs, R.W., 1981. The role of syntactic context in word recognition. Mem. Cogn. 9 (6), 580–586. https://doi.org/10.3758/BF03202352.
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.
Graessner, A., Zaccarella, E., Hartwigsen, G., 2021a. Differential contributions of left-hemispheric language regions to basic semantic composition. Brain Struct. Funct. 226, 501–518. https://doi.org/10.1007/s00429-020-02196-2, 0123456789.
Graessner, A., Zaccarella, E., Friederici, A.D., Obrig, H., Hartwigsen, G., 2021b. Dissociable contributions of frontal and temporal brain regions to basic semantic composition. Brain Commun. 3 (2) https://doi.org/10.1093/braincomms/fcab090.
Gurjanov, M., Lukatela, G., Moskovljević, J., Savić, M., Turvey, M.T., 1985. Grammatical priming of inflected nouns by inflected adjectives. Cognition 19 (1), 55–71. https://doi.org/10.1016/0010-0277(85)90031-9.
Gurjanov, M., Lukatela, G., Lukatela, K., Savić, M., Turvey, M.T., 1985. Grammatical priming of inflected nouns by the gender of possessive adjectives. J. Exp. Psychol.: Learn., Mem., Cogn. 11 (4), 692–701. https://doi.org/10.1037/0278-7393.11.1-4.692.
Gwilliams, L., 2020. How the brain composes morphemes into meaning. Philos. Trans. R. Soc. B: Biol. Sci. 375 (1791), 20190311. https://doi.org/10.1098/rstb.2019.0311.
Hagoort, P., 2016. MUC (Memory, Unification,Control): A Model on the Neurobiology of Language Beyond Single Word Processing. In: Hickok, G., Small, S.L. (Eds.), Neurobiology of Language. Academic Press, pp. 339–347. https://doi.org/10.1016/B978-0-12-407794-2.00028-6.
Hagoort, P., Indefrey, P., 2014. The neurobiology of language beyond single words. Annu. Rev. Neurosci. 37, 347–362. https://doi.org/10.1146/annurev-neuro-071013-013847.
Hahne, A., Friederici, A.D., 1999. Electrophysiological evidence for two steps in syntactic analysis. Early automatic and late controlled processes. J. Cogn. Neurosci. 11 (2), 194–205. https://doi.org/10.1162/089892999563328.
Hale, J.T. (2014). Automaton Theories of Human Sentence Comprehension. CSLI Publications.
Hale, J., Dyer, C., Kuncoro, A., & Brennan, J. R. 2018. Finding Syntax in Human Encephalography with Beam Search. *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, Melbourne, Australia*, 2727–2736. https://doi.org/10.48550/arXiv.1806.04127.
Hanna, J., Mejias, S., Schelstraete, M.A., Pulvermüller, F., Shtyrov, Y., van der Lely, H.K. J., 2014. Early activation of Broca's area in grammar processing as revealed by the syntactic mismatch negativity and distributed source analysis. Cogn. Neurosci. 5 (2), 66–76. https://doi.org/10.1080/17588928.2013.860087.
Hanna, J., Shtyrov, Y., Williams, J., Pulvermüller, F., 2016. Early neurophysiological indices of second language morphosyntax learning. Neuropsychologia 82, 18–30. https://doi.org/10.1016/j.neuropsychologia.2016.01.001.
Hardy, S.M., Jensen, O., Wheeldon, L., Mazaheri, A., Segaert, K., 2022. Modulation in alpha band activity reflects syntax composition: an MEG study of minimal syntactic binding. Cereb. Cortex 1–15. https://doi.org/10.1093/cercor/bhac080.
Hartwigsen, G., 2015. The neurophysiology of language: Insights from non-invasive brain stimulation in the healthy human brain. Brain Lang. 148, 81–94. https://doi.org/10.1016/j.bandl.2014.10.007.
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. J. Cogn. Neurosci. 20 (7), 1207–1219. https://doi.org/10.1162/jocn.2008.20083.
Hasting, A.S., Kotz, S.A., Friederici, A.D., 2007. Setting the stage for automatic syntax processing: the mismatch negativity as an indicator of syntactic priming. J. Cogn. Neurosci. 19 (3), 386–400. https://doi.org/10.1162/jocn.2007.19.3.386.
Heim, S., Van Ermingen, M., Huber, W., Amunts, K., 2010. Left cytoarchitectonic BA 44 processes syntactic gender violations in determiner phrases. Hum. Brain Mapp. 31 (10), 1532–1541. https://doi.org/10.1002/hbm.20957.
Herrmann, B., Maess, B., Hasting, A.S., Friederici, A.D., 2009. Localization of the syntactic mismatch negativity in the temporal cortex: an MEG study. NeuroImage 48 (3), 590–600. https://doi.org/10.1016/j.neuroimage.2009.06.082.

Herrmann, B., Obleser, J., Kalberlah, C., Haynes, J.-D., Friederici, A.D., 2012. Dissociable neural imprints of perception and grammar in auditory functional imaging. Hum. Brain Mapp. 33 (3), 584–595. https://doi.org/10.1002/hbm.21235.

Hultén, A., Schoffelen, J.M., Uddén, J., Lam, N.H.L., Hagoort, P., 2019. How the brain makes sense beyond the processing of single words – an MEG study. NeuroImage 186, 586–594. https://doi.org/10.1016/j.neuroimage.2018.11.035.

Iwabuchi, T., Nakajima, Y., Makuuchi, M., 2019. Neuralarchitecture of human language: Hierarchical structure building is independent from working memory. Neuropsychologia 132 (July), 107137.

Jakuszeit, M., Kotz, S.A., Hasting, A.S., 2013. Generating predictions: lesion evidence on the role of left inferior frontal cortex in rapid syntactic analysis. Cortex 49 (10), 2861–2874. https://doi.org/10.1016/j.cortex.2013.05.014.

Kang, A.M., Constable, R.T., Gore, J.C., Avrutin, S., 1999. An event-related fMRI study of implicit phrase-level syntactic and semantic processing. NeuroImage 10, 555–561. https://doi.org/10.1006/nimg.1999.0493.

Katz, L., Boyce, S., Goldstein, L., Lukatela, G., 1987. Grammatical information effects in auditory word recognition. Cognition 25 (3), 235–263. https://doi.org/10.1016/S0010-0277(87)80005-7.

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

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.

Kutas, M., Federmeier, K.D., 2011. Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). Annu. Rev. Psychol. 62, 621–647. https://doi.org/10.1146/annurev.psych.093008.131123.

Lam, N.H.L., Schoffelen, J.M., Uddén, J., Hultén, A., Hagoort, P., 2016. Neural activity during sentence processing as reflected in theta, alpha, beta, and gamma oscillations. NeuroImage 142, 43–54. https://doi.org/10.1016/j.neuroimage.2016.03.007.

Lau, E., Stroud, C., Plesch, S., Phillips, C., 2006. The role of structural prediction in rapid syntactic analysis. Brain Lang. 98 (1), 74–88. https://doi.org/10.1016/j.bandl.2006.02.003.

Lee, C.L., Federmeier, K.D., 2006. To mind the mind: An event-related potential study of word class and semantic ambiguity. Brain Res. 1081 (1), 191–202. https://doi.org/10.1016/j.brainres.2006.01.058.

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

Lucchese, G., Pulvermüller, F., Stahl, B., Dreyer, F.R., Mohr, B., 2017a. Therapy-induced neuroplasticity of language in chronic post stroke aphasia: a mismatch negativity study of (A)grammatical and meaningful/less mini-constructions. Front. Hum. Neurosci. 10, 1–14. https://doi.org/10.3389/fnhum.2016.00669.

Lucchese, G., Hanna, J., Autenrieb, A., Miller, T.M., Pulvermüller, F., 2017b. Electrophysiological evidence for early and interactive symbol access and rule processing in retrieving and combining language constructions. J. Cogn. Neurosci. 29 (2), 254–266. https://doi.org/10.1162/jocn_a_01038.

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

Lukatela, G., Moraca, J., Stojnov, D., Savic, M.D., Katz, L., Turvey, M.T., 1982. Grammatical priming effects between pronouns and inflected verb forms. Psychol. Res. 44 (4), 297–311. https://doi.org/10.1007/BF00309326.

Lukatela, G., Kostić, A., Feldman, L.B., Turvey, M.T., 1983. Grammatical priming of inflected nouns. Mem. Cogn. 11 (1), 59–63. https://doi.org/10.3758/BF03197662.

Lukatela, G., Carello, C., Kostic, A., Turvey, M.T., 1987. Low constraint facilitation in lexical decision with single word contexts. Am. J. Psychol. 90, 93–104.

Lukatela, G., Kostić, A., Todorović, D., Carello, C., Turvey, M.T., 1987. Type and number of violations and the grammatical congruency effect in lexical decision. Psychol. Res. 49 (1), 37–43. https://doi.org/10.1007/BF00309201.

Maess, B., Schröger, E., Widmann, A., 2016. High-pass filters and baseline correction in M/EEG analysis. Commentary on: "How inappropriate high-pass filters can produce artefacts and incorrect conclusions in ERP studies of language and cognition. J. Neurosci. Methods 266, 164–165. https://doi.org/10.1016/j.jneumeth.2015.12.003.

Makuuchi, M., Bahlmann, J., Anwander, A., Friederici, A.D., 2009. Segregatingthe core computational faculty of human language from working memory. Proceedings of the National Academy of Sciences of the United States of America 106 (20), 8362–8367.

Maran, M., Numssen, O., Hartwigsen, G., Friederici, A.D., & Zaccarella, E. (2021). Towards a causal role of Broca's area in language: A TMS-EEG study on syntactic prediction. *BioRxiv*, 2021.04.14.439631. https://doi.org/10.1101/2021.04.14.439631.

Marslen-Wilson, W.D., Tyler, L.K., 1980. The temporal structure of spoken language understanding. Cognition 8 (1), 1–71. https://doi.org/10.1016/0010-0277(80)90015-3.

Martin, A.E., Baggio, G., 2020. Modelling meaning composition from formalism to mechanism. Philos. Trans. R. Soc. B: Biol. Sci. 375 (1791), 1–7. https://doi.org/10.1098/rstb.2019.0298.

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

Matchin, W., Hickok, G., 2020. The cortical organization of syntax. Cereb. Cortex 30 (3), 1481–1498. https://doi.org/10.1093/cercor/bhz180.

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., İlkbaşaran, D., Hatrak, M., Roth, A., Villwock, A., Halgren, E., Mayberry, R. I., 2022. The cortical organization of syntactic processing is supramodal: evidence from American sign language. J. Cogn. Neurosci. 34 (2), 224–235. https://doi.org/10.1162/jocn_a_01790.

Meyer, L., 2018. The neural oscillations of speech processing and language comprehension: state of the art and emerging mechanisms. Eur. J. Neurosci. 48 (7), 2609–2621. https://doi.org/10.1111/ejn.13748.

Molinaro, N., Barber, H.A., Carreiras, M., 2011. Grammatical agreement processing in reading: ERP findings and future directions. Cortex 47 (8), 908–930. https://doi.org/10.1016/j.cortex.2011.02.019.

Molinaro, N., Barber, H.A., Caffarra, S., Carreiras, M., 2015. On the left anterior negativity (LAN): The case of morphosyntactic agreement: a reply to Tanner etal. Cortex 66, 156–159. https://doi.org/10.1016/j.cortex.2014.06.009.

Moreno, A., Limousin, F., Dehaene, S., Pallier, C., 2018. Brain correlates of constituent structure in sign language comprehension. NeuroImage 167, 151–161. https://doi.org/10.1016/j.neuroimage.2017.11.040.

Münte, T.F. (1992). *Hirnelektrische Korrelate von Sprache*. Medizinische Hochschule Hannover.

Münte, T.F., Heinze, H.J., 1994. ERP negativities during syntactic processing of written words. In: Heinze, H.J., Münte, T.F., Mangun, G.R. (Eds.), Cognitive Electrophysiology. Birkhäuser. https://doi.org/10.1007/978-1-4612-0283-7_9.

Münte, T.F., Heinze, H.J., Mangun, G.R., 1993. Dissociation of brain activity related to syntactic and semantic aspects of language. J. Cogn. Neurosci. 5 (3), 335–344. https://doi.org/10.1162/jocn.1993.5.3.335.

Murphy, E., 2015. The brain dynamics of linguistic computation. Front. Psychol. 6, 1–19. https://doi.org/10.3389/fpsyg.2015.01515.

Murphy, E., Woolnough, O., Rollo, P.S., Roccaforte, Z.J., Segaert, K., Hagoort, P., Tandon, N., 2022. Minimal phrase composition revealed by intracranial recordings. J. Neurosci. 42 (15), 3216–3227. https://doi.org/10.1523/jneurosci.1575-21.2022.

Näätänen, R., Gaillard, A.W.K., Mäntysalo, S., 1978. Early selective-attention effect on evoked potential reinterpreted. Acta Psychol. 42 (4), 313–329. https://doi.org/10.1016/0001-6918(78)90006-9.

Näätänen, R., Paavilainen, P., Rinne, T., Alho, K., 2007. The mismatch negativity (MMN) in basic research of central auditory processing: a review. Clin. Neurophysiol. 118 (12), 2544–2590. https://doi.org/10.1016/j.clinph.2007.04.026.

Näätänen, Risto, Pakarinen, S., Rinne, T., Takegata, R., 2004. The mismatch negativity (MMN): towards the optimal paradigm. Clin. Neurophysiol. 115 (1), 140–144. https://doi.org/10.1016/j.clinph.2003.04.001.

Nelson, M.J., El Karoui, I., Giber, K., Yang, X., Cohen, L., Koopman, H., Cash, S.S., Naccache, L., Hale, J.T., Pallier, C., Dehaene, S., 2017. Neurophysiological dynamics of phrase-structure building during sentence processing. Proc. Natl. Acad. Sci. U.S.A. 114 (18), E3669–E3678. https://doi.org/10.1073/pnas.1701590114.

Neufeld, C., Kramer, S.E., Lapinskaya, N., Heffner, C.C., Malko, A., Lau, E.F., 2016. The electrophysiology of basic phrase building. PLoS One 11 (10), 1–22. https://doi.org/10.1371/journal.pone.0158446.

Numssen, O., Zier, A.-L., Thielscher, A., Hartwigsen, G., Knösche, T.R., Weise, K., 2021. Efficient high-resolution TMS mapping of the human motor cortex by nonlinear regression. NeuroImage 245, 118654. https://doi.org/10.1016/j.neuroimage.2021.118654.

Pallier, C., Devauchelle, A.D., Dehaene, S., 2011. Cortical representation of the constituent structure of sentences. Proc. Natl. Acad. Sci. U.S.A. 108 (6), 2522–2527. https://doi.org/10.1073/pnas.1018711108.

Poulisse, C., Wheeldon, L., Segaert, K., 2019. Evidence against preserved syntactic comprehension in healthy aging. J. Exp. Psychol.: Learn. Mem. Cogn. 45 (12), 2290–2308. https://doi.org/10.1037/xlm0000707.

Poulisse, C., Wheeldon, L., Limachya, R., Mazaheri, A., Segaert, K., 2020. The oscillatory mechanisms associated with syntactic binding in healthy ageing. Neuropsychologia 146, 107523. https://doi.org/10.1016/j.neuropsychologia.2020.107523.

Pulvermüller, F., Assadollahi, R., 2007. Grammar or serial order?: discrete combinatorial brain mechanisms reflected by the syntactic mismatch negativity. J. Cogn. Neurosci. 19 (6), 971–980. https://doi.org/10.1162/jocn.2007.19.6.971.

Pulvermüller, F., Shtyrov, Y., 2003. Automatic processing of grammar in the human brain as revealed by the mismatch negativity. NeuroImage 20 (1), 159–172. https://doi.org/10.1016/S1053-8119(03)00261-1.

Pulvermüller, F., Shtyrov, Y., 2006. Language outside the focus of attention: the mismatch negativity as a tool for studying higher cognitive processes. Prog. Neurobiol. 79 (1), 49–71. https://doi.org/10.1016/j.pneurobio.2006.04.004.

Pulvermüller, F., Shtyrov, Y., Hasting, A.S., Carlyon, R.P., 2008. Syntax as a reflex: neurophysiological evidence for early automaticity of grammatical processing. Brain Lang. 104 (3), 244–253. https://doi.org/10.1016/j.bandl.2007.05.002.

Pulvermüller, F., Cook, C., Hauk, O., 2012. Inflection in action: semantic motor system activation to noun- and verb-containing phrases is modulated by the presence of overt grammatical markers. NeuroImage 60 (2), 1367–1379. https://doi.org/10.1016/j.neuroimage.2011.12.020.

Pyatigorskaya, E., Maran, M., & Zaccarella, E. (2021). Testing the automaticity of syntax using masked visual priming. *PsyArXiv*. https://doi.org/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.

Pylkkänen, L., 2020. Neural basis of basic composition: what we have learned from the red-boat studies and their extensions. Philos. Trans. R. Soc. B: Biol. Sci. 375 (1791) https://doi.org/10.1098/rstb.2019.0299.

Ratcliff, R., 1978. A theory of memory retrieval. Psychol. Rev. 85 (2), 59–108. https://doi.org/10.1037/0033-295X.85.2.59.

Ratcliff, R., Gomez, P., Mckoon, G., 2004. A diffusion model account of the lexical decision task. Psychol. Rev. 111 (1), 159–182. https://doi.org/10.1037/0033-295X.111.1.159.
Resnik, P., 1992. Left-Corner Parsing and Psychological Plausibility. *Proceedings of the 14th conference on Computational linguistics - Volume 1*, 191–197. https://doi.org/10.3115/992066.992098.
Sakai, K.L., Noguchi, Y., Takeuchi, T., Watanabe, E., 2002. Selective priming of syntactic processing by event-related transcranial magnetic stimulation of Broca's area. Neuron 35 (6), 1177–1182. https://doi.org/10.1016/S0896-6273(02)00873-5.
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), 1–8. https://doi.org/10.1111/psyp.13335.
Schell, M., Zaccarella, E., Friederici, A.D., 2017. Differential cortical contribution of syntax and semantics: An fMRI study on two-word phrasal processing. Cortex 96, 105–120. https://doi.org/10.1016/j.cortex.2017.09.002.
Schell, M., Friederici, A.D., Zaccarella, E., 2022. Neural classification maps for distinct word combinations in Broca 's area. PsyArXiv. https://doi.org/10.31234/osf.io/9yt8k.
Schmidt, R., 1986. Was weiß der Artikel vom Hauptwort? Ein Beitrag zur Verarbeitung syntaktischer Beziehungen beim Lesen. Z. Für Exp. Angew. Psychol. 33, 150–163.
Schoffelen, J.M., Hultén, A., Lam, N., Marquand, A.F., Uddén, J., Hagoort, P., 2017. Frequency-specific directed interactions in the human brain network for language. Proc. Natl. Acad. Sci. U.S.A. 114 (30), 8083–8088. https://doi.org/10.1073/pnas.1703155114.
Segaert, K., Mazaheri, A., Hagoort, P., 2018. Binding language: structuring sentences through precisely timed oscillatory mechanisms. Eur. J. Neurosci. 48 (7), 2651–2662. https://doi.org/10.1111/ejn.13816.
Sehyr, Z.S., Caselli, N., Cohen-Goldberg, A.M., Emmorey, K., 2021. The ASL-LEX 2.0 project: a database of lexical and phonological properties for 2,723 signs in American sign language. J. Deaf Stud. Deaf Educ. 26 (2), 263–277. https://doi.org/10.1093/deafed/enaa038.
Seidenberg, M.S., Waters, G.S., Sanders, M., Langer, P., 1984. Pre- and postlexical loci of contextual effects on word recognition. Mem. Cogn. 12 (4), 315–328. https://doi.org/10.3758/BF03198291.
Shtyrov, Y., Pulvermüller, F., Näätänen, R., Ilmoniemi, R.J., 2003. Grammar processing outside the focus of attention: an MEG study. J. Cogn. Neurosci. 15 (8), 1195–1206. https://doi.org/10.1162/089892903322598148.
Snijders, T.M., Petersson, K.M., Hagoort, P., 2010. Effective connectivity of cortical and subcortical regions during unification of sentence structure. NeuroImage 52 (4), 1633–1644. https://doi.org/10.1016/j.neuroimage.2010.05.035.
Steinhauer, K., Drury, J.E., 2012. On the early left-anterior negativity (ELAN) in syntax studies. Brain Lang. 120 (2), 135–162. https://doi.org/10.1016/j.bandl.2011.07.001.
Strijkers, K., Chanoine, V., Munding, D., Dubarry, A.S., Trébuchon, A., Badier, J.M., Alario, F.X., 2019. Grammatical class modulates the (left) inferior frontal gyrus within 100 ms when syntactic context is predictive. Sci. Rep. 9 (1), 1–13. https://doi.org/10.1038/s41598-019-41376-x.
Stroh, A.L., Rösler, F., Dormal, G., Salden, U., Skotara, N., Hänel-Faulhaber, B., Röder, B., 2019. Neural correlates of semantic and syntactic processing in German Sign Language. NeuroImage 200, 231–241. https://doi.org/10.1016/j.neuroimage.2019.06.025.
Tanner, D., 2015. On the left anterior negativity (LAN) in electrophysiological studies of morphosyntactic agreement: A Commentary on "Grammatical agreement processing in reading: ERP findings and future directions" by Molinaro etal., 2014. Cortex 66, 149–155. https://doi.org/10.1016/j.cortex.2014.04.007.
Tanner, D., Morgan-Short, L., Luck, S.J., 2015. How inappropriate high-pass filters can produce artifactual effects and incorrect conclusions in ERP studies of language and cognition. Psychophysiology 52 (8), 997–1009. https://doi.org/10.1111/psyp.12437.
Trettenbrein, P.C., Papitto, G., Friederici, A.D., Zaccarella, E., 2020. The functional neuroanatomy of sign language without speech. *Hum. Brain Mapp.* 1–14. 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). Behav. Res. Methods 53, 817–1832. 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. Front. Psychol. 12, 1–7. https://doi.org/10.3389/fpsyg.2021.628728.
Tsigka, S., Papadelis, C., Braun, C., Miceli, G., 2014. Distinguishable neural correlates of verbs and nouns: a MEG study on homonyms. Neuropsychologia 54 (1), 87–97. https://doi.org/10.1016/j.neuropsychologia.2013.12.018.
Tyler, L.K., Randall, B., Stamatakis, E.A., 2008. Cortical differentiation for nouns and verbs depends on grammatical markers. J. Cogn. Neurosci. 20 (8), 1381–1389. https://doi.org/10.1162/jocn.2008.20095.
Vaidya, A.R., Pujara, M.S., Petrides, M., Murray, E.A., Fellows, L.K., 2019. Lesion Studies in ContemporaryNeuroscience. Trends in Cognitive Sciences 23 (8), 653–671. https://doi.org/10.1016/j.tics.2019.05.009.
van den Broek, S.P., Reinders, F., Donderwinkel, M., Peters, M.J., 1998. Volume conduction effects in EEG and MEG. Electroencephalogr. Clin. Neurophysiol. 106, 522–534. https://doi.org/10.1016/S0013-4694(97)00147-8.
Vigliocco, G., Vinson, D.P., Druks, J., Barber, H., Cappa, S.F., 2011. Nouns and verbs in the brain: a review of behavioural, electrophysiological, neuropsychological and imaging studies. Neurosci. Biobehav. Rev. 35 (3), 407–426. https://doi.org/10.1016/j.neubiorev.2010.04.007.
Voss, A., Rothermund, K., Voss, J., 2004. Interpreting the parameters of the diffusion model: an empirical validation. Mem. Cogn. 32 (7), 1206–1220. https://doi.org/10.3758/BF03196893.
Weise, K., Numssen, O., Thielscher, A., Hartwigsen, G., Knösche, T.R., 2020. A novel approach to localize cortical TMS effects. NeuroImage 209. https://doi.org/10.1016/j.neuroimage.2019.116486.
Weiss, S., Mueller, H.M., 2012. "Too many betas do not spoil the broth": the role of beta brain oscillations in language processing. Front. Psychol. 3, 1–15. https://doi.org/10.3389/fpsyg.2012.00201.
West, R.F., Stanovich, K.E., 1982. Source of inhibition in experiments on the effect of sentence context on word recognition. J. Exp. Psychol.: Learn. Mem. Cogn. 8 (5), 385–399. https://doi.org/10.1037/0278-7393.8.5.385.
Widmann, A., Schröger, E., Maess, B., 2015. Digital filter design for electrophysiological data - a practical approach. J. Neurosci. Methods 250, 34–46. https://doi.org/10.1016/j.jneumeth.2014.08.002.
Wu, C.Y., 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. Hum. Brain Mapp. 40 (6), 1705–1717. https://doi.org/10.1002/hbm.24482.
Zaccarella, E., Friederici, A.D., 2015a. Reflections of word processing in the insular cortex: a sub-regional parcellation based functional assessment. Brain Lang. 142, 1–7. https://doi.org/10.1016/j.bandl.2014.12.006.
Zaccarella, E., Friederici, A.D., 2015b. Merge in the human brain: a sub-region based functional investigation in the left pars opercularis. Front. Psychol. 6, 1–9. https://doi.org/10.3389/fpsyg.2015.01818.
Zaccarella, E., Friederici, A.D., 2017. The neurobiological nature of syntactic hierarchies. Neurosci. Biobehav. Rev. 81, 205–212. https://doi.org/10.1016/j.neubiorev.2016.07.038.
Zaccarella, E., Meyer, L., Makuuchi, M., Friederici, A.D., 2017a. Building by syntax: the neural basis of minimal linguistic structures. Cereb. Cortex 27 (1), 411–421. https://doi.org/10.1093/cercor/bhv234.
Zaccarella, E., Schell, M., Friederici, A.D., 2017b. Reviewing the functional basis of the syntactic Merge mechanism for language: a coordinate-based activation likelihood estimation meta-analysis. Neurosci. Biobehav. Rev. 80, 646–656. https://doi.org/10.1016/j.neubiorev.2017.06.011.
Zaccarella, E., Papitto, G., Friederici, A.D., 2021. Language and action in Broca's area: computational differentiation and cortical segregation. Brain Cogn. 147, 105651 https://doi.org/10.1016/j.bandc.2020.105651.
Zorzi, M., Di Bono, M.G., Fias, W., 2011. Distinct representations of numerical and non-numerical order in the human intraparietal sulcus revealed by multivariate pattern recognition. NeuroImage 56 (2), 674–680. https://doi.org/10.1016/j.neuroimage.2010.06.035.

Glossary of terms

Alpha-band oscillations: Oscillatory activity occurring with a frequency between 8 and 12 Hz.
Baseline correction: Pre-processing step which consists in the subtraction of the average of pre-stimulus activity from the one recorded after the stimulus presentation, with the goal of removing drifts in the EEG and MEG signal (Luck, 2005).
Beta-band oscillations: Oscillatory activity occurring with a frequency between 12 and 30 Hz.
Closed-class elements: Function words such as determiners, pronouns, prepositions, conjunctions and modal verbs. The term "closed-class" refers to the fact that, in a given language, their number does not increase, contrary to "open-class" elements.
Cross-splicing technique: A technique which allows to control for low-level acoustic features of the experimental materials by extracting and concatenating audio segments (Hasting et al., 2007).
Early Left-Anterior Negativity (ELAN): An ERP component elicited by syntactic categorical violations, occurring approximately between 150 and 250 ms (Friederici, 2011).
Event-Related Potential (ERP): Stereotyped EEG voltage fluctuations that the brain produces in a fixed time relationship to a specific stimulus or event (Luck, 2005).
High-pass filtering: Pre-processing step which consists in the attenuation of frequencies below a cut-off value (e.g., 0.1 Hz) in the EEG and MEG signal (Luck, 2005; Widmann et al., 2015).
Left-Anterior Negativity (LAN): An ERP component elicited by syntactic agreement violations, occurring approximately between 300 and 500 ms (Molinaro et al., 2011).
N400: An ERP component elicited by lexical-semantic violations, with a peak occurring approximately at 400 ms (Kutas and Federmeier, 2011).
Masking technique: Fast presentation of a stimulus ("mask") in order to suppress the conscious perception of another stimulus of interest.
Morphemes: The smallest linguistic unit which can carry meaning (Gwilliams, 2020). Functional morphemes such suffixes (e.g., "-s") serve a primarily syntactic function.
Open-class elements: Content words such as nouns and verbs, whose number can change in a given language as new elements are added.
P600: An ERP component occurring approximately between 500 and 1000 ms indexing late linguistic repairing and re-analysis processes (Friederici, 2011).
Parsing: The process which consists in the incremental analysis of linguistic information, captured by parsing models described in the field of computational linguistics (Hale, 2014).
Pseudo-words: Word-like stimuli which respect the orthographic and phonological rules of a language but are not present in its lexicon (e.g., "boal" in English).
Re-referencing: EEG measurements reflect differences in voltage relative to a reference electrode, usually located on the subject's body. Re-referencing is a mathematical transformation of the recorded data to a new reference (Luck, 2005).

Recursivity: Property of an operation indicating the possibility of re-applying itself to its own output.

Red-boat paradigm: Experimental paradigm which consists in the comparison of a noun phrase (e.g., "red boat") against a list (e.g., "cup boat") to highlight semantic and conceptual composition.

Stimulus onset asynchrony: Temporal interval between the onset time of two consecutive stimuli.

Theta-band oscillations: Oscillatory activity occurring with a frequency between 4 and 8 Hz.

Transcranial magnetic stimulation (TMS): Non-invasive neurostimulation technique which allows to perturb the functioning of a target brain region (Hartwigsen, 2015).