Title

Lack of selectivity for syntax relative to word meanings throughout the language network

Abbreviated title

Syntax and word meanings in the brain

Authors

Evelina Fedorenko[1,2,3], Idan Blank[1,4], Matthew Siegelman[1,5], Zachary Mineroff[1,6]

Affiliations

[1]Department of Brain and Cognitive Sciences, MIT, Cambridge, MA 02139, USA
[2]McGovern Institute for Brain Research, MIT, Cambridge, MA 02139, USA
[3]Department of Psychiatry, MGH, Charlestown, MA 02129, USA
[4]Department of Psychology, UCLA, Los Angeles, CA 90095 USA
[5]Department of Psychology, Columbia University, New York, NY 10027, USA
[6]Eberly Center for Teaching Excellence & Educational Innovation, CMU, Pittsburgh, PA 15213, USA

Corresponding Author

Ev Fedorenko
evelina9@mit.edu; 43 Vassar Street, Room 46-3037G, Cambridge, MA, 02139

Abstract

To understand what you are reading now, your mind retrieves the meanings of words and constructions from a linguistic knowledge store (lexico-semantic processing) and identifies the relationships among them to construct a complex meaning (syntactic or combinatorial processing). Do these two sets of processes rely on distinct, specialized mechanisms or, rather, share a common pool of resources? Linguistic theorizing, empirical evidence from language acquisition and processing, and computational modeling have jointly painted a picture whereby lexico-semantic and syntactic processing are deeply inter-connected and perhaps not separable. In contrast, many current proposals of the neural architecture of language continue to endorse a view whereby certain brain regions selectively support syntactic/combinatorial processing, although the locus of such "syntactic hub", and its nature, vary across proposals. Here, we searched for selectivity for syntactic over lexico-semantic processing using a powerful individual-subjects fMRI approach across three sentence comprehension paradigms that have been used in prior work to argue for such selectivity: responses to lexico-semantic *vs.* morpho-syntactic violations (**Experiment 1**); recovery from neural suppression across pairs of sentences differing in only lexical items *vs.* only syntactic structure (**Experiment 2**); and same/different meaning judgments on such sentence pairs (**Experiment 3**). Across experiments, both lexico-semantic and syntactic conditions elicited robust responses throughout the left fronto-temporal language network. Critically, however, no regions were more strongly engaged by syntactic than lexico-semantic processing, although some regions showed the opposite pattern. Thus, contra many current proposals of the neural architecture of language, syntactic/combinatorial processing is not separable from lexico-semantic processing at the level of brain regions—or even voxel subsets—within the language network, in line with strong integration between these two processes that has been consistently observed in behavioral and computational language research. The results further suggest that the language network may be generally more strongly concerned with meaning than syntactic form, in line with the primary function of language—to share meanings across minds.

Introduction

The functional architecture of language processing

What is the functional architecture of human language? A core component is a set of knowledge representations, which include knowledge of words/constructions and their meanings, and the probabilistic constraints on how these linguistic elements can combine to create compound words, phrases, and sentences. During comprehension (decoding of linguistic utterances), we look for matches between elements/substrings in the incoming linguistic signal and these stored knowledge representations, and compose these retrieved representations, in an attempt to re-construct the intended meaning, and during production (encoding of linguistic utterances), we search our knowledge store for the right words/constructions and combine and arrange them in a particular way to express a target idea.

How is this rich set of representations and computations structured? Specifically, which aspects of language are functionally dissociable from one another? Traditionally, two principal distinctions have been drawn: one is between words (the lexicon) and rules (the grammar) (e.g., Chomsky, 1965, 1995; Fodor, 1983; Pinker & Prince, 1988; Pinker, 1991, 1999); and another is between linguistic representations themselves (i.e., our knowledge of the language) and their online processing (i.e., accessing them from memory and combining them to create new complex meanings and structures) (e.g., Chomsky, 1965; Fodor et al., 1974; Newmeyer, 2003). Because these dimensions are, in principle, orthogonal, we could have distinct mental capacities associated with i) knowledge of word (lexical) meanings, ii) knowledge of grammar (syntactic/combinatorial rules), iii) access (or "retrieval") of lexical representations, iv) access of syntactic/combinatorial rules, and v) combining retrieved representations into new complex representations (**Fig. 1a**).

However, both of these distinctions have been long debated. For example, as linguistic theorizing evolved and experimental evidence accumulated through the 1970s-90s, the distinction between the lexicon and grammar began to blur, both for the storage of linguistic knowledge representations and for online processing (e.g., **Fig. 1b**; see Snider & Arnon, 2012, for a summary and discussion). Many have observed that much of our grammatical knowledge does not operate over general categories like nouns and verbs, but instead requires reference to particular words or word classes (e.g., verbs that can occur in a particular construction, like the ditransitive) (e.g., Lakoff, 1970; Bybee, 1985, 1998, 2010; Levin, 1993; Goldberg, 1995, 2002; Jackendoff, 2002a,b, 2007; Sag et al., 2003; Culicover & Jackendoff, 2005; Levin & Rappaport Hovav, 2005; Audring & Jackendoff, 2020). As a result, current linguistic frameworks incorporate knowledge of "rules" (i.e., syntactic structures) into the mental lexicon, although they differ as to the degree of abstraction that exists above and beyond knowledge of how particular words combine with other words (e.g., Ambridge, 2018; see Hudson, 2007, for discussion), and in whether abstract syntactic representations (like the double object, passive, or question constructions) are always associated with meanings or functions (e.g., Pinker, 1989;

Goldberg, 1995; cf. Chomsky, 1957; Branigan & Pickering, 2017; see Jackendoff, 2002, for discussion).

In line with these changes in linguistic theorizing, experimental and corpus work in psycholinguistics have established that humans i) are exquisitely sensitive to contingencies between particular words and the constructions they occur in (e.g., Clifton et al., 1984; MacDonald et al., 1994; Trueswell et al., 1994; Garnsey et al., 1997; Traxler et al., 2002; Reali & Christiansen, 2007; Roland et al., 2007; Jaeger, 2010), and ii) store not just atomic elements (like morphemes and non-compositional lexical items), but also compositional phrases (e.g., "I don't know" or "give me a break"; e.g., Wray, 2005; Evert, 2008; Arnon & Snyder, 2010; Morgan & Levy, 2016; Christiansen & Arnon, 2017) and constructions (e.g., "the X-er the Y-er"; Goldberg, 1995; Culicover & Jackendoff, 1999). The latter suggested that the linguistic units people store are determined not by their nature (i.e., atomic vs. not) but instead, by their patterns of usage (e.g., Bybee 1998, 2006; Goldberg 2006; Barlow and Kemmer 2000; Langacker 1986, 1987; Tomasello 2003). Further, people's lexical abilities have been shown to strongly correlate with their grammatical abilities—above and beyond shared variance due to general fluid intelligence—both developmentally (e.g., Bates et al., 1988, 1995; Bates & Goodman, 1997; Dale et al., 2010; Dixon & Marchman, 2007; Snedeker et al., 2007, 2012; Hoff et al., 2018) and in adulthood (e.g., Dabrowska, 2018). Thus, linguistic mechanisms that have been previously proposed to be distinct are instead tightly integrated or, perhaps, are so cognitively inseparable as to be considered a single apparatus.

The distinction between stored knowledge representations and online computations has also been questioned (see Hasson et al., 2015, for a discussion of this issue in language and other domains). For example, by using the same artificial network to represent all linguistic experience, connectionist models dispense not only with the lexicon-grammar distinction but also the storage-computation one, and assume that the very same units that represent our linguistic knowledge support its online access and processing (e.g., Rumelhart & McClelland, 1986; Seidenberg, 1994; Devlin et al., 2019; see also Goldinger, 1996; Bod, 1998, 2006, for exemplar models, which also abandon the storage-computation divide).

Syntax selectivity in prior cognitive neuroscience investigations?

Alongside psycholinguistic studies, which inform debates about linguistic architecture by examining the behaviors generated by language mechanisms, and computational work, which aims to approximate human linguistic behavior using formal models, a different, complementary approach is offered by cognitive neuroscience studies. These studies aim to constrain the cognitive architecture by examining how cognitive processes are neurally implemented (e.g., Kanwisher, 2010; Mather et al., 2013). The assumption that links neuroimaging data (and neuropsychological patient data) to cognitive hypotheses is as follows: if distinct brain regions or sets of regions support the processing of manipulations targeting cognitive processes X and Y, we can infer that X and Y are dissociable. Such brain regions would be expected to show distinct patterns of response

in brain imaging studies, and their damage should lead to distinct patterns of cognitive deficits.

A large number of brain imaging investigations have observed distinct loci of activation for manipulations that target (lexico-)semantic vs. syntactic processing, and—following the reasoning above—have argued for a dissociation between the two (e.g., Dapretto & Bookheimer, 1999; Embick et al., 2000; Kuperberg et al., 2000; Ni et al., 2000; Newman et al., 2001; Kuperberg et al., 2003; Noppeney & Price, 2004; Cooke et al., 2006; Friederici et al., 2010; Glaser et al., 2013; Schell et al., 2017, inter alia). Consequently, many proposals of the neural architecture of language postulate a component that selectively supports syntactic, or more general combinatorial, processing relative to the storage/processing of individual word meanings (e.g., Grodzinsky & Santi, 2008; Baggio and Hagoort, 2011; Friederici, 2011, 2012; Tyler et al., 2011; Duffau et al., 2014; Ullman, 2016; Matchin & Hickok, 2019; Pylkkanen, 2019; cf. Bornkessel-Schlesewsky & Schlesewsky, 2009; Bornkessel-Schlesewsky et al., 2015 – we come back to these proposals in the Discussion). Although proposals vary in which component(s) of syntactic processing are emphasized—from morpho-syntactic agreement, to dependency structure building and composition, to word-order-related processes—the general idea of syntax selectivity remains prominent in cognitive neuroscience of language. How can we fit such selectivity with current linguistic theorizing, psycholinguistic evidence, and computational modeling work, which suggest strong integration between lexico-semantic and syntactic representations and processing, as discussed above? Here, we revisit past evidence and report a new study to ***argue against the existence of brain regions that are selective for syntactic/combinatorial processing over the processing of word meanings***.

There already exist ***reasons to doubt*** the existence of syntax-selective brain regions if we take a closer look at the cognitive neuroscience of language literature. ***First***, the specific brain regions that have been argued to support syntactic/combinatorial processing, and the construal of these regions' contributions, *differ across studies and theoretical proposals* (e.g., Baggio and Hagoort, 2011; Bemis & Pylkkanen, 2011; Friederici, 2011, 2012; Tyler et al., 2011; Duffau et al., 2014; Ullman, 2004, 2016; Matchin & Hickok, 2019; Pylkkanen, 2019). For example, the proposed location of the "core" syntactic/combinatorial hub varies between the inferior frontal cortex (e.g., Friederici et al., 2006; Hagoort, 2005, 2013), the posterior temporal cortex (e.g., Matchin & Hickok, 2019), and the anterior temporal cortex (e.g., Vandenberghe et al., 2002; Bemis & Pylkkanen, 2011; Pylkannen, 2019), with some positing multiple hubs (e.g., Pallier et al., 2011; Tyler et al., 2011).

Second, at least some syntactic/combinatorial manipulations appear to engage the *entire fronto-temporal language network* (e.g., Fedorenko et al., 2010; Pallier et al., 2011; Blank et al., 2016), putting into question the idea of syntactic processing being focally carried out. Relatedly, studies of patients with brain damage have failed to consistently link syntactic deficits with a particular region within the language network. Instead, damage to any component of the network appears to lead to similar syntactic difficulties—which also mirror patterns observed in neurotypical individuals under cognitive load (Miyake et al., 1994)—leading some to argue that syntactic processing is

supported by the language network as a whole (e.g., Caplan et al., 1996; Dick et al., 2001; Wilson and Saygin, 2004; Mesulam et al., 2014, 2015). Given that the language network has to additionally support lexico-semantic processing, these findings necessarily imply that at least some, and possibly all, of these syntax-responsive regions overlap with regions that support lexico-semantic processing.

Third, a number of studies have actually *failed to observe syntax selectivity*, showing that brain regions that respond to syntactic manipulations also show reliable, and sometimes stronger, responses to lexico-semantic manipulations (e.g., Chee et al., 1999; Keller et al., 2001; Roder et al., 2002; Luke et al., 2002; Heim et al., 2008; Rogalsky & Hickok, 2009; Fedorenko et al., 2010, 2012a, 2016; Bautista & Wilson, 2016; Blank et al., 2016; Shain et al., 2019; see Rodd et al., 2015, for a meta-analysis). Relatedly, some studies have reported activations for lexico-semantic manipulations, such as lexical ambiguity manipulations, in what appear to be the same regions as the ones implicated in other studies in syntactic/combinatorial processing: i.e., regions in the left inferior fontal and posterior temporal cortex (e.g., Rodd et al., 2005, 2010, 2012; Davis et al., 2007; Mason and Just, 2007; Zempleni et al., 2007; Bilenko et al., 2008; Bekinschtein et al., 2011).

And ***fourth***, many prior studies that have reported syntax selectivity suffer from *methodological and statistical limitations*. For example, although diverse paradigms have been used across studies to probe lexico-semantic vs. syntactic/combinatorial processing, any given study has typically used a single paradigm, raising the possibility that the results reflect paradigm-specific differences between conditions rather than a general difference between lexico-semantic and syntactic/combinatorial representations or computations. Furthermore, many studies that claimed to have observed a dissociation have not reported the required region-by-condition interactions, as needed to argue for a functional dissociation between brain regions (Nieuwenhuis et al., 2011). And some studies have argued for syntax selectivity based solely on sensitivity to syntactic complexity manipulations, without even examining responses to lexico-semantic processing (e.g., Stromswold et al., 1996; Ben-Shachar et al., 2003; Fiebach et al., 2005; Santi & Grodzinsky, 2010; see Friederici, 2011, for a meta-analysis). Although such studies (may) establish that a brain region is engaged in syntactic processing, they say little about its *selectivity* for syntactic over lexico-semantic processing. Finally, some studies have reported sensitivity to syntactic manipulations in regions that fall outside the boundaries of the fronto-temporal language network. For example, some studies of morpho-syntactic violations (e.g., Kuperberg et al., 2003; Nieuwland et al., 2012) have reported effects in regions that resemble the domain-general bilateral fronto-parietal network implicated in executive control (e.g., Duncan, 2010, 2013). This network is sensitive to unexpected events across domains (e.g., Corbetta & Shulman, 2002), and some of its regions lie in close proximity to the language regions (e.g., Fedorenko et al., 2012b; Fedorenko & Blank, in press). Although the precise nature of this network's contribution to language processing remains debated (e.g., Fedorenko, 2014; Diachek et al., 2019; Ryskin et al., 2020a), sensitivity of this network to a linguistic manipulation likely indexes a domain-general process and does not inform the question of whether different components of the *language network* support syntactic vs. lexico-semantic processing.

Motivation for the current study

The current study aims to ***resolve the conflict*** between i) converging evidence from linguistic theorizing, behavioral psycholinguistic work, and computational modeling, which have jointly painted a clear picture of strong integration between lexico-semantic and syntactic representations and processing, and ii) cognitive neuroscience studies and proposals, many of which continue to suggest the existence of syntax- or combinatorics-selective brain regions. To do so, we use fMRI to search for syntax selectivity using a *robust individual-subjects analytic approach*, including a well-validated task for identifying the language network (i.e., a "functional localizer"; Fedorenko et al., 2010) across *three classic paradigms* from the literature that contrast lexico-semantic and syntactic processing. We will now discuss and motivate both of these features of our study in greater detail, in the historical context of the field.

The need for robust, replicable, and cumulative science

Over the years, a large number of paradigms have been used across brain imaging studies to probe lexico-semantic and syntactic processes and their relationship (see references in Materials and Methods). Some paradigms (discussed below) have varied *the presence or absence* of lexico-semantic vs. syntactic information in the linguistic signal; others have *more strongly taxed* the processing of word meanings vs. syntactic structure; and still others have made the meaning of a particular word vs. the structure of the sentence *more salient / task-relevant.* Any of these kinds of manipulations could be informative, but no single manipulation—and certainly not from a single experiment—would suffice to argue for syntax selectivity. To compellingly argue that a brain region selectively supports (some aspect of) syntactic processing, one would need to demonstrate the *robustness* of the syntactic > lexico-semantic effect (e.g., replication in a new sample of participants, on a new set of experimental materials, and/or in another imaging modality) and its *generalizability* to other contrasts between conditions that engage the hypothesized computation and ones that do not. In particular, the selectivity of a brain region for the critical syntactic condition(s) would need to be established relative to a *broad range of control conditions*, given that any given syntactic vs. lexico-semantic pair of conditions will likely differ in ways beyond the target distinction between syntax and semantics. For example, showing that morpho-syntactic violations elicit a stronger response than semantic violations is not sufficient to argue for syntactic selectivity because these two conditions also differ in whether the error can be explained by a plausible noise process within a noisy-channel framework of sentence comprehension (e.g., Ferreira et al., 2002; Levy et al., 2009; Gibson et al., 2013); thus, a stronger response to morpho-syntactic violations could reflect the relevant correction process, which does not get engaged for typical semantic violations (cf. Ryskin et al., 2020b). Furthermore, it would be critical to ensure that—across studies—the critical effect arises in the *same* brain region, not just within the same broad macroanatomic area, like the left inferior frontal gyrus (LIFG) (see Hong et al., 2019, for a general discussion of challenges to determining what counts as a "replication" in fMRI research). To the best of our knowledge, nobody has demonstrated syntactic selectivity using this kind of a rigorous approach, or even attempted to do so. At

least two factors have likely contributed to the lack of such attempts, and to the resulting lack of clarity in the field.

First, until a decade ago, the issue of replicability has not been much discussed in the fields of psychology / cognitive science (e.g., Ioannidis et al., 2014) and cognitive neuroscience (e.g., Poldrack et al., 2017). And at least some of the findings in cognitive neuroscience of language that have been taken at face value based on a single report by a single group may not be robust and replicable (e.g., see Siegelman et al., 2019, for a recent attempt, and failure, to replicate a much cited report by Dapretto & Bookheimer, 1999). This issue is further compounded by the many hidden degrees of freedom (e.g., Simmons et al., 2011) that characterize the choices during the preprocessing and analysis of brain imaging data (Botvinik-Nazer et al., 2019) and the common use of "double dipping" (e.g., Krigeskorte et al., 2009) in many early studies.

And second, as we have previously argued (e.g., Fedorenko & Kanwisher, 2009; Fedorenko et al., 2010, 2012b; Blank et al., 2017; Fedorenko & Blank, 2020; see also Brett et al., 2002 and Saxe et al., 2006), establishing a cumulative research enterprise in cognitive neuroscience of language has been challenging due to the difficulty of comparing findings across studies that rely on the traditional group-averaging approach (e.g., Holmes & Friston, 1998). In this analytic approach—which has dominated the brain-imaging language research in the 1990s and 2000s and is still in common use despite being strongly disfavored in neuroimaging studies in other domains—individual activation maps are aligned in a common space, and the output is a set of coordinates in that space for voxels where significant effects obtain (typically the most reliable peak(s) in each activation cluster are reported). The main way to compare results across such studies is to compare the anatomical locations of these activation peaks. However, group-level activation peaks are noisy (e.g., Fedorenko et al., 2012b), due to the combined effect of two factors, both especially pronounced in the association cortex, which houses the language system: (1) high inter-individual variability in the locations of functional areas (e.g., Fedorenko et al., 2010; Mahowald & Fedorenko, 2016; Braga et al., 2019); and (2) lack of correspondence between functional areas and macroanatomic landmarks, like gyri and sulci (e.g., Frost & Goebel, 2011; Tahmasebi et al., 2011; Vázquez-Rodríguez et al., 2019). And even the most systematic comparisons in the form of meta-analyses of activation peaks from large numbers of studies (e.g., Bookheimer, 2002; Costafreda et al., 2006; Indefrey & Levelt, 2004; Kaan & Swaab, 2002; Lindenberg et al., 2007; Poldrack et al., 1999; Vigneau et al., 2006; Binder et al., 2009) are not very informative (Kvarven et al., 2019), and have been shown to lead to fundamentally wrong conclusions about the functional architecture of the human brain in some cases (e.g., Aguirre & Farah, 1998). So what is a solution?

A decade ago, we developed an alternative approach to the study of language in the brain (Fedorenko et al., 2010)—one that had been successful in other domains, including high-level vision (e.g., Kanwisher et al., 1997) and social cognition (e.g., Saxe & Kanwisher, 2003), and has now become widespread in the study of language (e.g., Axelrod et al., 2015; Lane et al., 2015; Poldrack et al., 2015; Wang et al., 2015; Matchin et al., 2019; Braga et al., 2019). In this approach, language-responsive areas are defined functionally

in individual brains without being constrained to fall precisely in the same anatomical locations across participants; and the localized regions are then probed for their responses to critical experimental manipulations. The use of the same "localizer" paradigm across individuals, studies, and research groups (and in domains, like vision, this is done across species, too; e.g., Tsao et al., 2008) provides a straightforward way to *directly relate findings* to one another.

Choice of paradigms

Using the individual-subjects functional localization approach, we have previously argued for the lack of syntax selectivity based on a paradigm that varies the presence of lexico-semantic vs. syntactic information in the linguistic signal (Fedorenko et al., 2010, 2012a, 2016; for earlier uses of this paradigm, see e.g., Mazoyer et al., 1993; Friederici et al., 2000; Humphries et al., 2001; Vandenberghe et al., 2002; for another variant, see Bautista & Wilson, 2016). In particular, we examined the processing of i) sentences, which have a syntactic structure and consist of real, interpretable words, ii) lists of unconnected words, which lack structure but are individually interpretable, iii) "Jabberwocky" sentences, which preserve a syntactic frame (word order and morpho-syntactic endings), but have the words replaced by nonwords, so the meanings of those strings cannot be interpreted with respect to our world knowledge, aside from very coarse-level semantics, and finally, iv) lists of unconnected nonwords, which lack both structure and interpretability. Across three replications with fMRI (Fedorenko et al., 2010; see Mollica et al., in prep. for another replication) and, in addition, in a more spatially and temporally sensitive method (electrocorticography, ECoG) (Fedorenko et al., 2016), we found that any language-responsive brain region or electrode that shows sensitivity to syntactic structure (i.e., stronger responses to sentences than word lists, and to Jabberwocky sentences than nonword lists) is at least as sensitive, and often more sensitive, to meanings of individual words (showing stronger responses to sentences than Jabberwocky sentences, and to word lists than nonword lists).

However, one could question the findings from this paradigm because the contrasts are rather crude and the materials are artificial/unnatural. If the overlap in the brain mechanisms that process individual word meanings and syntactic structure is a real and robust finding, the results should generalize to other, finer-grained comparisons between lexico-semantic and syntactic processing. As a result, we selected three paradigms from studies that have argued for syntax selectivity or for dissociations between lexico-semantic and syntactic processing, and that continue to be cited as evidence of such, and attempted to conceptually replicate (Schmidt, 2009) them.

Experiments 1 and 3 are designed to differentially tax lexico-semantic vs. syntactic processing by having a critical word in a sentence be incompatible with the context in terms of either its its meaning or morpho-syntactic properties (Experiment 1), or by forcing participants to focus on the meanings of the critical words or the structure of sentences (Experiment 3). Experiment 2 relies on the well-established neural adaptation to the repetition of a stimulus and recovery from such adaptation when some relevant feature of the stimulus changes: here, a change in the individual words (but not the

sentence structure) vs. the sentence structure (but not the words). Along with the manipulations varying the presence/absence of lexico-semantic and syntactic information in the linguistic signal discussed above, these manipulations span the space of available manipulations targeting lexico-semantic and syntactic processing quite comprehensively[1], and—for syntactic processing—cover both morpho-syntactic agreement (Experiment 1), and dependency structure building / word-order-related processes (Experiments 2 and 3).

If *any brain region* within the language network selectively supports syntactic processing, we would expect stronger responses to the syntactic than the lexico-semantic condition in that region in at least one paradigm. If this pattern holds—for the same brain region(s)—across two or all three paradigms, that would further help rule out paradigm-specific between-condition differences/confounds and strengthen the conclusion. Note that unlike the paradigms that vary the presence/absence of syntactic and lexical information in the linguistic signal—where one could, in principle, observe a pattern where a brain region is not *at all* engaged unless structure or meaning is present (although in practice, even lists of pseudowords, which lack both structure and meaning, elicit an above-baseline response across much of the language network; Fedorenko et al., 2010)—the current paradigms all use sentence materials across conditions, so all conditions are expected to elicit above-baseline responses throughout the language network. The critical question is whether any brain region(s) would exhibit stronger responses for the syntactic compared to lexico-semantic condition in one or more paradigms. If no brain region within the language network shows this pattern, this would strongly reinforce the conclusions drawn from the paradigms that have varied the presence/absence of lexico-semantic and syntactic information in the signal (e.g., Fedorenko et al., 2010; Bautista & Wilson, 2016).

To foreshadow the results: using the most sensitive analytic methods available in fMRI (e.g., Nieto-Castañon & Fedorenko, 2012), we find robust responses to both lexico-semantic and syntactic processing throughout the language network in each of the three experiments. Critically, every brain region in the language network that responds to syntactic manipulations responds at least as strongly to lexico-semantic manipulations. No region—or even set of non-contiguous voxels within these regions—shows a consistent preference, in the form of a stronger response, for syntactic processing (ruling out architectures in **Figure 1a-d**). However, in line with our prior work (e.g., Fedorenko et al., 2012a, 2016), some regions show the opposite preference—a stronger response to lexico-semantic processing. We therefore hope that this study brings clarity to the field

[1] One type of manipulation missing here is one that relies on ambiguity. Both lexical (e.g., Rodd et al., 2005, 2010, 2012; Davis et al., 2007; Mason and Just, 2007; Zempleni et al., 2007; Bekinschtein et al., 2011) and structural (e.g., Mason et al., 2003) ambiguity have been investigated in prior fMRI studies—although the former has received more attention—and both have been shown to elicit responses in the inferior frontal and posterior temporal cortex. Because, to the best of our knowledge, no arguments for syntactic selectivity have been made based on stronger responses to syntactic than lexical ambiguity in some part(s) of the language network, we did not include ambiguity manipulations here.

and helps build stronger bridges, grounded in robust empirical work, with behavioral and computational investigations of language processing.

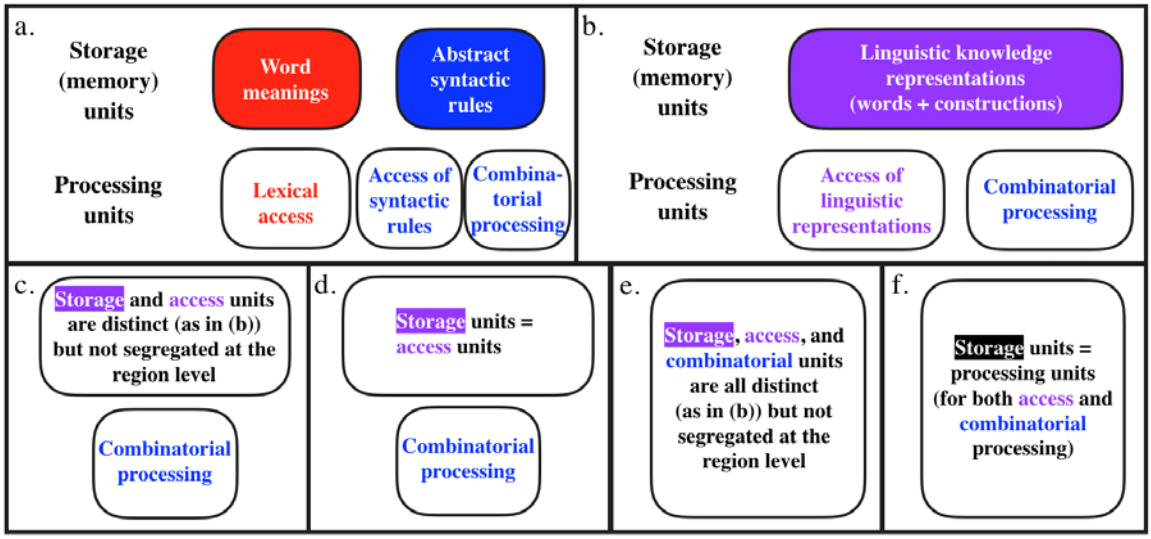


Figure 1: **A (non-exhaustive) set of theoretically possible architectures of language.** Distinct boxes correspond to distinct brain regions (or sets of brain regions; e.g., in 1a-d, "combinatorial processing" may recruit a single region or multiple regions, but critically, this region or these regions do not support other aspects of language processing, like understanding word meanings). The architectures differ in whether they draw a (region-level) distinction between the lexicon and grammar (a vs. b-f), between storage and access of linguistic representations (1a-b vs. 1c-f), and critically, in whether combinatorial processing is a separable component (1a-d vs. 1e-f).

Materials and Methods

General description of the paradigms and their use in prior studies

The *first paradigm* is commonly used in ERP investigations of language processing and relies on violations of expectations about an incoming word that are set up by the preceding context. In particular, the critical word does not conform to either the lexico-semantic or the syntactic expectations (e.g., Kutas & Hilliyard, 1980; Osterhout & Holcomb, 1992; Hagoort et al., 1993). This paradigm has been used in a number of prior fMRI studies (e.g., Embick et al., 2000; Newman et al., 2001; Kuperberg et al., 2003; Cooke et al., 2006; Friederici et al., 2010; Herrmann et al., 2012). The *second paradigm* relies on neural adaptation, wherein repeated exposure to a stimulus leads to a reduction in response, and a change in some feature(s) of the stimulus leads to a recovery of response (see e.g., Krekelberg et al., 2006, for a general overview of the approach). This paradigm has also been used in prior fMRI studies that examined adaptation, or recovery from adaptation, to the lexico-semantic vs. syntactic features of linguistic stimuli (e.g., Noppeney & Price, 2004; Devauchelle et al., 2009; Santi & Grodzinsky, 2010; Menenti et al., 2012; Segaert et al., 2012). Finally, the *third paradigm* was introduced in a classic study by Dapretto & Bookheimer (1999): pairs of sentences vary in a single word vs. in word order / syntactic structure. Either of these manipulations can result in the same meaning being expressed across sentences (if a word is replaced by a synonym, or if a syntactic alternation, like active→passive, is used) or in a different meaning (if a word is

replaced by a non-synonym, or if the thematic roles are reversed). Participants make same/different meaning judgments on the resulting sentence pairs.

General analytic approaches

In an effort to maximize sensitivity and functional resolution (e.g., Nieto-Castañon & Fedorenko, 2012), we adopt an approach where all the key contrasts are performed within individual participants. We perform two kinds of analyses. In one set of analyses, we identify language-responsive areas in each individual participant with an independent language "localizer" task based on a contrast between the reading of sentences vs. sequences of nonwords (Fedorenko et al., 2010), and compare the responses of these areas to the lexico-semantic vs. syntactic conditions in each critical paradigm. As discussed in the Introduction, the use of the same language localizer task across experiments allows for a straightforward comparison of their results, obviating the need to rely on coarse anatomy and reverse-inference reasoning (Poldrack, 2006, 2011) for interpreting activations in functional terms.

To further ensure that we are not missing syntax selectivity by averaging across (relatively) large sets of language-responsive voxels (see Friston et al., 2006, for discussion of this potential issue), we supplement these analyses with analyses where we only use data from the critical paradigms. In particular, we use some of the data from a given critical task to search for the most syntactic-selective (or lexico-semantic-selective, examined for completeness) voxels (e.g., in Experiment 1, voxels that respond more strongly to morpho-syntactic than semantic violations), and then test the replicability of this selectivity in left-out data, as detailed below.

Participants

Forty-nine individuals (age 19-32, 27 females) participated for payment (Experiment 1: n=22; Experiment 2: n=14; and Experiment 3: n=15; 2 participants overlapped between Experiments 1 and 3). Forty-seven were right-handed, as determined by the Edinburgh handedness inventory (Oldfield, 1971), or self-report; the two left-handed individuals showed typical left-lateralized language activations in the language localizer task (see Willems et al., 2014, for arguments to include left-handers in cognitive neuroscience research). All participants were native speakers of English from Cambridge, MA and the surrounding community. One additional participant was scanned (for Experiment 2) but excluded from the analyses due to excessive sleepiness and poor behavioral performance. All participants gave informed consent in accordance with the requirements of MIT's Committee on the Use of Humans as Experimental Subjects (COUHES).

Design, stimuli, and procedure

Each participant completed a language localizer task (Fedorenko et al., 2010) and a critical task (35 participants performed the localizer task in the same session as the critical task, the remaining 14 performed the localizer in an earlier session; see Mahowald & Fedorenko, 2016 and Braga et al., 2019, for evidence that localizer activations are

stable across scanning sessions). Most participants completed one or two additional tasks for unrelated studies. The entire scanning session lasted approximately 2 hours.

Language localizer task. The task used to localize the language network is described in detail in Fedorenko et al. (2010). Briefly, we used a reading task that contrasted sentences and lists of unconnected, pronounceable nonwords in a standard blocked design with a counterbalanced order across runs (for timing parameters, see **Table 1**). This contrast targets higher-level aspects of language including, critically, both lexico-semantic and syntactic/combinatorial processing, to the exclusion of perceptual (speech or reading-related) and articulatory processes (see e.g., Fedorenko & Thompson-Schill, 2014, for discussion). As discussed above, brain regions thus localized indeed show sensitivity to lexico-semantic processing (as evidenced by stronger responses to sentences than Jabberwocky sentences, and to word lists than to nonword lists) and to syntactic processing (as evidenced by stronger responses to sentences that word lists, and to Jabberwocky sentences than nonword lists) (e.g., Fedorenko et al., 2010, 2012a, 2016; Mollica et al., in prep.). Stimuli were presented one word/nonword at a time. For 19 participants, each trial ended with a memory probe and they had to indicate, via a button press, whether or not that probe had appeared in the preceding sequence of words/nonwords. The remaining 30 participants read the materials passively and performed a simple button-press task at the end of each trial (included in order to help participants remain alert). Importantly, this localizer has been shown to generalize across different versions: the sentences > nonwords contrast, and similar contrasts between language and a degraded control condition, robustly activates the fronto-temporal language network regardless of the task, materials, and modality of presentation (Fedorenko et al., 2010; Fedorenko, 2014; Scott et al., 2016). Furthermore, the same network robustly emerges from naturalistic-cognition paradigms (e.g., resting state, listening to stories, watching movies) using the functional correlation approach (e.g., Blank et al., 2014; Tie et al., 2014; Branco et al., 2019; Braga et al., 2019), suggesting that this network constitutes a natural kind in the brain, and our localizer contrast is simply a quick and efficient way to identify the relevant areas.

Critical experiments. The key details about the three experiments are presented in **Figure 2** (sample stimuli), **Figure 3** (trial structure), and **Table 2** (partitioning of stimuli into experimental lists and runs). To construct experimental stimuli, we first generated for each experiment a set of "base items" and then edited each base item to create several, distinct versions corresponding to different experimental conditions. The resulting stimuli were divided into several experimental lists following a Latin Square design, such that in each list (i) stimuli were evenly split across experimental conditions, and (ii) only one version of each item was used. Each participant saw materials from a single list, divided into a few experimental runs. All experiments used an event-related design. Condition orders were determined with the optseq2 algorithm (Dale, 1999), which was also used to distribute inter-trial fixation periods so as to optimize our ability to de-convolve neural responses to different experimental conditions. The materials for all experiments are available from OSF (https://osf.io/abjy9/).

	Experiment 1: Violations task: passive reading		Experiment 2: Recovery from Adaptation task: did a probe word appear in either of the two sentences?		Experiment 3: Meaning Judgment task: same or different meaning?	
	Example 1	Example 2	Example 1	Example 2	Example 1	Example 2
	Although his ears were damaged [...]	If the post-office is closed [...]	The protestor quoted the leader	The officer alerted the authorities	The scientist flattered David	The woman sold Anna a dress
Lexico-semantic	[...] the man could still cook	[...] John cannot visit the letter	The striker cited the chief	The deputy notified the higher-ups	The scientist misled David	The woman sold Anna a gown
	[meaning violation]		[different words]		[non-synonym; different]	[synonym; same]
Syntactic	[...] the man could still hears	[...] John cannot mails the letter	The leader was quoted by the protestor	The authorities were alerted by the officer	David flattered the scientist	The woman sold a dress to Anna
	[morpho-syntactic violation]		[different structure]		[words flipped; different]	[from DO to PP; same]
Other	[...] the man could still hear	[...] John cannot mail the letter	The leader quoted the protestor	The authorities alerted the officer		
	[font violation]		[different global meaning]			
Control	[...] the man could still hear	[...] John cannot mail the letter	The protestor quoted the leader	The leader was quoted by the protestor		
	[no violation]		[all same]	[all different]		

Figure 2: Sample stimuli for each condition in Experiments 1-3. Two examples are provided for each condition in each experiment. For Experiment 1, the top row shows the beginning of a sentence, and the next rows show different possible continuations. For Experiments 2-3, the top row shows one sentence from a pair, and the next rows show different possibilities for the other sentence in that pair. Red: Lexico-semantic condition; Blue: Syntactic condition; Green: other experimental conditions; Black: control condition. [NB1: For Experiment 1, the task was passive reading for the critical materials, but, as described in the text, a small number of (filler) trials contained a memory probe task. NB2: For Experiment 2, three versions of the same base item (corresponding to the Lexico-semantic, Syntactic, and Global meaning conditions) are presented for illustrative purposes. As detailed in the text, in the actual materials, distinct sets of base items were used for the three critical conditions in order to match the number of trials across conditions while avoiding sentence repetition.]

Experiment 1: Lexico-semantic vs. (morpho-)syntactic violations. Participants passively read stimuli, and their expectations were violated in several ways. The items were 10-word sentences, with four versions each (**Figure 2**): the critical verb (i) resulted in a lexico-semantic violation (stimuli that typically elicit an N400 component in ERP studies; see Kutas & Federmeier, 2011, for a review); (ii) resulted in a morpho-syntactic violation (stimuli that typically elicit a P600 component in ERP studies; e.g., Osterhout & Holcomb, 1992; Hagoort et al., 1993); (iii) contained no violations (control condition); or (iv) was presented in a different font (a low-level oddball violation, included as an additional, stricter, control condition). Lexico-semantic violations were created by shuffling the critical verbs across the base items. Syntactic violations were created by

either omitting a required morpheme (30%) or adding an unnecessary morpheme (70%). (The reason we included a higher proportion of added compared to missing morphemes is because form-based error correction mechanisms are so robust during language comprehension that grammatical errors are often missed during proofreading (e.g., Schotter et al., 2014), and noticing missing elements seems harder than the extra ones.)

Overall, there were 240 items. They included 139 items with a sentence-final critical verb, taken (and sometimes slightly edited) from Kuperberg et al. (2003), as well as 61 additional items (to increase power) constructed in a similar manner. Further, to render the timing of violations less predictable, we adapted another 40 items from Kuperberg et al. such that the critical verb appeared before the final (10th) word: 6 items had the verb in each of the 3rd through 8th positions, and 4 items had it in the 9th position. Critical verbs were not repeated across the 240 items, with two exceptions ("practice" and "read" were used twice each). For each participant, 10 additional sentences were included in each of the four conditions to serve as fillers. These fillers were followed by a memory-probe task (deciding whether the probe word appeared in the preceding sentence; **Figure 3**) to ensure that participants paid attention to the task; they were excluded from data analysis.

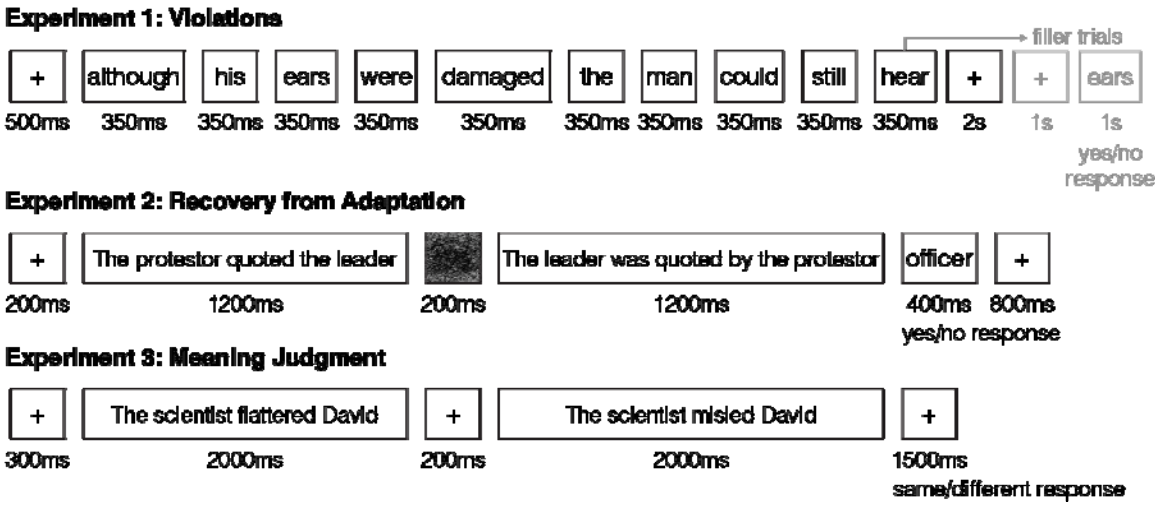


Figure 3: Trial structure for Experiments 1-3. One sample trial is shown for each experiment.

Experiment 2: Recovery from adaptation to word meanings vs. syntactic structure. Participants were asked to attentively read pairs of sequentially presented sentences and perform a memory probe task at the end of each pair (i.e., decide whether a probe word appeared in either of the two sentences). The sentences were simple transitive sentences consisting of an agent, a verb, and a patient. Because of constraints on these materials (as elaborated below), we constructed three sets of items (**Figure 2**): (i) sentence pairs that differed only in lexical items (but had the same syntactic structure and global meaning), created by replacing the verb and the agent and patient noun phrases with synonyms or words closely related in meaning; (ii) pairs that differed only in their syntactic structure (but had the same lexical items and global meaning), created by using the Active/Passive

alternation; and (iii) pairs that differed only in the global meaning (but had the same lexical items and syntactic structure), created by switching the two noun phrases, leading to opposite thematic role assignment. The third set was included in order to examine sensitivity to overall propositional meaning and to probe combinatorial semantic processing. Overall, there were 432 items (144 per set).

In each set, each sentence pair {A,B} had six versions (**Table 2**): sentence A followed by sentence B, and sentence B followed by sentence A ("Critical" condition); sentence A followed by sentence A, and sentence B followed by sentence B ("Same" condition); and, finally, sentence A followed by a completely different sentence X (lexical items, syntactic structure, and global meaning were all different), and sentence B followed by a completely different sentence Y, where the pair {X,Y} was taken from another item ("Different" condition). Every sentence was used once in the Different condition of some other item. Therefore, within each of the three sets of items, every sentence appeared twice in each condition (Critical, Same, Different). Across the three sets, there were overall 5 experimental conditions: Critical Lexico-semantic, Critical Syntactic, Critical Global meaning, Same, and Different. In order to clearly mark the distinctness of the two identical sentences in the Same condition, trials across all conditions included a brief visual mask between the two sentences.

To keep the materials semantically diverse, items in the first two sets were constructed to be evenly distributed among three types of agent-patient relationships: (1) animate agent + inanimate patient; (2) animate agent + animate patient, where the relationship is biased so that one of the noun phrases is much more likely to be the agent (e.g., *The hit man killed the politician*); and (3) animate agent + animate patient, where the two nouns are similarly likely to be the agent (e.g., *The protestor quoted the leader*). By virtue of the manipulation of global meaning in the third set, all items had to be semantically reversible (i.e., of the third type).

Experiment 3: Same-different meaning judgment on sentences that differ in word meanings vs. syntactic structure. This design was adapted from Dapretto & Bookheimer (1999). Participants were asked to decide whether or not a pair of sequentially presented sentences had roughly the same meaning. The items were 80 sentence pairs, and each pair had four versions (**Figure 2**; **Table 2**): two versions in which the sentences differed in a single word (Lexico-semantic condition), replaced by either a synonym (Same meaning version) or a non-synonym (Different meaning version); and two versions (Syntactic condition) in which the sentences were either syntactic alternations differing in both structure and word order (Same meaning version), or in only structure / only word order (Different meaning version). Half of the items used the Active/Passive constructions (as in Dapretto & Bookheimer), and half – the Double Object (DO) / Prepositional Phrase Object (PP) constructions.

A number of features varied and were balanced across items (**Figure 2**). First, the construction was always the *same* across the two sentences in the Lexico-semantic condition (balanced between active and passive for the Active/Passive items, and between DO and PP for the DO/PP items). However, in the Syntactic condition, the

construction was always *different* in the Same-meaning version because this is how the propositional meaning was preserved (again, balanced between active and passive for the Active/Passive items, and between DO and PP for the DO/PP items). For the Different-meaning version, the construction could either be the same (in which case the order of the two relevant nouns was switched) or different (in which case the order of the two relevant nouns was preserved). In cases where the construction differed across the two sentences, we balanced whether the first sentence was active *vs.* passive (for the Active/Passive items), or whether it was DO *vs.* PP (for the DO/PP items). The second feature that varied across the materials was whether the first-mentioned noun was a name or an occupation noun. All items contained one instance of each, with order of presentation balanced across stimuli. And third, for the Lexico-semantic condition, we varied how exactly the words in the second sentence in a pair differed from the words in the first. (This does not apply to the Syntactic condition because the content words were identical across the two sentences within each pair.) In particular, for the Active/Passive items, either the occupation noun or the verb could be replaced (by a synonym or a word with a different meaning); and for the DO/PP items, either the occupation noun or the direct object (inanimate) noun could be replaced.

Data acquisition, preprocessing, and first-level modeling

Data acquisition. Whole-brain structural and functional data were collected on a whole-body 3 Tesla Siemens Trio scanner with a 32-channel head coil at the Athinoula A. Martinos Imaging Center at the McGovern Institute for Brain Research at MIT. T1-weighted structural images were collected in 176 axial slices with 1mm isotropic voxels (repetition time (TR) = 2,530ms; echo time (TE) = 3.48ms). Functional, blood oxygenation level-dependent (BOLD) data were acquired using an EPI sequence with a 90° flip angle and using GRAPPA with an acceleration factor of 2; the following parameters were used: thirty-one 4.4mm thick near-axial slices acquired in an interleaved order (with 10% distance factor), with an in-plane resolution of 2.1mm × 2.1mm, FoV in the phase encoding (A >> P) direction 200mm and matrix size 96 × 96 voxels, TR = 2000ms and TE = 30ms. The first 10s of each run were excluded to allow for steady state magnetization.

Preprocessing. Data preprocessing was carried out with SPM5 (using default parameters, unless specified otherwise) and supporting, custom MATLAB scripts. (Note that SPM was only used for preprocessing and basic modeling—aspects that have not changed much between versions; we used an older version of the SPM software because we have projects that use data collected over the last 15 years, and we want to keep preprocessing and first-level analysis the same across the 800+ subjects in our database, which are pooled in the analyses for many projects. For several datasets, we have directly compared the outputs of data preprocessed and modeled in SPM5 vs. SPM12, and the outputs are nearly identical.) Preprocessing of anatomical data included normalization into a common space (Montreal Neurological Institute (MNI) template) and resampling into 2mm isotropic voxels. Preprocessing of functional data included motion correction (realignment to the mean image of the first functional run using 2nd-degree b-spline interpolation), normalization (estimated for the mean image using trilinear interpolation),

resampling into 2mm isotropic voxels, smoothing with a 4mm FWHM Gaussian filter and high-pass filtering at 200s.

Data modeling. For both the language localizer task and the critical tasks, a standard mass univariate analysis was performed in SPM5 whereby a general linear model (GLM) estimated the effect size of each condition in each experimental run. These effects were each modeled with a boxcar function (representing entire blocks/events) convolved with the canonical Hemodynamic Response Function (HRF). The model also included first-order temporal derivatives of these effects, as well as nuisance regressors representing entire experimental runs and offline-estimated motion parameters.

Definition and validation of language-responsive functional regions of interest (fROIs)

For each participant (in each experiment), we defined a set of language-responsive functional ROIs using group-constrained, participant-specific localization (Fedorenko et al., 2010). In particular, each individual participant's map for the sentences > nonwords contrast from the language localizer task was intersected with a set of six binary masks. These masks were derived from a probabilistic activation overlap map for the language localizer contrast in a large set of participants (n=220) using the watershed parcellation, as described in Fedorenko et al. (2010), and corresponded to relatively large areas within which most participants showed activity for the target contrast. These masks covered the fronto-temporal language network: three in the left frontal lobe falling within the IFG, its orbital portion, and the MFG, and three in the temporal and parietal cortex (**Figure 5**). Within each mask, a participant-specific language fROI was defined as the top 10% of voxels with the highest *t*-values for the localizer contrast. This top *n*% approach ensures that fROIs can be defined in every participant and that their sizes are the same across participants, allowing for generalizable results (e.g., Nieto-Castañón and Fedorenko, 2012).

Before examining the data from the critical experiments, we ensured that the language fROIs show the expected signature response (i.e., that the response is reliably greater to sentences than nonwords). To do so, we used an across-runs cross-validation procedure (e.g., Nieto-Castañón & Fedorenko, 2012), where one run of the localizer is used to define the fROIs, and the other run to estimate the responses, insuring independence (e.g., Kriegeskorte et al., 2009). As expected, and replicating prior work (e.g., Fedorenko et al., 2010; Fedorenko et al., 2011; Blank et al., 2016; Mahowald & Fedorenko, 2016), the language fROIs showed a robust sentences > nonwords effect ($ts_{(48)}>8.44$; $ps<0.0001$), correcting for the number of regions (six) using the False Discovery Rate (FDR) correction (Benjamini & Yekutieli, 2001).

Estimating the responses of the language fROIs to the conditions of the critical experiments

We estimated the responses in the language fROIs to the conditions of each critical experiment: the Control condition, Lexico-semantic violations, Syntactic violations, and Font violations in Experiment 1; the Same condition, Different condition, and three

Critical conditions (differing in only lexical items, only syntactic structure, or only global meaning) in Experiment 2; and the Lexico-semantic and Syntactic conditions (each collapsed across same and different pairs) in Experiment 3. Statistical comparisons were performed on the estimated percent BOLD signal change (PSC) values in each region in each experiment.

We analyzed each experiment separately to allow for the possibility that syntax selectivity would be observed in just one of the experiments. Such a pattern could still be potentially informative and would be missed in an analysis that pools data from the three experiments. Furthermore, we examined each region separately, in line with our research question: whether *any* region within the language network is selective for syntactic over lexico-semantic processing.

Reality-check analyses: Testing for sensitivity to lexico-semantic and syntactic processing. First, we tested for basic sensitivity to lexico-semantic and syntactic manipulations. In each region, we used two-tailed paired-samples *t*-tests to compare the response to each critical (lexico-semantic and syntactic) condition to one or more control conditions. In Experiment 1, we compared the response to each critical violation condition (Lexico-semantic or Syntactic) against a) the Control condition with no violations, and, as an additional, stricter, baseline, b) the Font violation condition. In Experiment 2, we compared the Same and Different conditions to each other (a reality check to test for recovery from adaptation in the language regions when all the features of the sentence change), and then we compared each of the Critical conditions to the Same condition (where the same sentence is repeated exactly) to test for recovery from adaptation when the lexical items or the syntactic structure changes. The predictions are similar for Experiments 1 and 2: if a brain region is sensitive to lexical processing, the Lexico-semantic condition should elicit a stronger response than the control condition(s); similarly, if a brain region is sensitive to syntactic processing, the Syntactic condition should elicit a stronger response than the control condition(s). Finally, in Experiment 3, we compared the response to each condition (Lexico-semantic and Syntactic) against the low-level fixation baseline to ensure robust responses in the language regions to sentence comprehension. (Note that fixation was used here because, unlike in the other two experiments, there was no other baseline condition following Dapretto & Bookheimer's (1999) design.) In each of these reality-check analyses, the results were FDR-corrected (Benjamini & Yekutieli, 2001) for the six regions.

Critical analyses (a): Directly comparing the Lexico-semantic and Syntactic conditions in the language fROIs. Next, we directly compared the Lexico-semantic and Syntactic conditions in each region in each experiment. If a brain region is selectively or preferentially engaged in syntactic processing, then we would expect to observe a reliably stronger response to the Syntactic condition than the Lexico-semantic condition. And if a brain region is selectively/preferentially engaged in lexico-semantic processing, we would expect to observe the opposite pattern. In these analyses, we report the results without a correction for the number of regions because we wanted to give syntactic selectivity—which we are arguing *against*—the best chance to reveal itself. (Of course, if an uncorrected *p*-value fails to reach significance, then the corrected one does, too.)

Critical analyses (b): Searching for voxels selective for syntactic (or lexico-semantic, for completeness) processing. One potential concern with the use of language fROIs is that each fROI is relatively large and the responses are averaged across voxels (e.g., Friston et al., 2006). Thus, fROI-based analyses may obscure underlying functional heterogeneity and potential selectivity for syntactic processing. For example, if a fROI contains a subset of voxels that show a stronger response to lexico-semantic than syntactic processing, and another subset of voxels that show a stronger response to syntactic than lexico-semantic processing, we may not detect a difference at the level of the fROI as a whole. To circumvent this concern, we supplemented the analyses of language fROIs, with analyses that i) use some of the data from each critical experiment to directly search for voxels—within the same broad masks encompassing the language network—that respond more strongly to syntactic than lexico-semantic processing (i.e., top 10% of voxels based on the Syntactic>Lexico-Semantic contrast), or vice versa (for completeness), and then ii) examine the replicability of this pattern of response in a left-out portion of the data. We performed this analysis for each of Experiments 1-3. If any (even non-contiguous) voxels with reliably stronger responses to syntactic processing exist anywhere within the fronto-temporal language network, this analysis should be able to detect them. For these analyses, we used one-tailed paired-samples *t*-tests because these hypotheses are directional. For example, when examining voxels that show stronger responses to syntactic than lexico-semantic processing to test whether this preference is replicable in left-out data, the critical contrast is Syntactic>Lexico-Semantic. As in the last set of analysis, these results were not corrected for the number of regions because we wanted to give syntactic selectivity the best chance to reveal itself.

Results

Behavioral results

Error rates and reaction times (RTs), for trials with a recorded response, in each of the three experiments are summarized in **Figure 4**. Performance on the memory probe task in the filler trials in Experiment 1 was close to ceiling (between 95.4% and 96.6% across conditions), with no reliable difference between the two critical—Lexico-semantic and Syntactic—conditions, in accuracies or RTs ($ts_{(21)}<1$, *n.s.*). In Experiment 2, performance on the memory probe task varied between 72.6% and 95.7% across conditions. As expected, participants were faster and more accurate in the Same condition, where the same sentence was repeated, than in the Different condition, where the two sentences in the pair differed in all respects ($ts_{(13)}>15.1$, $ps<0.05$). Furthermore, participants were faster and more accurate in the Syntactic condition than in the Lexico-semantic condition ($ts(_{13)}>3.55$, $ps<0.05$), plausibly because the words were repeated between the two sentences in the pair in the Syntactic, but not Lexico-semantic condition. However, the difference was small, with high performance (>89.7%) in both conditions. Finally, in line with the behavioral results in Dapretto & Bookheimer (1999), in Experiment 3, performance on the meaning judgment task did not differ between the Lexico-semantic and Syntactic conditions, in accuracies or RTs ($ts_{(14)}<1$, *n.s.*). In summary, in each of the three experiments, performance was high across conditions, with no reliable differences

between the Lexico-semantic and Syntactic conditions in Experiments 1 and 3, and only a small difference in Experiment 2. We can thus proceed to examine neural differences between lexico-semantic and syntactic processing without worrying about those differences being driven by systematic differences in processing difficulty.

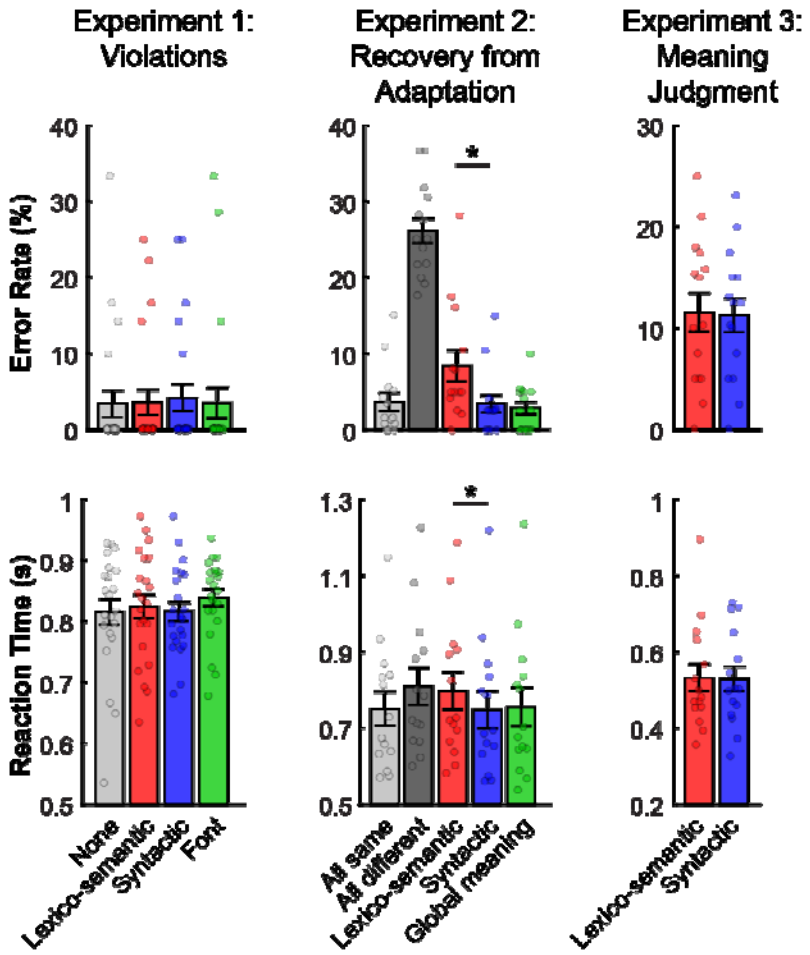


Figure 4: Summary of the behavioral results from Experiments 1-3. Each column shows the results from a different experiment: left – Experiment 1: Violations; middle – Experiment 2: Recovery from Adaptation; right – Experiment 3: Meaning Judgment. Top graphs: error rates; bottom graphs: reaction times. Here and in all subsequent figures, for each condition and each measure, dots correspond to individual participants; the bar shows the average across these participants, and the error bar shows standard error of the mean (across participants). Significant differences between the critical, Lexico-semantic (red) and Syntactic (blue), conditions are marked with *'s.

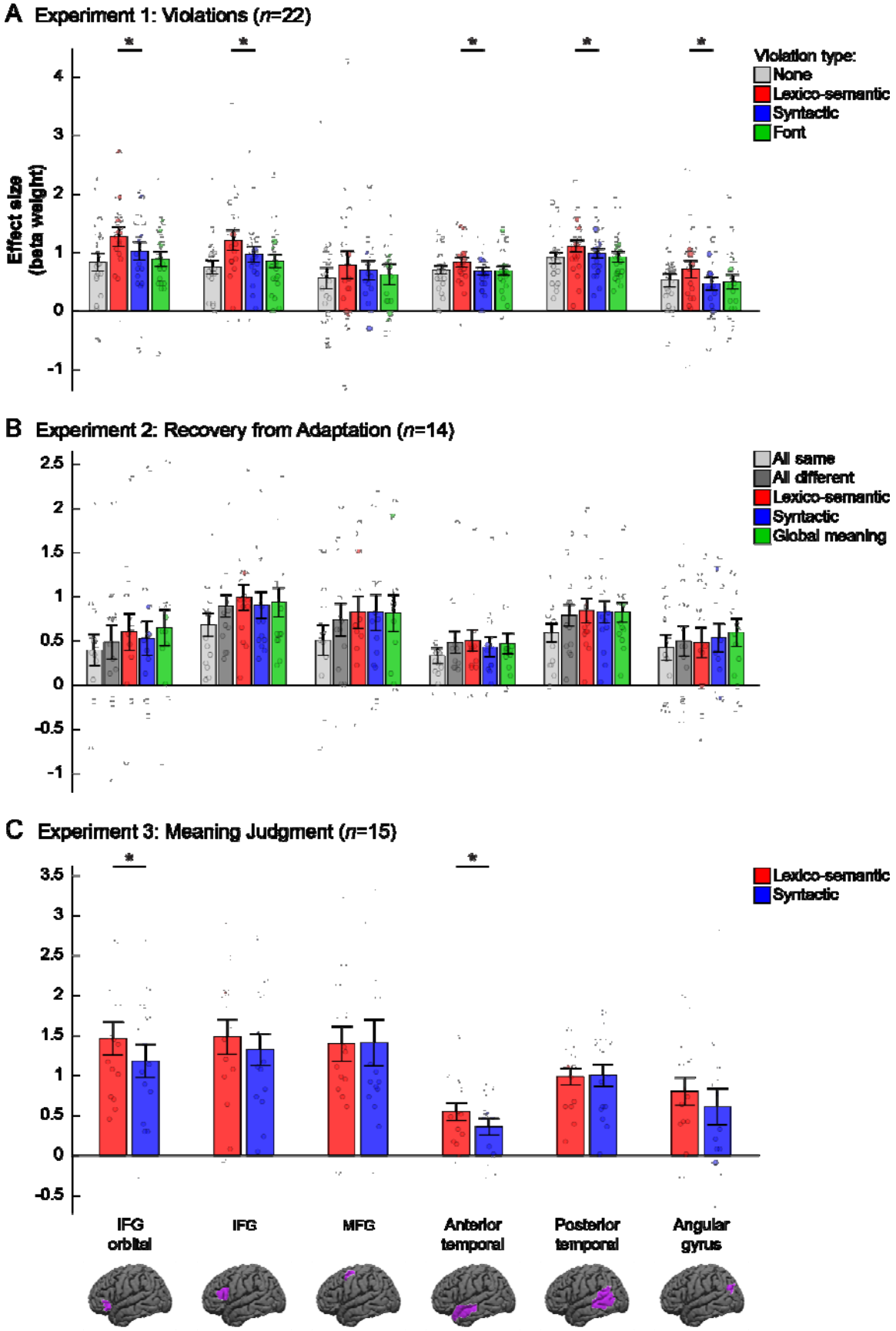
fMRI results

1. Reality-check analyses: Testing for sensitivity to lexico-semantic and syntactic processing. The results for the three experiments are summarized in **Figure 5** and **Table 3**. In Experiment 1, the Lexico-semantic condition elicited a reliably stronger response than the Control (no violations) condition in each of six language fROIs ($ts_{(21)}$>2.77,

*p*s<0.05), and than the Font violations condition in five of the six fROIs ($ts_{(21)}$>3.43, *p*s<0.05), with the MFG fROI not showing a significant effect. The Syntactic condition elicited a reliably stronger response than the Control condition in two language fROIs: IFGorb and IFG ($ts_{(21)}$>2.84, *p*s<0.05). However, the Syntactic condition did not reliably differ from the Font violations condition in any of the fROIs ($ts_{(21)}$<1.45, *n.s.*), suggesting that language regions are not recruited more strongly when people encounter syntactic violations than they are when people encounter low-level perceptually unexpected features in the linguistic input (see also Vissers et al., 2006; van de Meerendonk et al., 2011; see also Mollica et al., 2020, for extension to word-order violations). In Experiment 2, the Different condition—where the two sentences in a pair differed in lexical items, syntactic structure, and global meaning—elicited a reliably stronger response than the Same condition, where the two sentences in a pair were identical, in four language fROIs: IFG, MFG, AntTemp, and PostTemp (*t*s(13)>3.45, *p*s<0.05); the effect was not reliable in the IFGorb and AngG fROIs. Further, the Lexico-semantic condition elicited a response that was reliably stronger than the Same condition in all language fROIs, except for the AngG fROI ($ts_{(13)}$>2.71, *p*s<0.05), and similarly, the Syntactic condition elicited a response that was reliably stronger than the Same condition in all language fROIs, except for the AngG fROI ($ts_{(13)}$>2.33, *p*s<0.05)[2]. Finally, in Experiment 3, both experimental conditions elicited responses that were reliably above the fixation baseline in all six fROIs (Lexico-semantic: $ts_{(14)}$>4.58, *p*s<0.05; Syntactic: $ts_{(14)}$>2.66, *p*s<0.05).

2. Critical analyses (a): Directly comparing the Lexico-semantic and Syntactic conditions in the language fROIs. The results for the three experiments are summarized in **Figure 5** and **Table 3**. In Experiment 1, a direct comparison between the Lexico-semantic and Syntactic conditions revealed reliably stronger responses to the Lexico-semantic condition in all language fROIs except for the MFG fROI ($ts_{(21)}$>2.12, *p*s<0.05). In Experiment 2, no language fROI showed reliably stronger recovery from adaptation in the Lexico-semantic than Syntactic condition or vice versa ($ts_{(13)}$<1.44, *n.s.*). Finally, in Experiment 3, we observed a stronger response to the Lexico-semantic than Syntactic condition in two language fROIs—IFGorb and AntTemp ($ts_{(14)}$>2.27, *ps*<0.05). No language fROI showed the opposite pattern.

[2] It is worth noting that, similar to the Lexico-semantic and Syntactic conditions, the Global meaning condition also elicited a response that was reliably stronger than the Same condition in all language fROIs (*p*s<0.05). This effect provides evidence that language regions are sensitive to differences in complex meanings above and beyond the meanings of individual words (given that the only thing that differs between the sentences in a pair in the Global-meaning condition is word order).


A Experiment 1: Violations (n=22)
Effect size (beta weight)
Violation type:
None
Lexico-semantic
Syntactic
Font
B Experiment 2: Recovery from Adaptation (n=14)
All same
All different
Lexico-semantic
Syntactic
Global meaning
C Experiment 3: Meaning Judgment (n=15)
Lexico-semantic
Syntactic
IFG orbital
IFG
MFG
Anterior temporal
Posterior temporal
Angular gyrus

882

Figure 5: Responses in language fROIs to the conditions in Experiments 1-3. Responses (beta weights from a GLM) are measured as PSC relative to the fixation baseline. Each panel shows the results from a different experiment: (A) Experiment 1: Violations; (B) Experiment 2: Recovery from Adaptation; (C) Experiment 3: Meaning Judgment. Within each panel, each group of bars shows data from a different fROI (legend at the bottom). Significant differences between the critical, Lexico-semantic (red) and Syntactic (blue), conditions are marked with *'s. (The significance of the reality-check analyses that establish the sensitivity of the language fROIs to the lexico-semantic and syntactic manipulations relative to the control conditions is not marked; please see Results and Table 3.) Brain images at the bottom show the broad masks used to constrain the selection of the individual fROIs (these are not the fROIs themselves, which were defined using the sentences>nonwords contrast separately in each participant, as described in the text, and could thus vary within the borders of the masks depicted here).

3. Critical analyses (b): Searching for voxels selective for syntactic (or lexico-semantic, for completeness) processing. The results for the three experiments are summarized in **Figures 6** and **7** and **Table 4**. In Experiment 1, when we defined the individual fROIs by the Syntactic>Lexico-semantic contrast, we did not find a replicable (across runs) Syntactic>Lexico-semantic effect within any of the masks. In contrast, when we defined the fROIs by the Lexico-semantic>Syntactic contrast, we found a replicable Lexico-semantic>Syntactic effect within all the language masks, except for the MFG mask ($ts_{(21)}$>2.75, ps<0.05), consistent with our finding of stronger responses for the Lexico-semantic than Syntactic condition in the fROIs that were defined by the language localizer contrast.

In Experiment 2, when we defined the individual fROIs by the Syntactic>Lexico-semantic contrast, we found a slightly higher response to the Syntactic than Lexico-semantic condition within the PostTemp mask ($t_{(14)}$=1.84, p<0.05). This effect would not survive correction for the number of regions. And even if, with a larger number of participants, the statistical robustness of the Syntactic>Lexico-semantic effect increased, the difference between the two conditions would remain small. When we defined the fROIs by the Lexico-semantic>Syntactic contrast, we did not find a replicable Lexico-semantic>Syntactic effect in any of the language masks, in line with the two critical conditions not eliciting differential responses in the fROIs defined by the language localizer contrast in this experiment.

Finally, in Experiment 3, as in Experiment 1, when we defined the individual fROIs by the Syntactic>Lexico-semantic contrast, we did not find a replicable Syntactic>Lexico-semantic effect within any of the masks; but when we defined the fROIs by the Lexico-semantic>Syntactic contrast, we found a replicable Lexico-semantic>Syntactic effect within all the language masks, except for the MFG mask ($ts_{(14)}$>2.14, ps<0.05).

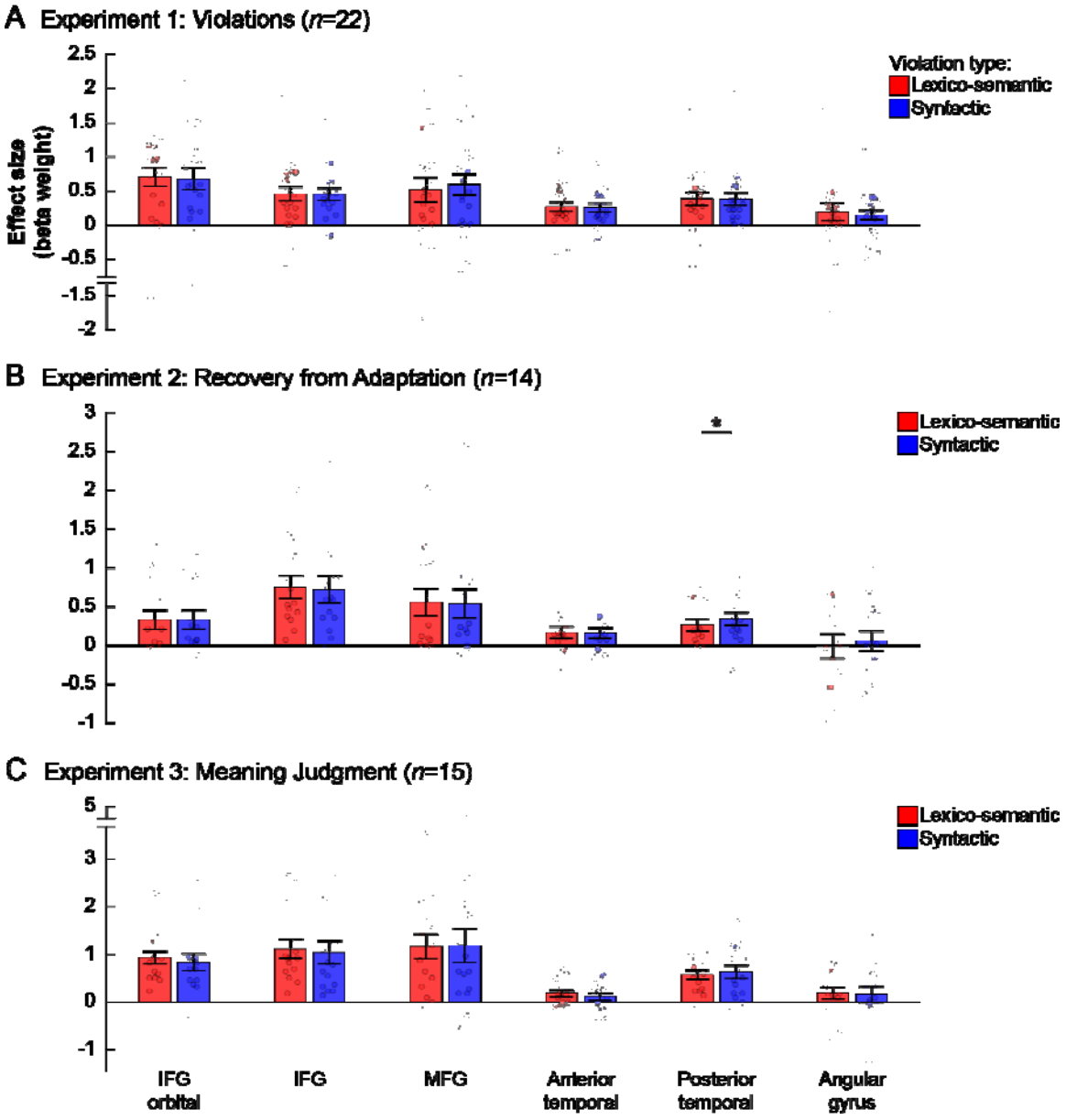


Figure 6: Responses in fROIs defined by the Syntactic>Lexico-semantic contrast to the critical conditions in Experiments 1-3. Participant-specific fROIs were defined, within the borders of each mask (Figure 5), as the top 10% of voxels showing the strongest Syntactic>Lexico-semantic contrast effect in the corresponding experiment. These fROIs were defined based on half the data from that experiment, and then the other (independent) half were used to estimate the effect size of this same contrast (i.e., estimate the replicability of the contrast effect). Conventions are the same as in Figure 5, with one exception: in panels A and C, parts of the *y*-axis at the top or bottom have been cut out (marked by two parallel horizontal tick marks) in order to stretch the bars more and accentuate differences across conditions when those appeared. In these visually edited cases, distance between the most extreme 1-2 data points and their corresponding bars are not at scale. Those data points are colored in gray. Differences between the Lexico-semantic (red) and Syntactic (blue) conditions are marked with *'s.

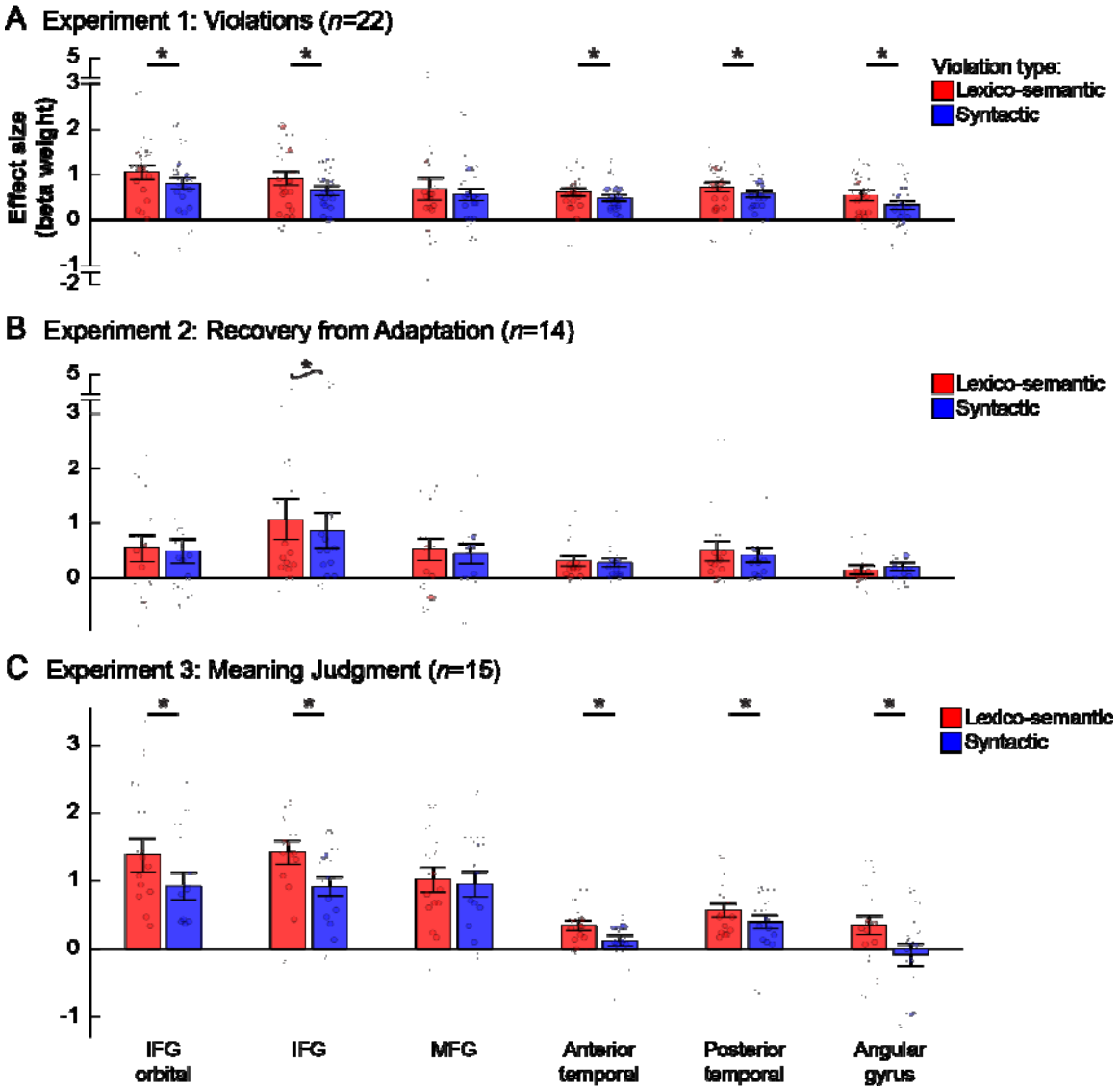


Figure 7: Responses in fROIs defined by the Lexico-semantic>Syntactic contrast to the critical conditions in Experiments 1-3. This figure depicts data from a parallel analysis to that depicted in Figure 6; here, participant-specific fROIs were defined as the top 10% of voxels showing the strongest Lexico-semantic>Syntactic contrast effect in the corresponding experiment, and the size of this contrast was then estimated in an independent part of the data (this is the opposite contrast to the one used in Figure 6). Conventions are the same as in Figures 5,6, with the addition of the following: non-significant effects with $p<0.10$ are marked with *s above tildes.

Discussion

Summary of the findings and their relationship to earlier proposals of the neural architecture of language

Across three fMRI experiments that used paradigms from the prior language literature but relied on a more sensitive and rigorous analytic approach—functional localization at the

individual-subject level (Saxe et al., 2006; Fedorenko et al., 2010; Nieto-Castañón & Fedorenko, 2012)—we searched for syntactic selectivity within the fronto-temporal language network, asking whether any part of this network responds reliably more strongly to syntactic than lexico-semantic processing. No such selectivity was found. In particular, no language fROI, defined with a robust language localizer contrast (Fedorenko et al., 2010), showed stronger responses to syntactic processing than lexico-semantic processing in any experiment (**Figure 5**). Even sets of voxels defined by their stronger responses to syntactic than lexico-semantic conditions using half of the data did not show a replicable Syntactic>Lexico-semantic effect in the other half of the data (**Figure 6**), with one exception of a small difference between the Syntactic and Lexico-semantic conditions in one experiment that did not survive correction for the number of regions.

Importantly, this failure to uncover any syntactically-selective regions/voxels within the language network is not due to lack of power. Specifically, the lack of syntactic selectivity stands in sharp contrast to a) sensitivity to both lexico-semantic and syntactic manipulations—in at least a subset of the language fROIs—relative to the control conditions, where present (in Experiments 1 and 2) (**Figure 5**, **Table 3**); and b) stronger responses to lexico-semantic than syntactic conditions, replicable across some experiments (Experiments 1 and 3) and analyses (**Figures 5** and **7**, **Tables 3-4**). And although the lack of syntactic selectivity in the current study appears to run contrary to earlier brain imaging reports of dissociable effects of syntactic and semantic processing (e.g., Dapretto & Bookheimer, 1999; Embick et al., 2000; Kuperberg et al., 2000; Ni et al., 2000; Newman et al., 2001; Kuperberg et al., 2003; Noppeney & Price, 2004; Cooke et al., 2006; Friederici et al., 2010; Glaser et al., 2013; Schell et al., 2017, among others), none of those earlier studies had compellingly established selectivity for syntactic processing in a robust (across multiple sets of participants/materials) and generalizable (across diverse manipulations that aim to isolate the same cognitive process, and relative to diverse control conditions) way, as discussed in the Introduction.

The current results—along with earlier results from paradigms that have varied the presence/absence of lexico-semantic vs. syntactic information in the linguistic signal (e.g., Fedorenko et al., 2010; 2012a; 2016; Bedny et al., 2011; Pallier et al., 2011; Bautista & Wilson, 2016)—pose a challenge for proposals of the neural architecture of language that postulate syntax- or combinatorics-selective brain regions (e.g., Grodzinsky & Santi, 2008; Baggio and Hagoort, 2011; Friederici, 2011, 2012; Tyler et al., 2011; Duffau et al., 2014; Ullman, 2016; Matchin & Hickok, 2019; Pylkkanen, 2019). Across such proposals (see next section), a syntax-/combinatorics-selective component is argued to not support the storage and processing of individual word meanings, as illustrated in the architectures in **Figure 1a-d**. In contrast to these proposals, it appears that any brain region / set of voxels within the language network that shows sensitivity to syntactic manipulations also shows sensitivity to manipulations targeting the processing of individual word meanings.

Not all cognitive neuroscience proposals of the language architecture have postulate a distinction between syntactic and semantic representations/processing, or between

combinatorial processing and stored knowledge representations (e.g., Bornkessel-Schlesewsky & Schlesewsky, 2009; Bornkessel-Schlesewsky et al., 2015). For example, Bornkessel-Schlesewsky et al. (2015) have suggested that the language network, as a whole, supports composition—combining smaller linguistic units into larger ones—in the service of meaning extraction (see also Mollica et al., 2020, for further empirical support of this idea). The current results align well with these proposals.

The bias of the language network toward lexico-semantic processing

In two of our experiments, lexico-semantic conditions elicited numerically, and sometimes reliably, stronger responses than syntactic conditions in many of the language fROIs. Furthermore, in Experiment 1, unlike the lexico-semantic violation condition, the syntactic violation condition did not elicit a response that was higher than a low-level (font violation) condition in the language regions. These overall stronger responses to semantic than syntactic conditions are in line with two prior findings. First, using multivariate analyses, we have previously found that lexico-semantic information is represented more robustly than syntactic information in the language system (Fedorenko et al., 2012a). In particular, pairs of conditions that differ in whether or not they contain lexico-semantic information (e.g., sentences vs. Jabberwocky sentences, or lists of words vs. lists of nonwords) are more robustly dissociable in the fine-grained patterns of activity than pairs of conditions that differ in whether or not they are structured (e.g., sentences vs. lists of words, or Jabberwocky sentences vs. lists of nonwords). And second, in ECoG, we observed reliably stronger responses to conditions that only contain lexico-semantic information (word lists) than conditions that only contain syntactic information (Jabberwocky) in many language-responsive electrodes (Fedorenko et al., 2016), but no electrodes showed the opposite pattern. Along with the current study, these results demonstrate that the magnitude and spatial organization of responses in the human language network are determined more by meaning than structure.

This bias toward lexico-semantic processing, fits with the view that the goal of language is communication, i.e., the transfer of *meanings* across minds (e.g., Hurford, 1998, 2007; Goldberg, 2006; Jackendoff, 2011; Kirby et al., 2015; Gibson et al., 2018; Hahn et al., 2020), and with the fact that most of our knowledge of language has to do with lexical semantics (word meanings), with only a small number of bits needed to store all of our syntactic knowledge (Mollica and Piantadosi, 2019). And it is not consistent with syntax-centric views of language (e.g., Chomsky and DiNozzi, 1972; Pinker, 1995; Hauser et al., 2002; Friederici et al., 2006; Berwick et al., 2013; Friederici, 2018).

One implication of these, and earlier behavioral, results discussed in the Introduction is that artificial grammar learning and processing paradigms (e.g., Reber, 1967)—where structured sequences of meaningless units (e.g., syllables) are used in an attempt to approximate human syntax (e.g., Friederici et al., 2006; Petersson et al., 2010; Wang et al., 2015)—have limited utility for understanding human language, given that syntactic representations and processing seem to be inextricably linked with representations of linguistic meaning (see also Fedor et al., 2012).

Limitations of scope

It is worth acknowledging the limitations of the current study. ***First***, perhaps some *cells / circuits / cortical-layer-specific areas* are selective for (some aspect) of syntactic/combinatorial processing (e.g., the architecture in **Figure 1e**), but we are not able to detect this selectivity due to the relatively coarse spatial resolution of our method. This possibility is hard to rule out without resorting to approaches like single-cell recordings (e.g., Engel et al., 2005; Mukamel & Fried, 2012) or laminar imaging (e.g., Norris & Polimeni, 2019). However, it is worth noting that at least for the paradigm that varies the presence/absence of lexico-semantic vs. syntactic information in the linguistic signal, the results from an ECoG study (Fedorenko et al., 2016)—where the spatial resolution is substantially higher than in fMRI—closely mirrored the fMRI results (e.g., Fedorenko et al., 2010).

Another possibility is that no single brain region is selective for syntactic/combinatorial processing, but inter-region *interaction/synchronization*—perhaps restricted to particular frequency bands (e.g., Giraud & Poeppel, 2012; Meyer, 2017; Martin & Doumas, 2019)—is critical for syntactic structure building. Some have argued that the arcuate / superior longitudinal fasciculus—the dorsal tract that connects posterior temporal and inferior frontal language areas—is critical for syntactic processing (e.g., Friederici, 2009; Brauer et al., 2011; Papoutsi et al., 2011; Wilson et al., 2011). However, the *selectivity* of this tract for syntactic/combinatorial processing is unclear, as it has also been implicated in non-syntactic computations, including, most commonly, articulation (e.g., Duffau et al., 2003; Hickok & Poeppel, 2007; Rauscheker & Scott, 2009), but also aspects of semantic processing (e.g., Glasser & Rillings, 2008). Thus, we would argue that, at present, no unequivocal evidence of syntax selectivity exists for inter-regional connections either.

Second, our research question focused on the "core" fronto-temporal language network, consisting of regions on the lateral surfaces of left frontal and left temporal cortex (e.g., Fedorenko & Thompson-Schill, 2014). Some areas *outside of this network's boundaries* have been implicated in syntactic processing, including, for example, parts of the basal ganglia (e.g., Ullman, 2001, 2004; cf. Grossman et al., 2002; Longworth et al., 2005) or the cerebellum (see Marien et al., 2014 for a review). However, we would argue that, similar to the cortical language regions, noone has compellingly established selectivity for syntactic processing for any brain region outside of the core language network.

Of relevance to this question is a claim that a region residing in or around Broca's area supports abstract hierarchical structure processing across domains, including language, arithmetic, music, and action observation/planning (e.g., Koechlin & Jubault, 2006; Tettamanti & Weniger, 2006; Fadiga et al., 2009; Fitch & Martins, 2014). This claim does not find empirical support. In particular, the part of Broca's area that responds to the presence of structure in language (e.g., showing stronger responses to structured linguistic stimuli, like sentences, than to lists of unconnected words) is highly selective for language relative to non-linguistic tasks, including ones that involve hierarchical structure and/or recursion, like arithmetic and music (e.g., Fedorenko et al., 2011; Monti

et al., 2012; see Fedorenko & Varley, 2016, for a review). These results make sense given the strong links between linguistic structure and meaning discussed in this manuscript: in other words, given what we now know about linguistic syntax, the idea that it would be supported by a mechanism that is not sensitive to the nature of the representation does not seem tenable.

One possible explanation for some of the findings that have been used as evidence for a domain-general structure processor is that those manipulations activated a domain-general component of Broca's area. This component belongs to the domain-general multiple demand (MD) network implicated in executive control and goal-directed behaviors and robustly sensitive to effort (e.g., Duncan, 2010, 2013; Fedorenko et al., 2013). Manipulations of hierarchical complexity have often been confounded with difficulty, such that the more structurally complex conditions required greater cognitive effort. As a result, they would be likely to elicit responses in the MD network, including its inferior frontal component residing in Broca's area (see Fedorenko & Blank, in press, for more discussion). Although it is possible that outside of the domain of language—which appears to rely on domain-specific processing mechanisms (Fedorenko et al., 2011)—this component of the MD network, or the MD network as a whole, is important for structured behavior, we should keep in mind that this network is also sensitive to manipulations that don't involve structured/hierarchical representations (e.g., Crittenden & Duncan, 2012). The latter findings argue against the idea of complex syntactic operations being the core computation of the MD network.

Third, we have here focused on language comprehension. Could the architecture of language processing be different for language *production*? Although we plausibly access the same knowledge representations to interpret (comprehend) and generate (produce) linguistic utterances—in line with substantial overlap that has been observed between comprehension and production in fMRI (e.g., Menenti et al., 2011; Silbert et al., 2014)—the computational demands of language production differ from those of language comprehension. In particular, the goal of comprehension is to infer the intended meaning from the linguistic signal, and abundant evidence now suggests that the representations we extract and maintain during comprehension are probabilistic and noisy (e.g., Ferreira et al., 2002; Levy et al., 2008; Gibson et al., 2013). In contrast, in production, the target meaning is (typically) clear and precise, and the goal is to express that particular meaning. To do so, we have to utter a precise sequence of words where each word takes a particular morpho-syntactic form, and the words appear in a particular order. This pressure for linearization of words, morphemes, and sounds might lead to a clearer temporal, and perhaps spatial, segregation among the different stages of the production process compared to comprehension (e.g., Hagoort & Indefrey, 2014), and/or require additional mechanisms implemented in regions that do not support comprehension. Indeed, recent evidence from intracranial stimulation suggests that a small region in posterior superior temporal cortex may be selective for encoding and enacting morpho-syntactic inflections (Lee, Fedorenko, et al., 2018; see Fedorenko et al., 2018, for discussion). It is therefore possible that some aspects of language production are implemented in focal and functionally selective regions. However, this conjecture remains to be evaluated further in future work.

Finally, in the current study, we investigated the relationship between syntactic and lexico-semantic processing in a single Germanic language: English. Some of the theoretical linguistic work, experimental psycholinguistic work, and computational modeling work have spanned multiple languages (e.g., Norcliffe et al., 2015). However, much/most cognitive neuroscience research has been conducted on English and a handful of other languages/families (e.g., see Bornkessel-Schlesewsky & Schlesewsky, 2016 for discussion). Thus, the conclusions drawn here remain to be generalized to typologically diverse languages.

Other findings that have been interpreted as evidence for syntax selectivity

Three other lines of research—on phenomena that are, or have been, taken as strong evidence for syntax selectivity—deserve discussion. ***First***, the early ERP literature on language processing appeared to have provided evidence of distinct components associated with lexico-semantic processing (N400; Kutas & Hilliyard, 1980) vs. with syntactic processing (P600; Osterhout & Holcomb, 1992; Hagoort et al., 1993). However, the interpretation of the P600 as an index of syntactic processing has been challenged from the earliest days following its discovery (e.g., Coulson et al., 1998), and the current dominant interpretation of this component is as a domain-general error detection or correction signal (e.g., Kolk & Chwila, 2007; Vissers et al., 2007; van de Meerendonk et al., 2010; Sassenhagen et al., 2014; Ryskin et al., 2020b). Some other, earlier, ERP components (e.g., eLAN; Friederici, 2002) have been argued to index syntactic processes. However, the robustness and nature of these components have been questioned (Steinhauer & Drury, 2012), and the interpretation that seems to capture the empirical findings best has to do with violations of word-form expectations rather than syntactic parsing per se (e.g., Dikker et al., 2009, 2010; Rosenfelt et al., 2009, 2011).

Second, syntactic priming—re-use of a syntactic frame based on recent linguistic experiences (Bock, 1986; see Pickering and Ferreira, 2008; Branigan & Pickering, 2016 for reviews)—has often been cited as evidence of abstract syntactic representations independent of meaning (e.g., Bock & Loebell, 1990), including in relatively recent cognitive neuroscience papers (Pallier et al., 2011). However, a large body of work has now established that the effect is strongly modulated by lexical overlap (e.g., Mahowald et al. 2016; Scheepers et al., 2017; Ziegler et al., 2019) and driven by the meaning-related aspects of the utterance (e.g., Hare & Goldberg, 1999; Cai et al., 2012; Ziegler & Snedeker, 2018; Ziegler et al., 2018).

And ***third***, a class of linguistic phenomena known as “syntactic islands” (Ross, 1967) have been argued to be due to abstract properties of syntactic structures unrelated to meaning. In particular, some structures are disfavored when a phrase is “extracted” from its “canonical” structural location in a sentence (e.g., Chomsky, 1973; Schütze et al., 2015). However, this interpretation has been challenged: in particular, some researchers (e.g., Erteschik-Shir, 1973; Kuno, 1987; Goldberg, 2013) have argued that semantic and discourse factors can explain these phenomena (see Abeillé et al., in press, for empirical support).

Beyond syntax and semantics: dissociations of other linguistic processes

Language processing encompasses a broad array of computations in both comprehension and production, and some aspects of language are robustly dissociable and supported by distinct sets of brain regions. Here, we have argued that during language comprehension the mechanisms that process the *structure* of sentences are also deeply sensitive to the meanings of *individual words*. Is the reverse also true? Are there any brain areas that process individual word meanings but are not sensitive to syntactic processing?

One area that deserves a mention lies in the left temporal pole, extending onto the lateral and ventral surface of the temporal lobe. According to one influential hypothesis, this area has been implicated in lexical retrieval (e.g., Damasio et al., 2004; Drane et al., 2008; Grabowski et al., 2001; Tranel, 2006, 2009; Mesulam et al., 2013, 2015). According to another hypothesis, motivated chiefly by investigations of semantic dementia (Gorno-Tempini et al., 2011), this region has been linked to general object knowlege (e.g., Lambon-Ralph et al., 2001; Rogers et al., 2004, 2006; Patterson et al., 2007). Critically, syntactic abilities in patients with damage to this area appear to be relatively preserved (e.g., Mesulam et al., 2013, 2015). This area overlaps with our LAntTemp fROI, but extends beyond it. In our experiments, the LAntTemp language fROI showed stronger responses during lexico-semantic than syntactic processing in two of the three experiments (**Figure 5**). However, this region still responds reliably to syntactic processing in at least some manipulations: for example, it responds more strongly to a) structured but meaningless Jabberwocky sentences compared to lists of unconnected nonwords (Fedorenko et al., 2010; see Mollica et al., in prep., for a replication), and b) structurally more complex sentences with object-extracted relative clauses compared to those with subject-extracted relative clauses (Blank et al., 2016). Further, based on evidence from MEG, parts of the left anterior temporal lobe have been implicated in semantic composition, above and beyond the processing of single words (e.g., Bemis & Pylkkanen, 2011; see Pylkkanen, 2019 for a review). The brain imaging evidence therefore suggests that this region is engaged in some syntax-/combinatorics-relevant processes, in addition to lexico-semantic processing. But because the temporal pole and the anterior ventral temporal cortex are challenging to study with fMRI given the signal dropout due to proximity to air-filled sinuses (e.g., Devlin et al., 2000), it is possible that we are missing some areas—anterior to our LAntTemp fROI and/or on the ventral surface of the temporal lobe—that are truly selective for lexico-semantic/conceptual over syntactic processing. The patient evidence mentioned above implies the existence of such areas, although the evidence is not unequivocal (e.g., Bi et al., 2010). More work is therefore needed to functionally characterize left anterior temporal areas and their relationship with the core fronto-temporal language network.

Going beyond syntax and semantics, lower-level speech perception and reading processes as well as speech production (articulation) recruit areas that are robustly distinct from the high-level areas that we focused on here. In particular, speech perception recruits parts of the auditory cortex in the superior temporal gyrus and sulcus (e.g., Scott et al., 2000, Mesgarani et al., 2014; Overath et al., 2015), and these areas are highly selective for

speech over many other types of auditory stimuli (Norman-Haignere et al., 2015). Reading recruits a small area on the ventral surface of the temporal lobe (see McCandliss et al., 2003, for a review), and this "visual word-form area" is highly selective for letters in a familiar script over a broad range of other visual stimuli (Baker et al., 2007; Hamame et al., 2013). And articulation draws on a set of areas, including portions of the precentral gyrus, supplementary motor area, inferior frontal cortex, superior temporal cortex, and cerebellum (e.g., Wise et al., 1999; Bohland and Guenther, 2006; Eickhoff et al., 2009; Basilakos et al., 2017).

On the other end of linguistic processes, discourse-level processing draws on areas distinct from those that support word and sentence-level comprehension (e.g., Ferstl & von Cramon, 2001; Lerner et al., 2011; Jacoby et al., 2018; Blank & Fedorenko, 2020), and aspects of non-literal language have been argued to draw on brain regions in the right hemisphere (e.g., Joanette et al., 1990) and on the system that supports social cognition (e.g., Kline et al., 2018; Hagoort, 2019).

Thus, many aspects of language are robustly dissociable, in line with distinct patterns of deficits reported in the aphasia literature (e.g., Goodglass, 1993). However, syntactic and lexico-semantic processing do not appear to be separable during language comprehension. Some brain areas not easily accessible to fMRI may be selective for lexico-semantic processing, as discussed above, but no area within the language network appears to be selective for syntactic processing based on both brain imaging studies and patient investigations.

Concluding remarks

To conclude, across three fMRI experiments, we found robust responses to both lexico-semantic and syntactic processing throughout the language network, with generally stronger responses to lexico-semantic processing, and no regions, or even sets of non-contiguous voxels within those regions, that respond reliably more strongly to syntactic processing than lexico-semantic processing. These results constrain the space of possible neural architectures of language. In particular, they rule out architectures that postulate a distinct region (or set of regions) that selectively supports syntactic/combinatorial processing (i.e., architectures shown in **Figures 1a-d**). These findings, illuminating how minds are instantiated in brains, are mirrored by studies of how minds are implemented in machines, where modern-day connectionist networks achieve remarkable performance on a wide variety of language tasks (e.g., Mikolov et al., 2010; Sutskever et al., 2014; Bahdanau et al., 2016), including those that involve complex syntactic phenomena (e.g., Linzen et al., 2016; Gulordava et al., 2018; Futrell et al., 2018, 2019; Prasad et al., 2019; Wilcox et al., 2019a,b), apparently without a clearly separable syntax-selective mechanism. Taking all the available data into consideration, it therefore seems that a cognitive architecture whereby *syntactic processing is not separable from the processing of individual word meanings* is most likely.

Author contributions

EF conceived and designed the study. All authors collected the data and performed first-level data analyses. IB performed second-level analyses. IB, EF, and MS created the figures. EF drafted the manuscript and IB provided critical revisions, with MS and ZM providing additional comments. All authors approved the final version of the manuscript.

Acknowledgements

We would like to acknowledge the Athinoula A. Martinos Imaging Center at the McGovern Institute for Brain Research at MIT, and its support team (Steve Shannon and Atsushi Takahashi). We thank former and current EvLab members (especially Zuzanna Balewski and Brianna Pritchett) for their help with data collection, Gina Kuperberg for providing the materials used in adapted form in Experiment 1, Michael Behr for creating the script for Experiment 1, Zuzanna Balewski for help with creating the materials and script for Experiment 2, Josef Affourtit for helping put together the OSF page, Nancy Kanwisher for discussions of the experimental design for all three experiments, and Inbal Arnon, Yonatan Belinkov, Ted Gibson, Adele Goldberg, Maryellen MacDonald, and Jayden Ziegler for comments on the manuscript. We also thank the audience at the 2017 CUNY Sentence Processing conference (Cambridge, MA) for feedback, as well as three anonymous reviewers whose comments helped to greatly improve the manuscript. EF was supported by NIH awards R00-HD057522, R01-DC016607, R01-DC016950, by a grant from the Simons Foundation to the Simons Center for the Social Brain at MIT, and by support from the Brain and Cognitive Sciences Department and the McGovern Institute for Brain Research at MIT.

Conflict of interest

The authors declare no competing financial interests.

Table 1. Timing parameters for the different versions of the language localizer task.

	Version		
	A	**B**	**C**
Number of participants	30	14	5
Task: Button Press or Memory?	BP	M	M
Words / nonwords per trial	12	12	12
Trial duration (ms)	6,000	6,000	6,000
Fixation	100	---	---
Presentation of each word / nonword	450	350	350
Fixation	---	300	300
Button icon / Memory probe	400	1,000	1,000
Fixation	100	500	500
Trials per block	3	3	3
Block duration (s)	18	18	18
Blocks per condition (per run)	8	8	6
Conditions	Sentences Nonwords	Sentences Nonwords	Sentences Nonwords Word-lists*
Fixation block duration (s)	14	18	18
Number of fixation blocks	5	5	4
Total run time (s)	358	378	396
Number of runs	2	2	2-3

* Used for an unrelated experiment.

Table 2. Stimulus and Procedure details for each experiment.

	Experiment 1: Violations	**Experiment 2: Recovery from Adaptation**	**Experiment 3: Meaning Judgment**[e]
Item description	10-word sentence	Two transitive sentences A & B	Two sentences A & B
Base items	240[c]	144 × 3 sets	80
Versions per base item[a]	4 Lexico-semantic Syntactic Font Control	6 Critical[d]: A then B / B then A Same: A then A / B then B Different: X then A / Y then B (X,Y: another item)	4 Lexico-semantic, same Lexico-semantic, different Syntactic, same Syntactic, different
Unique trials	240 × 4 = 960	(144 × 3) × 6 = 2592	80 × 4 = 320
Experimental lists	4 (240 trials each, 60 per version)	6 (432 trials each, 144 per subset, 24 per version per subset)	4 (80 trials each, 20 per version)
Trial breakdown in a single list	60 Lexico-semantic 60 Syntactic 60 Font 60 Control (+10 fillers each)	48 Critical Lexico-semantic 48 Critical Syntactic 48 Critical Global meaning 48 × 3 = 144 Same 48 × 3 = 144 Different	40 Lexico-semantic 40 Syntactic
Runs per list	5 (sometimes 4)	6 (sometimes 5)	2
Trials per version per run	12 (+2 fillers)	4 (24 per subset)	10
Trial length (s)[b]	6	4	6
Fixation per run (s)	72	32	120
Run duration (s)	408 (6min 48s)	320 (5min 20s)	360 (6min)
No. condition orders	10	6	8
Subjects per list	4-6	2-4	3-4

[a] See Figure 2.
[b] See Figure 3.
[c] Verbs did not repeat across items, except for "*practice*" and "*read*", which were each used twice. 179 base items were adapted from Kuperberg et al. (2003).
[d] Each of the three sets of base items had a different critical condition: Lexico-semantic, Syntactic, or Global meaning (see Figure 2).
[e] Stimuli and procedure followed the design of Dapretto & Bookheimer (1999).

Table 3. Responses of language fROIs in Experiments 1-3: mean PSC with standard error (by participants), effect size (Cohen's *d*), *t*-value, and *p*-value.[a]

fROI (size)	Exp. 1: violations					Exp. 2: adaptation recovery				Exp. 3: meaning judgments		
	LexSem *vs.* Control	Synt *vs.* Control	LexSem *vs.* Font	Synt *vs.* Font	LexSem *vs.* Synt	Diff *vs.* Same	LexSem *vs.* Same	Synt *vs.* Same	LexSem *vs.* Synt	LexSem *vs.* Baseline	Synt *vs.* Baseline	LexSem *vs.* Synt
IFGorb (37 voxels)	0.44 (±0.07) *d*=1.37 *t*=6.40 $p<10^{-5}$	0.19 (±0.07) *d*=0.61 *t*=2.85 *p*=.0096	0.38 (±0.08) *d*=0.97 *t*=4.54 *p*=.0002	0.13 (±0.09) *d*=0.31 *t*=1.44 *p*=.16	0.25 (±0.08) *d*=0.64 *t*=3.01 *p*=.0067	0.09 (±0.05) *d*=0.52 *t*=1.93 p=.075	0.21 (±0.08) *d*=0.73 *t*=2.72 *p*=.18	0.13 (±0.04) *d*=0.83 *t*=3.10 *p*=.0084	0.07 (±0.07) *d*=0.28 *t*=1.05 *p*=0.31	1.46 (±0.21) *d*=1.80 *t*=6.96 $p<10^{-5}$	1.18 (±0.21) *d*=1.46 *t*=5.67 $p<10^{-4}$	0.28 (±0.12) *d*=0.59 *t*=2.28 *p*=.039
IFG (74 voxels)	0.46 (±0.08) *d*=1.17 *t*=5.51 $p<10^{-4}$	0.22 (±0.06) *d*=0.72 *t*=3.37 *p*=.0029	0.35 (±0.09) *d*=0.88 *t*=4.12 *p*=.0005	0.12 (±0.09) *d*=0.28 *t*=1.30 *p*=.21	0.24 (±0.08) *d*=0.64 *t*=3.02 *p*=.0065	0.21 (±0.05) *d*=1.08 *t*=4.02 *p*=.0014	0.31 (±0.08) *d*=1.02 *t*=3.80 *p*=.0022	0.22 (±0.08) *d*=0.74 *t*=2.77 *p*=.016	0.09 (±0.08) *d*=0.29 *t*=1.09 *p*=0.29	1.48 (±0.21) *d*=1.78 *t*=6.89 $p<10^{-5}$	1.32 (±0.20) *d*=1.72 *t*=6.68 $p<10^{-4}$	0.16 (±0.12) *d*=0.34 *t*=1.30 *p*=.21
MFG (46 voxels)	0.22 (±0.08) *d*=0.59 *t*=2.78 *p*=.011	0.14 (±0.09) *d*=0.33 *t*=1.56 *p*=.13	0.16 (±0.09) *d*=0.39 *t*=1.81 p=.084	0.08 (±0.09) *d*=0.17 *t*=0.82 *p*=.042	0.09 (±0.12) *d*=0.16 *t*=0.74 *p*=0.47	0.23 (±0.07) *d*=0.93 *t*=3.47 *p*=.0042	0.32 (±0.10) *d*=0.90 *t*=3.35 *p*=.0052	0.32 (±0.09) *d*=0.92 *t*=3.46 *p*=.0043	0.00 (±0.08) *d*=0.01 *t*=0.04 *p*=0.97	1.39 (±0.22) *d*=1.64 *t*=6.36 $p<10^{-4}$	1.41 (±0.29) *d*=1.25 *t*=4.85 *p*=.0002	-0.01 (±0.12) *d*=-0.03 *t*=-0.12 *p*=.90
AntTemp (162 voxels)	0.13 (±0.03) *d*=0.83 *t*=3.90 *p*=.0008	-0.02 (±0.02) *d*=-0.15 *t*=-0.71 *p*=.49	0.14 (±0.04) *d*=0.85 *t*=3.99 *p*=.0007	-0.01 (±0.04) *d*=-0.04 *t*=-0.18 *p*=0.86	0.15 (±0.04) *d*=0.87 *t*=4.10 *p*=.0005	0.15 (±0.04) *d*=0.94 *t*=3.51 *p*<.0038	0.17 (±0.05) *d*=0.95 *t*=3.56 *p*=.0035	0.10 (±0.04) *d*=0.62 *t*=2.34 *p*=.036	0.07 (±0.05) *d*=0.38 *t*=1.43 *p*=0.17	0.54 (±0.11) *d*=1.26 *t*=4.87 *p*=.0002	0.36 (±0.11) *d*=0.87 *t*=3.36 *p*=.0047	0.19 (±0.06) *d*=0.81 *t*=3.14 *p*=.0072
PostTemp (294 voxels)	0.20 (±0.04) *d*=1.16 *t*=5.42 $p<10^{-4}$	0.07 (±0.04) *d*=0.38 *t*=1.80 *p*=.086	0.18 (±0.05) *d*=0.80 *t*=3.76 *p*=.0011	0.06 (±0.07) *d*=0.18 *t*=0.85 *p*=.40	0.12 (±0.06) *d*=0.45 *t*=2.13 *p*=.045	0.20 (±0.04) *d*=1.22 *t*=4.58 *p*=.0005	0.25 (±0.06) *d*=1.05 *t*=3.91 *p*=.0018	0.24 (±0.06) *d*=1.07 *t*=3.99 *p*=.0015	0.01 (±0.07) *d*=0.05 *t*=0.20 *p*=0.84	0.98 (±0.10) *d*=2.43 *t*=9.43 $p<10^{-6}$	1.00 (±0.14) *d*=1.85 *t*=7.16 $p<10^{-5}$	-0.02 (±0.08) *d*=-0.06 *t*=-0.22 *p*=.83
AngG (64 voxels)	0.19 (±0.06) *d*=0.73 *t*=3.40 *p*=.0027	-0.06 (±0.04) *d*=-0.30 *t*=-1.42 *p*=.17	0.22 (±0.06) *d*=0.73 *t*=3.44 *p*=.0024	-0.03 (±0.06) *d*=-0.12 *t*=-0.58 *p*=.57	0.25 (±0.06) *d*=0.95 *t*=4.48 *p*=.0002	0.07 (±0.05) *d*=0.43 *t*=1.59 *p*=.13	0.06 (±0.06) *d*=0.28 *t*=1.04 *p*=.32	0.11 (±0.05) *d*=0.58 *t*=2.16 p=.054	-0.05 (±0.07) *d*=-0.21 *t*=-0.79 *p*=0.44	0.80 (±0.17) *d*=1.18 *t*=4.59 *p*<.0004	0.61 (±0.23) *d*=0.69 *t*=2.67 *p*=.018	0.19 (±0.14) *d*=0.34 *t*=1.33 *p*=.20

[a] non-significant effects at α=0.05 (FDR-corrected for the number of regions) are shaded in grey; marginal effects (0.05<*p*<0.1) are shaded in light grey.

Table 4. Replicability (in left-out data) of the critical contrasts in Experiments 1-3.[a,b]

fROI (size)	**Exp. 1: violations**		**Exp. 2: adaptation recovery**		**Exp. 3: meaning judgments**	
	LexSem> Synt	Synt> LexSem	LexSem> Synt	Synt> LexSem	LexSem> Synt	Synt> LexSem
IFGorb (37 voxels)	0.24 (±0.08) *d*=0.75 *t*=3.54 *p*=.0010	0.03 (±0.08) *d*=0.07 *t*=0.35 *p*=.63	0.05 (±0.12) *d*=0.11 t=0.40 *p*=.35	0.00 (± 0.04) *d*=-0.01 *t*=-0.03 *p*=.49	0.46 (±0.15) *d*=0.79 *t*=3.06 *p*=.004	-0.10 (±0.10) *d*=-0.25 *t*=-0.95 *p*=.82
IFG (74 voxels)	0.26 (±0.11) *d*=0.59 *t*=2.76 *p*=.0059	0.00 (±0.07) *d*=-0.01 *t*=-0.02 *p*=.51	0.21 (±0.15) *d*=0.37 *t*=1.40 *p*=.09	-0.03 (±0.06) *d*=-0.13 *t*=-0.47 *p*=.68	0.5 (±0.14) *d*=0.94 *t*=3.63 *p*=.0014	-0.08 (±0.08) *d*=-0.24 *t*=-0.93 *p*=.81
MFG (46 voxels)	0.13 (±0.16) *d*=0.20 *t*=0.93 *p*=.18	0.08 (±0.11) *d*=-0.15 *t*=0.72, *p*=.24	0.08 (±0.10) *d*=0.20 *t*=0.74 *p*=.24	-0.02 (±0.07) *d*=-0.07 *t*=-0.25 *p*=.60	0.07 (±0.09) *d*=0.19 *t*=0.72 *p*=.24	0.03 (±0.13) *d*=0.05 *t*=0.21 *p*=.42
AntTemp (162 voxels)	0.13 (±0.04) *d*=0.74 *t*=3.50 *p*=.0011	-0.02 (±0.03) *d*=-0.10 *t*=-0.48 *p*=.68	0.03 (±0.06) *d*=0.15 *t*=0.57 *p*=.29	-0.01 (±0.03) *d*=-0.05 *t*=-0.19 *p*=.57	0.22 (±0.06) *d*=0.97 *t*=3.77 *p*=.0010	-0.06 (±0.04) *d*=-0.44 *t*=-1.7 *p*=.94
PostTemp (294 voxels)	0.15 (±0.05) *d*=0.63 *t*=2.97 *p*=.0036	0.00 (±0.05) *d*=-0.02 *t*=0-.08 *p*=.53	0.08 (±0.08) *d*=0.26 *t*=0.96 *p*=.18	0.08 (±0.03) *d*=0.80 *t*=2.99 *p*=.0051	0.17 (±0.07) *d*=0.67 *t*=2.59 *p*=.011	0.06 (±0.07) *d*=0.21 *t*=0.83 *p*=.21
AngG (64 voxels)	0.21 (±0.06) *d*=0.71 *t*=3.34 *p*=.0015	0.05 (±0.08) *d*=-0.12 *t*=-0.58 *p*=.71	-0.05 (±0.05) *d*=-0.26 *t*=-0.99 *p*=.83	0.07 (±0.07) *d*=0.26 *t*=0.96 *p*=.18	0.43 (±0.16) *d*=0.68 *t*=2.59 *p*=.0099	-0.03 (±0.10) *d*=-0.08 *t*=-0.31 *p*=.62

[a] The Lexico-semantic>Syntactic effect is tested in Lexico-semantic>Syntactic fROIs, and the Syntactic>Lexico-semantic effect is tested in Syntactic>Lexico-semantic fROIs.

[b] Conventions are the same as in Table 3.

References

Abeillé, A., Hemforth, B., Winckel, E., & Gibson, E. (in press). Extraction from subjects: differences in acceptability depend on the discourse function of the construction. *Cognition.*

Aguirre, G. K., & Farah, M. J. (1998). Human visual object recognition: What have we learned from neuroimaging?. *Psychobiology*, *26*(4), 322-332.

Ambridge, B., Pine, J. M., Rowland, C. F., Freudenthal, D., & Chang, F. (2014). Avoiding dative overgeneralisation errors: semantics, statistics or both? *Language Cognition and Neuroscience*, *29*(2), 218-243.

Ambridge, B. (2018). Against stored abstractions: A radical exemplar model of language acquisition. *First Language*, 0142723719869731.

Axelrod, V., Bar, M., Rees, G., & Yovel, G. (2015). Neural Correlates of Subliminal Language Processing. *Cerebral Cortex*, *25*(8), 2160–2169.

Arnon, I., & Snider, N. (2010). More than words: Frequency effects for multi-word phrases. *Journal of Memory and Language*, *62*(1), 67-82.

Baggio, G., & Hagoort, P. (2011). The balance between memory and unification in semantics: A dynamic account of the N400. *Language and Cognitive Processes*, *26*(9), 1338-1367.

Bahdanau, D., Chorowski, J., Serdyuk, D., Brakel, P., & Bengio, Y. (2016). End-to-end attention-based large vocabulary speech recognition. In *Acoustics, Speech and Signal Processing (ICASSP), 2016 IEEE International Conference* (pp. 4945-4949).

Baker, C. I., Liu, J., Wald, L. L., Kwong, K. K., Benner, T., & Kanwisher, N. (2007). Visual word processing and experiential origins of functional selectivity in human extrastriate cortex. *Proceedings of the National Academy of Sciences*, *104*(21), 9087–9092.

Barlow, M., & Kemmer, S. (Eds.). (2000). Usage Based Models of Language.

Basilakos, A., Smith, K. G., Fillmore, P., Fridriksson, J., & Fedorenko, E. (2018). Functional characterization of the human speech articulation network. *Cerebral Cortex*, *28*(5), 1816-1830.

Bates, E., Bretherton, I., & Snyder, L. (1988). From first words to grammar. New York: Cambridge University Press.

Bates, E., Dale, P. S., & Thal, D. (1995). Individual differences and their implications for theories of language development. *The handbook of child language* (Fletcher, P. and MacWhinnery, B., eds.), 96-151. Basil Blackwell.

Bates, E., & Goodman, J.C. (1997). On the inseparability of grammar and the lexicon: Evidence from acquisition, aphasia and real time processing. *Language and Cognitive Processes*, 12, 507-586.

Bautista, A., & Wilson, S. M. (2016). Neural responses to grammatically and lexically degraded speech. *Language Cognition and Neuroscience*, *31*(4), 567-574.

Bedny, M., Pascual-Leone, A., Dodell-Feder, D., Fedorenko, E., & Saxe, R. (2011). Language processing in the occipital cortex of congenitally blind adults. *Proceedings of the National Academy of Sciences*, *108*(11), 4429-4434.

Bekinschtein, T. A., Davis, M. H., Rodd, J. M., & Owen, A. M. (2011). Why Clowns Taste Funny: The Relationship between Humor and Semantic Ambiguity. *Journal of Neuroscience*, *31*(26), 9665–9671.

Bemis, D. K., & Pylkkanen, L. (2011). Simple Composition: A Magnetoencephalography Investigation into the Comprehension of Minimal Linguistic Phrases. *Journal of Neuroscience*, *31*(8), 2801–2814.

Ben-Shachar, M., Hendler, T., Kahn, I., Ben-Bashat, D., & Grodzinsky, Y. (2003). The neural reality of syntactic transformations: Evidence from functional magnetic resonance imaging. *Psychological Science*, *14*(5), 433-440.

Benjamini, Y., & Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *Annals of Statistics*, 29(4), 1165-1188.

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

Bi, Y., Wei, T., Wu, C., Han, Z., Jiang, T., & Caramazza, A. (2011). The role of the left anterior temporal lobe in language processing revisited: evidence from an individual with ATL resection. *Cortex*, *47*(5), 575-587.

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex, 19*(12), 2767–2796.

Bilenko, N. Y., Grindrod, C. M., Myers, E. B., & Blumstein, S. E. (2008). Neural correlates of semantic competition during processing of ambiguous words. *Journal of Cognitive Neuroscience, 21*(5), 960-975.

Blank, I., Kanwisher, N., & Fedorenko, E. (2014). A functional dissociation between language and multiple-demand systems revealed in patterns of BOLD signal fluctuations. *Journal of Neurophysiology, 112*(5), 1105–1118.

Blank, I. A., Kiran, S., & Fedorenko, E. (2017). Can neuroimaging help aphasia researchers? Addressing generalizability, variability, and interpretability. *Cognitive Neuropsychology, 34*(6), 377–393.

Blank, I., Balewski, Z., Mahowald, K., & Fedorenko, E. (2016). Syntactic processing is distributed across the language system. *NeuroImage, 127*, 307-323.

Bock, J. K. (1986). Meaning, sound, and syntax: Lexical priming in sentence production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*(4), 575-586.

Bock, K., & Loebell, H. (1990). Framing sentences. *Cognition, 35*(1), 1-39.

Bod, R. (1998). Beyond grammar: An experience-based theory of language. Stanford, CA: CSLI Publications.

Bod, R. (2006). Exemplar-based syntax: How to get productivity from examples. *The linguistic review*, *23*(3), 291-232.

Bohland, J. W., & Guenther, F. H. (2006). An fMRI investigation of syllable sequence production. *Neuroimage*, *32*(2), 821-841.

Bookheimer, S. (2002). Functional MRI of language: New approaches to understanding the cortical organization of semantic processing. *Annual Review of Neuroscience*, *25*, 151-188.

Bornkessel-Schlesewsky, I., Kretzschmar, F., Tune, S., Wang, L. M., Genc, S., Philipp, M., Roehm, D., Schlesewsky, M. (2011). Think globally: Cross-linguistic variation in electrophysiological activity during sentence comprehension. *Brain and Language*, *117*(3), 133-152.

Bornkessel-Schlesewsky, I., Schlesewsky, M., Small, S. L., & Rauschecker, J. P. (2015). Neurobiological roots of language in primate audition: Common computational properties. *Trends in Cognitive Sciences*, *19*(3), 142-150.

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2016). The importance of linguistic typology for the neurobiology of language. *Linguistic Typology*, *20*(3), 615-621.

Botvinik-Nezer, R., Holzmeister, F., Camerer, C. F., Dreber, A., Huber, J., Johannesson, M., ... & Avesani, P. (2019). Variability in the analysis of a single neuroimaging dataset by many teams. *bioRxiv*.

Braga, R. M., DiNicola, L. M., & Buckner, R. L. (2019). Situating the Left-Lateralized Language Network in the Broader Organization of Multiple Specialized Large-Scale Distributed Networks. *bioRxiv*.

Branco, P., Seixas, D., & Castro, S. L. (2020). Mapping language with resting-state functional magnetic resonance imaging: A study on the functional profile of the language network. *Human Brain Mapping*, *41*(2), 545-560.

Branigan, H. P., & Pickering, M. J. (2017). An experimental approach to linguistic representation. *Behavioral and Brain Sciences*, *40*.

Brauer, J., Anwander, A., & Friederici, A. D. (2011). Neuroanatomical Prerequisites for Language Functions in the Maturing Brain. *Cerebral Cortex*, *21*(2), 459-466.

Brett, M., Johnsrude, I. S., & Owen, A. M. (2002). The problem of functional localization in the human brain. *Nature Reviews. Neuroscience*, *3*(3), 243–249.

Bybee, J. L. (1985). *Morphology: A Study of the Relation between Meaning and Form* (Vol. 9). Amsterdam: John Benjamins Publishing Company.

Bybee, J. L. (1998). A Functionalist Approach to Grammar and Its Evolution. *Evolution of Communication*, *2*(2), 249-278.

Bybee, J. (2006). From usage to grammar: The mind's response to repetition. *Language*, *82*, 711-733.

Bybee, J. (2010). *Language, usage and cognition* (Vol. 98). Cambridge University Press.

Cai, Z. G. G., Pickering, M. J., & Branigan, H. P. (2012). Mapping concepts to syntax: Evidence from structural priming in Mandarin Chinese. *Journal of Memory and Language*, *66*(4), 833-849.

Chang, F., Bock, K., & Goldberg, A. E. (2003). Can thematic roles leave traces of their places? *Cognition*, *90*(1), 29-49.

Charniak, E. (1997). Statistical parsing with a context-free grammar and word statistics. *AAAI/IAAI, 2005*(598-603), 18.

Chee, M. W., Caplan, D., Soon, C. S., Sriram, N., Tan, E. W., Thiel, T., & Weekes, B. (1999). Processing of visually presented sentences in Mandarin and English studied with fMRI. *Neuron, 23*(1), 127-137.

Chomsky, N. (2002). *Syntactic structures*. Walter de Gruyter.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky, N., & Dinozzi, R. (1972). *Language and mind*: Harcourt Brace Jovanovich.

Chomsky, N. (1973). Conditions on transformations. *A festschrift for Morris Halle*.

Chomsky, N. (1995). Language and Nature. *Mind, 104*(413), 1–61. JSTOR.

Christiansen, M. H., & Arnon, I. (2017). More than words: The role of multiword sequences in language learning and use. *Topics in cognitive science, 9*(3), 542-551.

Clifton, C., Frazier, L., & Connine, C. (1984). Lexical expectations in sentence comprehension. *Journal of Memory and Language, 23*(6), 696.

Cooke, A., Grossman, M., DeVita, C., Gonzalez-Atavales, J., Moore, P., Chen, W., . . . Detre, J. (2006). Large-scale neural network for sentence processing. *Brain and Language, 96*(1), 14-36.

Costafreda, S. G., Fu, C. H., Lee, L., Everitt, B., Brammer, M. J., & David, A. S. (2006). A systematic review and quantitative appraisal of fMRI studies of verbal fluency: role of the left inferior frontal gyrus. *Human brain mapping, 27*(10), 799-810.

Coulson, S., King, J. W., & Kutas, M. (1998). Expect the unexpected: Event-related brain response to morphosyntactic violations. *Language and Cognitive Processes, 13*(1), 21-58.

Crittenden, B. M., & Duncan, J. (2012). Task difficulty manipulation reveals multiple demand activity but no frontal lobe hierarchy. *Cerebral Cortex, 24*(2), 532-540.

Culicover, P. W., & Jackendoff, R. (1999). The view from the periphery: The English comparative correlative. *Linguistic inquiry, 30*(4), 543-571.

Culicover, P. W., Jackendoff, R. S., & Jackendoff, R. (2005). *Simpler syntax*. Oxford University Press on Demand.

Dąbrowska, E. (2018). Experience, aptitude and individual differences in native language ultimate attainment. *Cognition, 178*, 222-235.

Dale, A. M. (1999). Optimal experimental design for event-related fMRI. *Human Brain Mapping, 8*(2-3), 109-114.

Dale, P.S., Dionne, G., Eley, T.C., Plomin, R. (2000). Lexical and grammatical development: A behavioural genetic perspective. Journal of Child Language, 27, 619-642.

Dale, P. S., Harlaar, N., Haworth, C. M. A., & Plomin, R. (2010). Two by Two: A Twin Study of Second-Language Acquisition. *Psychological Science, 21*(5), 635–640.

Damasio, H., Tranel, D., Grabowski, T., Adolphs, R., & Damasio, A. (2004). Neural systems behind word and concept retrieval. *Cognition, 92*(1-2), 179-229.

Dapretto, M., & Bookheimer, S. Y. (1999). Form and content: Dissociating syntax and semantics in sentence comprehension. *Neuron, 24*(2), 427-432.

Davis, M. H., Coleman, M. R., Absalom, A. R., Rodd, J. M., Johnsrude, I. S., Matta, B. F., Owen, A. M., & Menon, D. K. (2007). Dissociating speech perception and comprehension at reduced levels of awareness. *Proceedings of the National Academy of Sciences of the United States of America, 104*(41), 16032-16037.

Devauchelle, A. D., Oppenheim, C., Rizzi, L., Dehaene, S., & Pallier, C. (2009). Sentence syntax and content in the human temporal lobe: an fMRI adaptation study in auditory and visual modalities. *Journal of Cognitive Neuroscience, 21*(5), 1000-1012.

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

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *ArXiv:1810.04805 [Cs]*. http://arxiv.org/abs/1810.04805

Dick, F., Bates, E., Wulfeck, B., Utman, J. A., Dronkers, N., & Gernsbacher, M. A. (2001). Language deficits, localization, and grammar: Evidence for a distributive model of language breakdown in aphasic patients and neurologically intact individuals. *Psychological Review, 108*(4), 759-788. doi:10.1037//0033-295x.108.4.759

Dikker, S., Rabagliati, H., & Pylkkänen, L. (2009). Sensitivity to syntax in visual cortex. *Cognition*, *110*(3), 293-321.

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

Dixon, J. A., & Marchman, V. A. (2007). Grammar and the lexicon: Developmental ordering in language acquisition. *Child Development*, *78*(1), 190-212.

Drane, D. L., Ojemann, G. A., Aylward, E., Ojemann, J. G., Johnson, L. C., Silbergeld, D. L., ... & Tranel, D. (2008). Category-specific naming and recognition deficits in temporal lobe epilepsy surgical patients. *Neuropsychologia*, *46*(5), 1242-1255.

Duffau, H., Capelle, L., Denvil, D., Gatignol, P., Sichez, N., Lopes, M., ... & Van Effenterre, R. (2003). The role of dominant premotor cortex in language: a study using intraoperative functional mapping in awake patients. *Neuroimage*, *20*(4), 1903-1914.

Duffau, H., Moritz-Gasser, S., & Mandonnet, E. (2014). A re-examination of neural basis of language processing: Proposal of a dynamic hodotopical model from data provided by brain stimulation mapping during picture naming. *Brain and Language*, *131*, 1-10. doi:10.1016/j.bandl.2013.05.011

Duncan, J. (2010). The multiple-demand (MD) system of the primate brain: Mental programs for intelligent behaviour. *Trends in Cognitive Sciences*, *14*(4), 172-179.

Duncan, J. (2013). The Structure of Cognition: Attentional Episodes in Mind and Brain. *Neuron*, *80*(1), 35-50.

Eickhoff, S. B., Heim, S., Zilles, K., & Amunts, K. (2009). A systems perspective on the effective connectivity of overt speech production. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, *367*(1896), 2399–2421.

Embick, D., Marantz, A., Miyashita, Y., O'Neil, W., & Sakai, K. L. (2000). A syntactic specialization for Broca's area. *Proceedings of the National Academy of Sciences*, *97*(11), 6150-6154.

Engel, A. K., Moll, C. K., Fried, I., & Ojemann, G. A. (2005). Invasive recordings from the human brain: clinical insights and beyond. *Nature Reviews Neuroscience*, *6*(1), 35-47.

Erteschik-Shir, N. (1973). *On the nature of island constraints* (Doctoral dissertation, Massachusetts Institute of Technology).

Evert, S. (2008). Corpora and collocations. *Corpus linguistics. An international handbook*, *2*, 1212-1248.

Fadiga, L., Craighero, L., & D'Ausilio, A. (2009). Broca's Area in Language, Action, and Music. *Annals of the New York Academy of Sciences*, *1169*(1), 448–458.

Fedor, A., Varga, M., & Szathmáry, E. (2012). Semantics boosts syntax in artificial grammar learning tasks with recursion. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(3), 776.

Fedorenko, E. (2014). The role of domain-general cognitive control in language comprehension. *Frontiers in Psychology*, *5*, 335.

Fedorenko, E., Behr, M. K., & Kanwisher, N. (2011). Functional specificity for high-level linguistic processing in the human brain. *Proceedings of the National Academy of Sciences*, *108*(39), 16428–16433.

Fedorenko & Blank (in press). Broca's area is not a natural kind. *Trends in Cognitive Sciences.*

Fedorenko, E., Hsieh, P. J., Nieto-Castanon, A., Whitfield-Gabrieli, S., & Kanwisher, N. (2010). New method for fMRI investigations of language: Defining ROIs functionally in individual subjects. *Journal of Neurophysiology*, *104*(2), 1177-1194.

Fedorenko, E., & Kanwisher, N. (2009). Neuroimaging of language: why hasn't a clearer picture emerged? *Language and linguistics compass*, *3*(4), 839-865.

Fedorenko, E., Nieto-Castanon, A., & Kanwisher, N. (2012a). Lexical and syntactic representations in the brain: An fMRI investigation with multi-voxel pattern analyses. *Neuropsychologia*, *50*(4), 499-513.

Fedorenko, E., Nieto-Castañon, A. & Kanwisher, N. (2012b). Syntactic processing in the human brain: What we know, what we don't know, and a suggestion for how to proceed. Brain and Language, 120(2), 187-207.

Fedorenko, E., & Thompson-Schill, S. L. (2014). Reworking the language network. *Trends in cognitive sciences*, *18*(3), 120-126.

Fedorenko, E., & Varley, R. (2016). Language and thought are not the same thing: Evidence from neuroimaging and neurological patients. *Annals of the New York Academy of Sciences*, *1369*(1), 132-153.

Fedorenko, E., Williams, Z. M., & Ferreira, V. S. (2018). Remaining Puzzles about Morpheme Production in the Posterior Temporal Lobe. Neuroscience, 392, 160-163.

Ferreira, F., Bailey, K. G. D., & Ferraro, V. (2002). Good-enough representations in language comprehension. *Current Directions in Psychological Science, 11*(1), 11-15.

Ferreira, V. S., & Bock, K. (2006). The functions of structural priming. *Language and Cognitive Processes, 21*(7-8), 1011-1029.

Ferstl, E. C., & von Cramon, D. Y. (2001). The role of coherence and cohesion in text comprehension: an event-related fMRI study. *Cognitive brain research, 11*(3), 325-340.

Fiebach, C. J., Schlesewsky, M., Lohmann, G., Von Cramon, D. Y., & Friederici, A. D. (2005). Revisiting the role of Broca's area in sentence processing: syntactic integration versus syntactic working memory. *Human brain mapping, 24*(2), 79-91.

Fischl, B., Rajendran, N., Busa, E., Augustinack, J., Hinds, O., Yeo, B. T., . . . Zilles, K. (2008). Cortical folding patterns and predicting cytoarchitecture. *Cereb Cortex, 18*(8), 1973-1980.

Fitch, W. T., & Martins, M. D. (2014). Hierarchical processing in music, language, and action: Lashley revisited. *Annals of the New York Academy of Sciences, 1316*, 87–104.

Fodor, J. A. (1983). *The Modularity of Mind*. MIT Press.

Fodor, J. A., Bever, T. G., & Garrett, M. F. (1974). *The psychology of language: An introduction to psycholinguistics and generative grammar*. McGraw-Hill.

Friederici, A. D., Meyer, M., & von Cramon, D. Y. (2000). Auditory Language Comprehension: An Event-Related fMRI Study on the Processing of Syntactic and Lexical Information. *Brain and Language, 74*(2), 289–300.

Friederici, A. D., Bahlmann, J., Heim, S., Schubotz, R. I., & Anwander, A. (2006). The brain differentiates human and non-human grammars: functional localization and structural connectivity. *Proceedings of the National Academy of Sciences, 103*(7), 2458-2463.

Friederici, A. D. (2011). The brain basis of language processing: from structure to function. *Physiological reviews, 91*(4), 1357-1392.

Friederici, A. D. (2012). The cortical language circuit: from auditory perception to sentence comprehension. *Trends in Cognitive Sciences, 16*(5), 262-268.

Friederici, A. D. (2018). The neural basis for human syntax: Broca's area and beyond. *Current opinion in behavioral sciences, 21*, 88-92.

Friederici, A. D., Chomsky, N., Berwick, R. C., Moro, A., & Bolhuis, J. J. (2017). Language, mind and brain. *Nature Human Behaviour, 1*(10), 713.

Friederici, A. D., Kotz, S. A., Scott, S. K., & Obleser, J. (2010). Disentangling syntax and intelligibility in auditory language comprehension. *Human Brain Mapping, 31*(3), 448-457.

Friederici, A. D., Meyer, M., & von Cramon, D. Y. (2000). Auditory language comprehension: An event-related fMRI study on the processing of syntactic and lexical information. *Brain and Language, 74*(2), 289-300.

Friston, K. J., Rotshtein, P., Geng, J. J., Sterzer, P., & Henson, R. N. (2006). A critique of functional localisers. *Neuroimage, 30*(4), 1077-1087.

Frost, M. A., & Goebel, R. (2012). Measuring structural-functional correspondence: Spatial variability of specialised brain regions after macro-anatomical alignment. *Neuroimage, 59*(2), 1369-1381.

Futrell, R., Wilcox, E., Morita, T., & Levy, R. (2018). RNNs as psycholinguistic subjects: Syntactic state and grammatical dependency. *arXiv preprint arXiv:1809.01329*.

Futrell, R., Wilcox, E., Morita, T., Qian, P., Ballesteros, M., & Levy, R. (2019). Neural language models as psycholinguistic subjects: Representations of syntactic state. *arXiv preprint arXiv:1903.03260*.

Garnsey, S. M., Pearlmutter, N. J., Myers, E., & Lotocky, M. A. (1997). The contributions of verb bias and plausibility to the comprehension of temporarily ambiguous sentences. *Journal of Memory and Language, 37*(1), 58-93.

Gibson, E., Futrell, R., Piandadosi, S. T., Dautriche, I., Mahowald, K., Bergen, L., & Levy, R. (2019). How efficiency shapes human language. *Trends in Cognitive Sciences*.

Gibson, E., Bergen, L., & Piantadosi, S. T. (2013). Rational integration of noisy evidence and prior semantic expectations in sentence interpretation. *Proceedings of the National Academy of Sciences of the United States of America, 110*(20), 8051-8056.

Giraud, A. L., & Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. *Nature neuroscience, 15*(4), 511.

Glaser, Y. G., Martin, R. C., Van Dyke, J. A., Hamilton, A. C., & Tan, Y. (2013). Neural basis of semantic and syntactic interference in sentence comprehension. *Brain and Language*, *126*(3), 314–326.
Glasser, M. F., & Rilling, J. K. (2008). DTI Tractography of the Human Brain's Language Pathways. *Cerebral Cortex*, *18*(11), 2471–2482.
Goldberg, A. (2002). Construction Grammar. In *Encyclopedia of Cognitive Science*: Macmillan Reference Limited Nature Publishing Group.
Goldberg, A. E. (1995). *Constructions: A Construction Grammar Approach to Argument Structure*: University of Chicago Press.
Goldberg, A. E. (2006). *Constructions at work: The nature of generalization in language*. Oxford University Press on Demand.
Goldberg, A. E. (2013). 10 Backgrounded constituents cannot be "extracted". *Experimental syntax and island effects*, 221.
Goldinger, S. D. (1996). Words and voices: episodic traces in spoken word identification and recognition memory. *Journal of experimental psychology: Learning, memory, and cognition*, *22*(5), 1166.
Goodglass, H. (1993). *Understanding aphasia*. Academic Press.
Gorno-Tempini, M. L., Hillis, A. E., Weintraub, S., Kertesz, A., Mendez, M., Cappa, S. F., Ogar, J. M., Rohrer, J. D., Black, S., Boeve, B. F., Manes, F., Dronkers, N. F., Vandenberghe, R., Rascovsky, K., Patterson, K., Miller, B. L., Knopman, D. S., Hodges, J. R., Mesulam, M. M., & Grossman, M. (2011). Classification of primary progressive aphasia and its variants. *Neurology*, *76*(11), 1006.
Grabowski, T. J., Damasio, H., Tranel, D., Ponto, L. L. B., Hichwa, R. D., & Damasio, A. R. (2001). A role for left temporal pole in the retrieval of words for unique entities. *Human brain mapping*, *13*(4), 199-212.
Grodzinsky, Y., & Santi, A. (2008). The battle for Broca's region. *Trends in cognitive sciences*, *12*(12), 474-480.
Grossman, M., Lee, C., Morris, J., Stern, M.B. & Hurtig, H.I. (2002). Assessing resource demands during sentence processing in Parkinson's disease. *Brain and Language*, *80*, 603-616.
Gulordava, K., Bojanowski, P., Grave, E., Linzen, T., & Baroni, M. (2018). Colorless green recurrent networks dream hierarchically. *arXiv preprint arXiv:1803.11138*.
Hagoort, P. (2005). On Broca, brain, and binding: A new framework. *Trends in Cognitive Sciences*, *9*(9), 416–423.
Hagoort, P. (2013). MUC (Memory, Unification, Control) and beyond. *Frontiers in Psychology*, *4*.
Hagoort, P. (2019). The neurobiology of language beyond single-word processing. *Science*, *366*(6461), 55-58.
Hagoort, P., Brown, C., & Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Language and Cognitive Processes*, *8*(4), 439-483.
Hagoort, P., & Indefrey, P. (2014). The Neurobiology of Language Beyond Single Words. *Annual Review of Neuroscience*, *37*, 347.
Hahn, M., Jurafsky, D., & Futrell, R. (2020). Universals of word order reflect optimization of grammars for efficient communication. *Proceedings of the National Academy of Sciences*.
Hamamé, C. M., Szwed, M., Sharman, M., Vidal, J. R., Perrone-Bertolotti, M., Kahane, P., ... & Lachaux, J. P. (2013). Dejerine's reading area revisited with intracranial EEG: Selective responses to letter strings. *Neurology*, *80*(6), 602-603.
Hare, M. L., & Goldberg, A. E. (1999). Structural priming: Purely syntactic? *Proceedings of the Twenty First Annual Conference of the Cognitive Science Society*, 208-211.
Hasson, U., Chen, J., & Honey, C. J. (2015). Hierarchical process memory: memory as an integral component of information processing. *Trends in Cognitive Sciences*, *19*(6), 304-313.
Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science*, *298*(5598), 1569-1579.
Heim, S., Eickhoff, S. B., & Amunts, K. (2008). Specialisation in Broca's region for semantic, phonological, and syntactic fluency? *Neuroimage*, *40*(3), 1362-1368.
Herrmann, B., Obleser, J., Kalberlah, C., Haynes, J. D., & Friederici, A. D. (2012). Dissociable neural imprints of perception and grammar in auditory functional imaging. *Human Brain Mapping*, *33*(3), 584-595.
Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature reviews neuroscience*, *8*(5), 393-402.

Hoff, E., Quinn, J. M., & Giguere, D. (2018). What explains the correlation between growth in vocabulary and grammar? New evidence from latent change score analyses of simultaneous bilingual development. *Developmental science*, *21*(2), e12536.
Holmes, A., & Friston, K. (1998). Generalisability, random effects and population inference. *NeuroImage, 7*(4), S754.
Hong Y, Yoo Y, Wager T, Woo CW. (2019). False-positive neuroimaging: undisclosed flexibility in testing spatial hypotheses allows presenting anything as a replicated finding. bioRxiv :514-521.
Hudson, R. (2007). *Language Networks: The New Word Grammar*. Oxford University Press UK.
Humphries, C., Willard, K., Buchsbaum, B., & Hickok, G. (2001). Role of anterior temporal cortex in auditory sentence comprehension: An fMRI study. *NeuroReport*, *12*(8), 1749–1752.
Hurford, J. R. (1998). *Approaches to Evolution Language: Social and Cognitive Bases*. Cambridge University Press.
Hurford, J. R. (2007). *The Origins of Meaning*. Oxford University Press.
Indefrey, P., & Levelt, W. J. (2004). The spatial and temporal signatures of word production components. *Cognition*, *92*(1-2), 101-144.
Ioannidis JPA, Munafo MR, Fusar-Poli P, Nosek BA, David SP. (2014). Publication and other reporting biases in cognitive sciences: detection, prevalence, and prevention. *Trends in Cognitive Sciences, 18*(5), 235-241.
Jackendoff, R. (2002a). English particle constructions, the lexicon, and the autonomy of syntax. *Verb-particle explorations*, 67-94.
Jackendoff, R. (2002b). *Foundations of Language: Brain, Meaning, Grammar, Evolution*: Oxford University Press.
Jackendoff, R. (2007). A Parallel Architecture perspective on language processing. *Brain Research, 1146*, 2-22.
Jackendoff, R. (2011). What is the human language faculty? Two views. *Language*, *87*(3), 586–624.
Jackendoff, R., & Audring, J. (2020). *The Texture of the Lexicon: Relational Morphology and the Parallel Architecture*. Oxford University Press.
Jacoby, N., & Fedorenko, E. (2018). Discourse-level comprehension engages medial frontal Theory of Mind brain regions even for expository texts. *Language, Cognition and Neuroscience*, *0*(0), 1–17.
Jaeger, T. F. (2010). Redundancy and reduction: Speakers manage syntactic information density. *Cognitive psychology*, *61*(1), 23-62.
Joanette, Y., Goulet, P., & Hannequin, D. (1990). *Right Hemisphere and Verbal Communication*. Springer-Verlag.
Kaan, E., & Swaab, T. Y. (2002). The brain circuitry of syntactic comprehension. *Trends in cognitive sciences*, *6*(8), 350-356.
Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The Fusiform Face Area: A Module in Human Extrastriate Cortex Specialized for Face Perception. *Journal of Neuroscience*, *17*(11), 4302–4311.
Keller, T. A., Carpenter, P. A., & Just, M. A. (2001). The neural bases of sentence comprehension: a fMRI examination of syntactic and lexical processing. *Cerebral Cortex*, *11*(3), 223-237.
Kirby, S., Tamariz, M., Cornish, H., & Smith, K. (2015). Compression and communication in the cultural evolution of linguistic structure. *Cognition*, *141*, 87-102.
Kline, M., Gallée, J., Balewski Z., & Fedorenko, E. (2018). Understanding jokes draws most heavily on the Theory of Mind brain network. PsyArXiv preprint 10.31234/osf.io/h2nyx
Koechlin, E., & Jubault, T. (2006). Broca's area and the hierarchical organization of human behavior. *Neuron*, *50*(6), 963-974.
Kolk, H., & Chwilla, D. (2007). Late positivities in unusual situations. *Brain and Language*, *100*(3), 257-261.
Krekelberg, B., Boynton, G. M., & van Wezel, R. J. A. (2006). Adaptation: from single cells to BOLD signals. *Trends in Neurosciences*, *29*(5), 250-256.
Kriegeskorte, N., Simmons, W. K., Bellgowan, P. S., & Baker, C. I. (2009). Circular analysis in systems neuroscience: the dangers of double dipping. *Nature Neuroscience*, *12*(5), 535-540.
Kuno, S. (1987). *Functional syntax: Anaphora, discourse and empathy*. University of Chicago Press.
Kuperberg, G. R., McGuire, P. K., Bullmore, E. T., Brammer, M. J., Rabe-Hesketh, S., Wright, I. C., Lythgoe, D. J., Williams, S. C. R., & David, A. S. (2000). Common and Distinct Neural Substrates for Pragmatic, Semantic, and Syntactic Processing of Spoken Sentences: An fMRI Study. *Journal of Cognitive Neuroscience*, *12*(2), 321–341.

Kuperberg, G. R., Holcomb, P. J., Sitnikova, T., Greve, D., Dale, A. M., & Caplan, D. (2003). Distinct patterns of neural modulation during the processing of conceptual and syntactic anomalies. *Journal of Cognitive Neuroscience*, *15*(2), 272-293.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology, Vol 62, 62*, 621-647.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*(4427), 203-205.

Kvarven, A., Strømland, E., & Johannesson, M. (2019). Comparing meta-analyses and preregistered multiple-laboratory replication projects. *Nature Human Behaviour*, 1-12.

Lakoff, G. (1970). *Irregularity in syntax*. New York: Holt, Rinehart, and Winston.

Lambon-Ralph, M. A., McClelland, J. L., Patterson, K., Galton, C. J., & Hodges, J. R. (2001). No Right to Speak? The Relationship between Object Naming and Semantic Impairment: Neuropsychological Evidence and a Computational Model. *Journal of Cognitive Neuroscience*, *13*(3), 341–356.

Lane, C., Kanjlia, S., Omaki, A., & Bedny, M. (2015). "Visual" cortex of congenitally blind adults responds to syntactic movement. *Journal of Neuroscience*, *35*(37), 12859-12868.

Langacker, R. W. (1986). An introduction to cognitive grammar. *Cognitive science*, *10*(1), 1-40.

Langacker, R. W. (1987). *Foundations of cognitive grammar: Theoretical prerequisites* (Vol. 1). Stanford University Press.

Lee, D. K., Fedorenko, E., Simon, M. V., Curry, W. T., Nahed, B. V., Cahill, D. P., & Williams, Z. M. (2018). Neural encoding and production of functional morphemes in the posterior temporal lobe. *Nature communications*, *9*(1), 1-12.

Lerner, Y., Honey, C. J., Silbert, L. J., & Hasson, U. (2011). Topographic Mapping of a Hierarchy of Temporal Receptive Windows Using a Narrated Story. *Journal of Neuroscience*, *31*(8), 2906–2915.

Levin, B. (1993). *English verb classes and alternations: A preliminary investigation*. University of Chicago press.

Levin, B., & Rappaport-Hovav, M. (2005). *Argument Realization*. Cambridge, UK: Cambridge University Press.

Levy, R. (2008). A noisy-channel model of human sentence comprehension under uncertain input. In *Proceedings of the 2008 conference on empirical methods in natural language processing* (pp. 234-243).

Lindenberg, R., Fangerau, H., & Seitz, R. J. (2007). "Broca's area" as a collective term? *Brain and Language*, *102*(1), 22–29.

Linzen, T., Dupoux, E., & Goldberg, Y. (2016). Assessing the ability of LSTMs to learn syntax-sensitive dependencies. *Transactions of the Association for Computational Linguistics*, *4*, 521-535.

Longworth, C. E., Keenan, S. E., Barker, R. A., Marslen-Wilson, W. D., & Tyler, L. K. (2005). The basal ganglia and rule-governed language use: evidence from vascular and degenerative conditions. *Brain*, *128*(3), 584-596.

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, *101*, 676-703.

Mahowald, K., & Fedorenko, E. (2016). Reliable individual-level neural markers of high-level language processing: A necessary precursor for relating neural variability to behavioral and genetic variability. *Neuroimage*, *139*, 74-93.

Mahowald, K., James, A., Futrell, R., & Gibson, E. (2016). A meta-analysis of syntactic priming in language production. *Journal of Memory and Language*, *91*, 5-27.

Mariën, P., Ackermann, H., Adamaszek, M., Barwood, C. H., Beaton, A., Desmond, J., ... & Leggio, M. (2014). Consensus paper: language and the cerebellum: an ongoing enigma. *The Cerebellum*, *13*(3), 386-410.

Martin, A. E., & Doumas, L. A. (2019). Predicate learning in neural systems: using oscillations to discover latent structure. *Current Opinion in Behavioral Sciences*, *29*, 77-83.

Mason, R. A., Just, M. A., Keller, T. A., & Carpenter, P. A. (2003). Ambiguity in the brain: What brain imaging reveals about the processing of syntactically ambiguous sentences. *Journal of Experimental Psychology-Learning Memory and Cognition*, *29*(6), 1319–1338.

Mason, R. A., & Just, M. A. (2007). Lexical ambiguity in sentence comprehension. *Brain Research*, *1146*, 115–127.

Matchin, W., Brodbeck, C., Hammerly, C., & Lau, E. (2019). The temporal dynamics of structure and content in sentence comprehension: Evidence from fMRI-constrained MEG. *Human brain mapping*, *40*(2), 663-678.

Matchin, W., & Hickok, G. (2019). The cortical organization of syntax. *Cerebral Cortex*.

Mather, M., Cacioppo, J. T., & Kanwisher, N. (2013). How fMRI Can Inform Cognitive Theories. *Perspectives on Psychological Science*, *8*(1), 108–113.

Mazoyer, B. M., Tzourio, N., Frak, V., Syrota, A., Murayama, N., Levrier, O., Salamon, G., Dehaene, S., Cohen, L., & Mehler, J. (1993). The Cortical Representation of Speech. *Journal of Cognitive Neuroscience*, *5*(4), 467-479.

McCandliss, B. D., Cohen, L., & Dehaene, S. (2003). The visual word form area: Expertise for reading in the fusiform gyrus. *Trends in Cognitive Sciences*, *7*(7), 293–299.

McClelland, J. L., Rumelhart, D. E., & PDP Research Group. (1986). Parallel distributed processing. *Explorations in the Microstructure of Cognition*, *2*, 216-271.

Menenti, L., Gierhan, S. M. E., Segaert, K., & Hagoort, P. (2011). Shared Language: Overlap and Segregation of the Neuronal Infrastructure for Speaking and Listening Revealed by Functional MRI. *Psychological Science*, *22*(9), 1173–1182.

Menenti, L., Petersson, K. M., & Hagoort, P. (2012). From reference to sense: how the brain encodes meaning for speaking. *Frontiers in Psychology*, *3*.

Mesulam, M. M., Wieneke, C., Hurley, R., Rademaker, A., Thompson, C. K., Weintraub, S., & Rogalski, E. J. (2013). Words and objects at the tip of the left temporal lobe in primary progressive aphasia. *Brain*, *136*(2), 601-618.

Mesulam, M.-M. *et al.* (2014). Primary progressive aphasia and the evolving neurology of the language network. *Nat. Rev. Neurol.* 10, 554-569.

Mesulam, M.-M., Thompson, C.K., Weintraub, S., Rogalski, E.J. (2015). The Wernicke conundrum and the anatomy of language comprehension in primary progressive aphasia. *Brain*, 138(8), 2423-2437.

Meyer, L. (2018). The neural oscillations of speech processing and language comprehension: state of the art and emerging mechanisms. *European Journal of Neuroscience*, *48*(7), 2609-2621.

Mikolov, T., Karafiát, M., Burget, L., Černocký, J., & Khudanpur, S. (2010). Recurrent neural network based language model. In *Eleventh Annual Conference of the International Speech Communication Association*.

Miyake, A., Carpenter, P. A., & Just, M. A. (1994). A capacity approach to syntactic comprehension disorders: Making normal adults perform like aphasic patients. Cognitive Neuropsychology, 11, 671-717.

Mollica, F., Shain, C., Affourtit, J., Kean, H., Siegelman, M., and Fedorenko, E. (in prep.) Another look at the constituent structure of sentences in the human brain.

Mollica, F., Siegelman, M., Diachek, E., Piantadosi, S. T., Mineroff, Z., Futrell, R., ... & Fedorenko, E. (in press). Composition is the core driver of the language-selective network. *Neurobiology of Language*, 1-67.

Mollica, F., & Piantadosi, S. T. (2019). Humans store about 1.5 megabytes of information during language acquisition. *Royal Society open science*, *6*(3), 181393.

Monti, M. M., Parsons, L. M., & Osherson, D. N. (2012). Thought Beyond Language: Neural Dissociation of Algebra and Natural Language. *Psychological Science*, *23*(8), 914–922.

Morgan, E., & Levy, R. (2016). Abstract knowledge versus direct experience in processing of binomial expressions. *Cognition*, *157*, 384-402.

Mukamel, R., & Fried, I. (2012). Human intracranial recordings and cognitive neuroscience. *Annual review of psychology*, *63*, 511-537.

Nelson, M. J., El Karoui, I., Giber, K., Yang, X. F., Cohen, L., Koopman, H., . . . Dehaene, S. (2017). Neurophysiological dynamics of phrase-structure building during sentence processing. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(18), E3669-E3678.

Newman, A. J., Pancheva, R., Ozawa, K., Neville, H. J., & Ullman, M. T. (2001). An event-related fMRI study of syntactic and semantic violations. *Journal of psycholinguistic research*, *30*(3), 339-364.

Newmeyer, F. J. (2003). Grammar is grammar and usage is usage. *Language*, *79*(4), 682-707.

Nieto-Castañón, A., & Fedorenko, E. (2012). Subject-specific functional localizers increase sensitivity and functional resolution of multi-subject analyses. *NeuroImage*, *63*(3), 1646-1669.

Ni, W., Constable, R. T., Mencl, W. E., Pugh, K. R., Fulbright, R. K., Shaywitz, S. E., ... & Shankweiler, D. (2000). An event-related neuroimaging study distinguishing form and content in sentence processing. *Journal of Cognitive Neuroscience*, *12*(1), 120-133.

Nieuwenhuis, S., Forstmann, B. U., & Wagenmakers, E. J. (2011). Erroneous analyses of interactions in neuroscience: a problem of significance. *Nature Neuroscience, 14*(9), 1105-1107.

Noppeney, U., & Price, C. J. (2004). An fMRI study of syntactic adaptation. *Journal of Cognitive Neuroscience, 16*(4), 702-713.

Norcliffe, E., Harris, A. C., & Jaeger, T. F. (2015). Cross-linguistic psycholinguistics and its critical role in theory development: Early beginnings and recent advances.

Norman-Haignere, S., Kanwisher, N. G., & McDermott, J. H. (2015). Distinct Cortical Pathways for Music and Speech Revealed by Hypothesis-Free Voxel Decomposition. *Neuron*, *88*(6), 1281–1296.

Norris, D. G., & Polimeni, J. R. (2019). Laminar (f)MRI: A short history and future prospects. *Neuroimage*, *197*, 643.

O'Donnell, T. J. (2015). *Productivity and Reuse in Language: A Theory of Linguistic Computation and Storage*: MIT Press.

Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language, 31*(6), 785-806.

Overath, T., McDermott, J. H., Zarate, J. M., & Poeppel, D. (2015). The cortical analysis of speech-specific temporal structure revealed by responses to sound quilts. *Nature Neuroscience*, *18*(6), 903–911.

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

Papoutsi, M., Stamatakis, E. A., Griffiths, J., Marslen-Wilson, W. D., & Tyler, L. K. (2011). Is left fronto-temporal connectivity essential for syntax? Effective connectivity, tractography and performance in left-hemisphere damaged patients. *NeuroImage, 58*(2), 656–664.

Patterson, K., Nestor, P. J., & Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Reviews. Neuroscience*, *8*(12), 976–987.

Petersson, K. M., Folia, V., & Hagoort, P. (2012). What artificial grammar learning reveals about the neurobiology of syntax. *Brain and language*, *120*(2), 83-95.

Pickering, M. J., & Ferreira, V. S. (2008). Structural priming: A critical review. *Psychological Bulletin, 134*(3), 427-459.

Pinker, S. (1989). *Learnability and Cognition: The Acquisition of Argument Structure*. Cambridge, MA: MIT Press.

Pinker, S. (1991). Rules of language. *Science*, *253*(5019), 530–535.

Pinker, S. (1995). *The language instinct*: William Morrow and Company.

Pinker, S. (1999). *Words and rules: The ingredients of language*. London: Weidenfeld & Nicolson.

Pinker, S., & Prince, A. (1988). On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition*, *28*(1), 73–193.

Poldrack, R. A., Wagner, A. D., Prull, M. W., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. E. (1999). Functional Specialization for Semantic and Phonological Processing in the Left Inferior Prefrontal Cortex. *NeuroImage*, *10*(1), 15–35.

Poldrack, R. A. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Sciences, 10*(2), 59-63.

Poldrack, R. A. (2011). Inferring Mental States from Neuroimaging Data: From Reverse Inference to Large-Scale Decoding. *Neuron, 72*(5), 692-697.

Poldrack, R. A., Laumann, T. O., Koyejo, O., Gregory, B., Hover, A., Chen, M.-Y., Gorgolewski, K. J., Luci, J., Joo, S. J., Boyd, R. L., Hunicke-Smith, S., Simpson, Z. B., Caven, T., Sochat, V., Shine, J. M., Gordon, E., Snyder, A. Z., Adeyemo, B., Petersen, S. E., … Mumford, J. A. (2015). Long-term neural and physiological phenotyping of a single human. *Nature Communications*, *6*(1), 1–15.

Poldrack RA, Baker CI, Durnez J, Gorgolewski KJ, Matthews PM, Munafò MR, Yarkoni T. (2017). Scanning the horizon: towards trans- parent and reproducible neuroimaging research. Nat Rev Neurosci 18(2), 115.

Prasad, G., van Schijndel, M., & Linzen, T. (2019). Using Priming to Uncover the Organization of Syntactic Representations in Neural Language Models. *arXiv preprint arXiv:1909.10579*.

Pylkkänen, L. (2019). The neural basis of combinatory syntax and semantics. *Science*, *366*(6461), 62–66.
Rauschecker, J. P., & Scott, S. K. (2009). Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nature neuroscience*, *12*(6), 718-724.
Reali, F., & Christiansen, M. H. (2007). Processing of relative clauses is made easier by frequency of occurrence. *Journal of memory and language*, *57*(1), 1-23.
Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of verbal learning and verbal behavior*, *6*(6), 855-863.
Rodd, J. M., Davis, M. H., & Johnsrude, I. S. (2005). The neural mechanisms of speech comprehension: FMRI studies of semantic ambiguity. *Cerebral Cortex (New York, N.Y.: 1991)*, *15*(8), 1261–1269.
Rodd, J. M., Longe, O. A., Randall, B., & Tyler, L. K. (2010). The functional organisation of the fronto-temporal language system: Evidence from syntactic and semantic ambiguity. *Neuropsychologia*, *48*(5), 1324–1335.
Rodd, J. M., Berriman, R., Landau, M., Lee, T., Ho, C., Gaskell, M. G., & Davis, M. H. (2012). Learning new meanings for old words: Effects of semantic relatedness. *Memory & Cognition*, *40*(7), 1095-1108.
Rodd, J. M., Vitello, S., Woollams, A. M., & Adank, P. (2015). Localising semantic and syntactic processing in spoken and written language comprehension: An Activation Likelihood Estimation meta-analysis. *Brain and Language*, *141*, 89-102.
Roder, B., Stock, O., Neville, H., Bien, S., & Rosler, F. (2002). Brain activation modulated by the comprehension of normal and pseudo-word sentences of different processing demands: A functional magnetic resonance Imaging study. *Neuroimage*, *15*(4), 1003-1014.
Rogalsky, C., & Hickok, G. (2009). Selective attention to semantic and syntactic features modulates sentence processing networks in anterior temporal cortex. *Cerebral Cortex*, *19*(4), 786-796.
Rogers, T. T., Lambon Ralph, M. A., Garrard, P., Bozeat, S., McClelland, J. L., Hodges, J. R., & Patterson, K. (2004). Structure and deterioration of semantic memory: a neuropsychological and computational investigation. *Psychological review*, *111*(1), 205.
Rogers, T. T., Hocking, J., Noppeney, U. T. A., Mechelli, A., Gorno-Tempini, M. L., Patterson, K., & Price, C. J. (2006). Anterior temporal cortex and semantic memory: reconciling findings from neuropsychology and functional imaging. *Cognitive, Affective, & Behavioral Neuroscience*, *6*(3), 201-213.
Roland, D., Dick, F., & Elman, J. L. (2007). Frequency of basic English grammatical structures: A corpus analysis. *Journal of Memory and Language*, *57*(3), 348–379.
Rosenfelt, L., Barkley, C., Belvin, K., Lee, C., Federmeier, K., Kluender, R., and Kutas, M. (2009). No ERP evidence for first pass structure building: Pure word category violations do not elicit early negativity. CUNY Presentation (Davis, CA).
Rosenfelt, L., Kluender, R. & Kutas, M. (2011). Severing the tie between the eLAN and automatic syntactic processing. AMLaP presentation (Paris, France).
Ross, J. R. (1967). Constraints on variables in syntax.
Ryskin, R., Levy, R. P., & Fedorenko, E. (2020a). Do domain-general executive resources play a role in linguistic prediction? Re-evaluation of the evidence and a path forward. *Neuropsychologia*, *136*, 107258.
Sag, I. A., & Wasow, T. (2000). Syntactic Theory: A Formal Introduction. *Computational Linguistics*, *26*(2), 295–295.
Santi, A., & Grodzinsky, Y. (2010). fMRI adaptation dissociates syntactic complexity dimensions. *Neuroimage*, *51*(4), 1285-1293.
Sassenhagen, J., Schlesewsky, M., & Bornkessel-Schlesewsky, I. (2014). The P600-as-P3 hypothesis revisited: Single-trial analyses reveal that the late EEG positivity following linguistically deviant material is reaction time aligned. *Brain and language*, *137*, 29-39.
Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people: the role of the temporo-parietal junction in "theory of mind". *Neuroimage*, *19*(4), 1835-1842.
Saxe, R., Brett, M., & Kanwisher, N. (2006). Divide and conquer: A defense of functional localizers. *NeuroImage*, *30*(4), 1088-1096.
Scheepers, C., Raffray, C., & Myachykov, A. (2017). The lexical boost effect is not diagnostic of lexically-specific syntactic representations. *Memory and Language*, *95*, 102-115.
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.

Schotter, E. R., Tran, R., & Rayner, K. (2014). Don't Believe What You Read (Only Once): Comprehension Is Supported by Regressions During Reading. *Psychological Science*, *25*(6), 1218–1226.

Schmidt, S. (2009). Shall We Really Do It Again? The Powerful Concept of Replication Is Neglected in the Social Sciences. *Review of General Psychology*, *13*, 90–100.

Schütze, C. T., Sprouse, J., & Caponigro, I. (2015). Challenges for a theory of islands: A broader perspective on Ambridge, Pine, and Lieven. *Language*, *91*(2), e31-e39.

Scott, S. K., Blank, C. C., Rosen, S., & Wise, R. J. (2000). Identification of a pathway for intelligible speech in the left temporal lobe. *Brain: A Journal of Neurology*, *123 Pt 12*, 2400–2406.

Scott, T. L., Gallee, J., & Fedorenko, E. (2016). A new fun and robust version of an fMRI localizer for the frontotemporal language system. *Cognitive Neuroscience*, 1-10.

Segaert, K., Menenti, L., Weber, K., Petersson, K. M., & Hagoort, P. (2012). Shared syntax in language production and language comprehension - an fMRI Study. *Cerebral Cortex*, *22*(7), 1662-1670.

Shain, C., Blank, I. A., van Schijndel, M., Schuler, W., & Fedorenko, E. (2019). fMRI reveals language-specific predictive coding during naturalistic sentence comprehension. *Neuropsychologia*, 107307.

Siegelman, M., Blank, I. A., Mineroff, Z., & Fedorenko, E. (2019). An Attempt to Conceptually Replicate the Dissociation between Syntax and Semantics during Sentence Comprehension. *Neuroscience*, *413*, 219–229.

Silbert, L. J., Honey, C. J., Simony, E., Poeppel, D., & Hasson, U. (2014). Coupled neural systems underlie the production and comprehension of naturalistic narrative speech. *Proceedings of the National Academy of Sciences*, *111*(43), E4687-E4696.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant. *Psychological Science*, *22*(11), 1359–1366.

Snedeker, J., Geren, J., & Shafto, C. L. (2007). Starting over: International adoption as a natural experiment in language development. *Psychological Science*, *18*(1), 79–87.

Snedeker, J., Geren, J., & Shafto, C. L. (2012). Disentangling the effects of cognitive development and linguistic expertise: A longitudinal study of the acquisition of English in internationally-adopted children. *Cognitive Psychology*, *65*(1), 39–76.

Steinhauer, K., & Drury, J. E. (2012). On the early left-anterior negativity (ELAN) in syntax studies. *Brain and Language*, *120*(2), 135–162.

Stromswold, K., Caplan, D., Alpert, N., & Rauch, S. (1996). Localization of syntactic comprehension by positron emission tomography. *Brain and language*, *52*(3), 452-473.

Sutskever, I., Vinyals, O., & Le, Q. V. (2014). Sequence to Sequence Learning with Neural Networks. *ArXiv:1409.3215 [Cs]*. http://arxiv.org/abs/1409.3215

Tahmasebi, A. M., Davis, M. H., Wild, C. J., Rodd, J. M., Hakyemez, H., Abolmaesumi, P., & Johnsrude, I. S. (2012). Is the link between anatomical structure and function equally strong at all cognitive levels of processing? *Cerebral Cortex (New York, N.Y.: 1991)*, *22*(7), 1593–1603.

Tettamanti, M., & Weniger, D. (2006). Broca's area: a supramodal hierarchical processor? *Cortex*, *42*(4), 491-494.

Tie, Y., Rigolo, L., Norton, I. H., Huang, R. Y., Wu, W., Orringer, D., Mukundan, S., & Golby, A. J. (2014). Defining language networks from resting-state fMRI for surgical planning—A feasibility study. *Human Brain Mapping*, *35*(3), 1018–1030.

Tomasello, M. (2003). Constructing a Language: A Usage-Based Theory of Language Acquisition. Cambridge, MA: Harvard University Press.

Tranel, D. (2006). Impaired naming of unique landmarks is associated with left temporal polar damage. *Neuropsychology*, *20*(1), 1.

Tranel, D. (2009). The left temporal pole is important for retrieving words for unique concrete entities. *Aphasiology*, *23*(7-8), 867-884.

Traxler, M. J., Morris, R. K., & Seely, R. E. (2002). Processing subject and object relative clauses: Evidence from eye movements. *Journal of Memory and Language*, *47*(1), 69-90.

Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of memory and language*, *33*, 285-285.

Tsao, D. Y., Moeller, S., & Freiwald, W. A. (2008). Comparing face patch systems in macaques and humans. *Proceedings of the National Academy of Sciences*, *105*(49), 19514-19519.

Tyler, L. K., Marslen-Wilson, W. D., Randall, B., Wright, P., Devereux, B. J., Zhuang, J., . . . Stamatakis, E. A. (2011). Left inferior frontal cortex and syntax: function, structure and behaviour in patients with left hemisphere damage. *Brain*, *134*, 415-431. doi:10.1093/brain/awq369

Ullman, M. T. (2001). A neurocognitive perspective on language: the declarative/procedural model. *Nature Reviews Neuroscience*, *2*, 717-726.

Ullman, M. T. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition*, *92*(1-2), 231-270.

Ullman, M. T. (2016). The Declarative/Procedural Model : A Neurobiological Model of Language Learning, Knowledge, and Use. In *Neurobiology of Language*: Academic Press.

van de Meerendonk, N., Indefrey, P., Chwilla, D. J., & Kolk, H. H. J. (2011). Monitoring in language perception: Electrophysiological and hemodynamic responses to spelling violations. *NeuroImage*, *54*(3), 2350–2363.

van de Meerendonk, N., Kolk, H. H. J., Vissers, C., & Chwilla, D. J. (2010). Monitoring in Language Perception: Mild and Strong Conflicts Elicit Different ERP Patterns. *Journal of Cognitive Neuroscience*, *22*(1), 67-82.

van de Meerendonk, N., Kolk, H. J., & Chwilla, D. J. (2009). Monitoring in language perception *Language and Linguistics Compass*, *3*(5), 1211-1224.

Vandenberghe, R., Nobre, A. C., & Price, C. J. (2002). The response of left temporal cortex to sentences. *Journal of cognitive neuroscience*, *14*(4), 550-560.

Vázquez-Rodríguez, B., Suárez, L. E., Markello, R. D., Shafiei, G., Paquola, C., Hagmann, P., ... & Misic, B. (2019) Gradients of structure–function tethering across neocortex. *Proceedings of the National Academy of Sciences*, *116*(42), 21219-21227.

Vigneau, M., Beaucousin, V., Hervé, P. Y., Duffau, H., Crivello, F., Houdé, O., Mazoyer, B., & Tzourio-Mazoyer, N. (2006). Meta-analyzing left hemisphere language areas: Phonology, semantics, and sentence processing. *NeuroImage*, *30*(4), 1414–1432.

Vissers, C., Chwilla, D. J., & Kolk, H. H. J. (2006). Monitoring in language perception: The effect of misspellings of words in highly constrained sentences. *Brain Research*, *1106*, 150-163. doi:10.1016/j.brainres.2006.05.012

Vissers, C., Chwilla, D. J., & Kolk, H. H. J. (2007). The interplay of heuristics and parsing routines in sentence comprehension: Evidence from ERPs and reaction times. *Biological Psychology*, *75*(1), 8-18.

Wang, L., Uhrig, L., Jarraya, B., & Dehaene, S. (2015). Representation of numerical and sequential patterns in macaque and human brains. *Current Biology*, *25*(15), 1966-1974.

Wilcox, E., Levy, R., & Futrell, R. (2019s). Hierarchical representation in neural language models: Suppression and recovery of expectations. *arXiv preprint arXiv:1906.04068*

Wilcox, E., Levy, R. & Futrell, R. (2019b). What syntactic structures block dependencies in RNN language models? *arXiv preprint arXiv: 1905.10431*

Willems, R. M., der Haegen, L. V., Fisher, S. E., & Francks, C. (2014). On the other hand: Including left-handers in cognitive neuroscience and neurogenetics. *Nature Reviews Neuroscience*, *15*(3), 193-201.

Wilson, S. M., Galantucci, S., Tartaglia, M. C., Rising, K., Patterson, D. K., Henry, M. L., Ogar, J. M., DeLeon, J., Miller, B. L., & Gorno-Tempini, M. L. (2011). Syntactic Processing Depends on Dorsal Language Tracts. *Neuron*, *72*(2), 397–403.

Wilson, S. M., & Saygin, A. P. (2004). Grammaticality judgment in aphasia: Deficits are not specific to syntactic structures, aphasic syndromes, or lesion sites. *Journal of Cognitive Neuroscience*, *16*(2), 238–252.

Wise, R., Greene, J., Büchel, C., & Scott, S. (1999). Brain regions involved in articulation. *The Lancet*, *353*(9158), 1057–1061.

Wray, A. (2005). *Formulaic language and the lexicon*. Cambridge University Press.

Zempleni, M.-Z., Renken, R., Hoeks, J. C. J., Hoogduin, J. M., & Stowe, L. A. (2007). Semantic ambiguity processing in sentence context: Evidence from event-related fMRI. *NeuroImage*, *34*(3), 1270-1279.

Ziegler, J., Bencini, G., Goldberg, A., & Snedeker, J. (2019). How abstract is syntax? Evidence from structural priming. *Cognition*, *193*, 104045.

Ziegler, J. & Snedeker, J. (2018). How broad are thematic roles? Evidence from structural priming. *Cognition*, 179, 221-240.

Ziegler, J., Snedeker, J., & Wittenberg, E. (2018). Event structures drive semantic structural priming, not thematic roles: Evidence from idioms and light verbs. *Cognitive Science*.