Original Articles

# Grounding the neurobiology of language in first principles: The necessity of non-language-centric explanations for language comprehension

Uri Hasson[a,b,*,1], Giovanna Egidi[a], Marco Marelli[c,f], Roel M. Willems[d,e]

[a] Center for Mind/Brain Sciences, The University of Trento, Trento, Italy
[b] Center for Practical Wisdom, The University of Chicago, Chicago, IL, United States
[c] Department of Psychology, University of Milano-Bicocca, Milano, Italy
[d] Centre for Language Studies & Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, The Netherlands
[e] Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands
[f] NeuroMI - Milan Center for Neuroscience, Milano, Italy

## ARTICLE INFO

## ABSTRACT

Recent decades have ushered in tremendous progress in understanding the neural basis of language. Most of our current knowledge on language and the brain, however, is derived from lab-based experiments that are far removed from everyday language use, and that are inspired by questions originating in linguistic and psycholinguistic contexts. In this paper we argue that in order to make progress, the field needs to shift its focus to understanding the neurobiology of naturalistic language comprehension. We present here a new conceptual framework for understanding the neurobiological organization of language comprehension. This framework is non-language-centered in the computational/neurobiological constructs it identifies, and focuses strongly on context. Our core arguments address three general issues: (i) the difficulty in extending language-centric explanations to discourse; (ii) the necessity of taking context as a serious topic of study, modeling it formally and acknowledging the limitations on external validity when studying language comprehension outside context; and (iii) the tenuous status of the language network as an explanatory construct. We argue that adopting this framework means that neurobiological studies of language will be less focused on identifying correlations between brain activity patterns and mechanisms postulated by psycholinguistic theories. Instead, they will be less self-referential and increasingly more inclined towards integration of language with other cognitive systems, ultimately doing more justice to the neurobiological organization of language and how it supports language as it is used in everyday life.

## 1. Introduction

The last two decades have witnessed extensive methodological advances in the non-invasive study of brain activity. These advances allow researchers to address questions that have been at the core of the neurobiology of language since its inception, addressing the structural and functional basis of phonetic, semantic and syntactic processing. Neuropsychological analysis of brain damage was the dominant method for understanding neural function for over 150 years (see Levelt, 2012), but offered only a relatively gross picture of neural function in language processing, which could not capture the unfolding of neural events among transiently activated brain regions. In contrast, the current state of the art allows characterizing comprehension as a product of network-level interactions at different temporal and spatial scales, and offers promise towards explaining how the brain supports language comprehension in naturalistic, everyday language use.

Yet, the theoretical focus in research into the neurobiological organization of language has largely been maintained on those questions originating in linguistics and psycholinguistics, and theoretical progress in the neurobiology of language has maintained a remarkably stable and linear course of advancement. Experimental work has largely followed a research program wherein: (i) a linguistic function or operation

---

is defined, (ii) an experimental paradigm manipulates a variable that operationalizes this function, and (iii) regions or networks whose activity or connectivity varies systematically with the levels of the independent variable are interpreted as supporting, subserving, mediating, computing, implementing, or otherwise performing the cognitive process in question. This neurolinguistic approach has catalyzed and advanced research in the neurobiology of language: matching activity patterns to function assists diagnosis in cases of stroke or trauma, is core to studying brain-behavior correlations, and allows sophisticated meta-analyses to draw conclusions about brain regions associated with linguistic functions. Beyond their compatibility with the premises of the lesion-symptom mapping work that initiated the neurobiology of language, several other factors contribute to make neurolinguistic models the tool of choice for studying language in the brain. They set up constrained hypotheses, are for the most part conceptually precise and, by social convention, often limit themselves to examining a single brain region or tightly defined networks. All these factors contribute to clearly written scientific reports, communicating results that map onto the linguistic and psycholinguistic communities' common ground.

However, targeting the neurobiological basis of experimentally isolable processes has offered little by way of understanding how the brain supports language comprehension *as it is carried out in everyday naturalistic discourse*. This means comprehension of ideas presented across multiple sentences. The fact that neurobiological accounts of syntax and semantics drawn from tightly controlled studies may not scale to the discourse level is only one concern. More importantly, the existing paradigm simply does not intend to address questions that are crucial for understanding the neurobiology of Naturalistic Language Comprehension (NLC henceforth). Examples are the interaction between semantic processes and memory encoding or retrieval during comprehension, or the impact of context on online comprehension processes. Thus, there is a real question about whether data from simple experiments are relevant to understanding everyday language processing.

If extrapolation from basic paradigms is not a productive analytic method, how can we study and interpret which principles organize brain activity during NLC? Based on work by us and others, we argue here that addressing this question requires a different explanatory framework. This framework attributes a central role to neurobiological mechanisms that implement language but are not essentially linguistic; holds that context must be considered as fundamental for understanding the neurobiology of language comprehension rather than an additional consideration; and argues that neurobiological accounts of comprehension must divorce from the idea that it is largely dependent on activity within a central language network. In the next subsection we provide a synopsis of these main tenets of the framework, which we then present in detail in the subsequent sections of the manuscript.[2]

## 1.1. Theoretical tenets

*Limiting language-specific interpretations:* Brain activity observed during language comprehension is frequently interpreted in terms of core linguistic processes. Such computations include but are not limited to monotonic integration of information, establishing coherence, and prediction. Despite this, several studies suggest that these effects are parsimoniously explained by basic computations that are not limited to language comprehension, although often documented in areas associated with language comprehension. For this reason, linguistic-related constructs such as semantic or syntactic complexity should not constitute the default interpretive framework. Rather, from first principles,

such effects should be adopted after considering alternatives that can be formulated in terms of generic predictive and compositional processes not unique to language. Processes that co-occur during naturalistic comprehension, such as memory operations or emotional responses also fall within this category. Section 2 presents this argument.

*Broadening the notion of context and emphasizing its necessity:* There have been several demonstrations of the impact of context on language comprehension (for reviews, see Hagoort & van Berkum, 2007; van Berkum, 2008). Our argument for an essential role for context in neurobiological explanation derives from a synthesis of this work, as well as recent developments in computational modeling of language. First, during language comprehension, very diverse types of context appear to be integrated within the same time frame. These include prior textual context (co-text), the social context of the communication such as characteristics of the speaker, or personal context such as the beliefs of the comprehenders, or their mood. In addition, contextual integration appears to implicate a limited set of networks that are often involved in semantic processing. We present several generic, non-linguistic computational architectures that can support this broad sort of contextual integration. Our second argument for taking context as an organizing factor is based on studies that suggest that neurobiological conclusions drawn from studies where single-sentences are presented outside of context do not naturally extend to more naturalistic contexts. This is a problem of external/ecological validity. Section 3 presents the argument for opening up the notion of context.

*Letting go of the notion of a stable language network:* Concentrating on a neurobiological language network, defined anatomically or functionally, as a starting point for investigation results in an incomplete understanding of the diverse brain networks that implement NLC and their temporal dynamics. Neuroimaging research shows that during discourse comprehension, brain networks are brought online and offline dynamically, depending on the content comprehended, and that comprehenders' preferences with respect to the type of information on which they focus shape the organization of activity in regions central to comprehension. In addition, brain regions considered outside the canonical language network, such as, for example, posterior midline areas of the human brain often involved in vision, play crucial and largely ignored roles in comprehension. Findings from other neurobiological domains further suggest that assuming a fixed functional language network may be a weak starting position. In developing this point we address several misconceptions often used to support the existence of functional networks for language. First, anatomical connectivity imposes only moderate constraints on functional connectivity or on networks deployed for specific tasks, so it is tenuous to argue that anatomical connectivity constraints result in invariant functional networks. Second, functional networks are inherently non-stationary, and their core topological features are strongly influenced by context, making it difficult to speak of 'a network'. Section 4 presents this argument for letting go of the idea of a core stationary anatomical-functional language network as a neurobiological explanatory construct.

## 1.2. Main aims and structure

Our main aim is to argue for a substantial shift in perspective in how cognitive scientists, who are consumers of neurobiological research, and cognitive neuroscientists who produce this research think about the neurocognitive basis of language comprehension as understood in natural contexts (we avoid here a discussion of production due to the limited neurobiological literature on the topic). We present a research agenda that can advance our understanding of the principles that organize brain activity during language comprehension. Beyond advancing neurobiological knowledge, this progress is also likely to challenge theoretical positions on language in the cognitive sciences. To this end we present a large body of experimental findings, but only in order to illustrate the utility of this approach, as we do not purport to provide an overview of how the brain organizes language.

---

[2] For readability, we omit references in this section unless necessary for tracking ideas' provenance, as the empirical findings are presented in length in subsequent sections.

There are various neurobiological models of language, addressing issues such as white matter structure, sublexical speech perception, the neural basis of several sub-functions of language, and the role of dorsal and ventral streams (e.g., Binder & Desai, 2011; Bornkessel-Schlesewsky & Schlesewsky, 2013; Fedorenko & Thompson-Schill, 2014; Friederici, 2012; Hagoort & Indefrey, 2014; Hickok et al., 2007; Zwaan, 2014). We do not attempt to review those here. Instead we call for a shift in perspective, which holds the promise of more accurately modeling linguistic behavior, such as the neurobiological markers of successful comprehension and memory for content, as well as grounding of language in more basic computations. This endeavor can open the way to tighter integration of language studies with other domains of cognitive neuroscience from which they have been traditionally isolated. As we show, ideally this should induce timely revisions in the ontology guiding current thought about what computations are the building blocks for neurobiological accounts of language.

## 2. Limiting language-specific interpretations

Computationally, language comprehension relies on processes that are generally described as ranging from lower-level acoustic/phonetic processing to higher-level discourse comprehension. Studies examining the neurobiological implementation of discourse-level processes run into two main interpretive difficulties. One difficulty is more general, and relates to explaining activation in areas that were not predicted as involved in the computation of interest. Here, researchers need to be careful to avoid what has been called 'reverse inference' (Poldrack, 2006). This is a logically invalid inference where activity in a certain region is interpreted in reference to a cognitive process that is frequently associated with the region, but that was not directly controlled or tested in the study.[3] For instance, finding greater hippocampal involvement in one language task may be interpreted as suggesting that this task is accompanied by greater memory encoding demands (even though hippocampal activity can reflect novelty responses to verbal stimuli, Grunwald, Lehnertz, Heinze, Helmstaedter, & Elger, 1998). This problem is more prevalent in studies manipulating higher-level discourse features than, for example, in studies manipulating lower-level acoustic features, as these are not as likely to prompt differences in activation outside sensory regions.

The second problem pertains to the tendency to interpret brain activity during comprehension in terms of computations that are specific to language processing at the discourse level (e.g., evaluation of semantic constraints between a sentence and preceding sentences). With some notable exceptions (such as work based on oscillatory underpinnings of language), many studies of the neurobiology of language interpret brain activity in terms of complex language-specific computations – what we call a language-essential interpretation. For example, Ben-Shachar, Hendler, Kahn, Ben-Bashat, and Grodzinsky (2003) found that Broca's area (posterior left inferior frontal gyrus) differentiated sentences containing syntactic movement from those that did not, and interpreted this finding in reference to neural reality of syntactic transformations (p. 433). Only later work adjudicated between this syntactically-centered interpretation and a more basic one based on working-memory demands (see Santi & Grodzinsky, 2007). There is a long-standing work in the neuropsychology literature on whether Broca's area is necessary for sentences considered to contain traces (e.g., Caramazza, Capasso, Capitani, & Miceli, 2005), and even in manipulations of object vs. subject extraction (Caplan, Stanczak, & Waters, 2008), more complex structures do not necessarily involve this region. The involvement of Broca's area in sentence comprehension has been

interpreted in multiple ways since Ben-Shachar et al.'s study (for a review, see Rogalsky & Hickok, 2010).[4] This shows how language-essential interpretations can overlook more basic or domain-general computations that can similarly, and more parsimoniously, account for the differential involvement of a brain region given different types of language contents or tasks. The potential impact of reverse-inferences and language essential interpretations is compounded by the relative sparsity of such studies which has not allowed in depth investigation of many related questions.

Identifying core computations is an end in itself from the perspective of basic research, and can also result in a more complete understanding of co-occurring cognitive deficits in clinical states or explain correlations with cognitive deficits in apparently different domains. As detailed below (Section 2.1), effects of plausibility, consistency, predictability and event-change are often interpreted in language-essential terms in neurobiological studies. However, some results indicate that these effects are more parsimoniously explained by assuming basic computations that are not limited to language comprehension, but operate similarly over stimuli lacking semantics or meaningful syntax.

### 2.1. Basic functions for narrative construction, event segmentation and prediction

What evidence supports the claim that some computations considered crucial for discourse comprehension are better understood in terms of more basic functions? Neurobiologically, integration of spoken or written narratives into a meaningful structure has been linked to a network that includes perisylvian temporal and left inferior frontal regions, but also the medial prefrontal cortex and the precuneus (e.g., Ferstl, Neumann, Bogler, & von Cramon, 2008). Whether these regions are involved in the construction of non-linguistically communicated narratives has received considerably less consideration, but some data support this possibility. Humphries, Willard, Buchsbaum, and Hickok (2001) used fMRI to examine the processing of narratives communicated via sentences or via sequences of environmental sounds (gun shot, footsteps fading). They found that beyond the expected activity in lower-level regions that process any auditory input, both types of narratives induced activity in regions linked to higher-level language processing, including posterior lateral temporal and inferior frontal regions. Thus, comprehension of narratives, communicated by any means, may rely on regions typically associated with language processing.

Consistent with this possibility, electrophysiological signatures for integration of meaningful but non-linguistic information are similar to those found for language. Silent sequences of meaningful visual events produce signatures of integration-difficulty that are similar to those of semantic integration and syntactic repair in language studies (a N400 and P600-like ERP potential, Sitnikova, Holcomb, Kiyonaga, & Kuperberg, 2008). Integrating a word or picture with a preceding sentence produces similar evoked responses (Willems, Ozyurek, & Hagoort, 2008). This integration process is not limited to language contexts: when musical pieces are followed by words that match or mismatch the atmosphere evoked by the preceding music, mismatched words produce a stronger N400 electrophysiological response. The similarity of these patterns to those found for difficulty in semantic

---

[3] This is not to say that any abductive inference, or inference to the best explanation, is necessarily normatively invalid, but that any such inference should be made only after careful ruling out all other possibilities (Harman, 1965).

[4] These include mediation of sequential-related computations in linguistic and non-linguistic elements (Fiebach & Schubotz, 2006; Hoen, Pachot-Clouard, Segebarth, & Dominey, 2006), implementation of semantic unification at the single sentence or discourse level (Hagoort, 2005, Hagoort, 2017), or controlled retrieval of meanings in more demanding situations (e.g., Novick, Trueswell, & Thompson-Schill, 2005), to name a few. It is still an ongoing question whether the region mediates a single high-level function, multiple functions in different subregions, or a core though yet-undefined, lower-level computation that could account for the diverse sets of activation patterns found in this region (Rogalsky & Hickok, 2010).

integration suggests that integration vis-a-vis non-linguistic contexts may rely on similar processes (Koelsch et al., 2004). Finally, musical contents can also generate N400 responses. Daltrozzo and Schon (2009) presented participants with words that were followed by a musical piece, and found that N400 responses to the music tracked its perceived fit with the preceding word. This suggests that music can convey conceptual information, which is evaluated by neural computations that also underlie language processing. This set of findings, as a whole, is most parsimoniously explained by an account on which language-evoked semantic integration is implemented by a general mechanism for mapping new information onto a relevant context. This basic computation is sensitive to the contextual fit of incoming linguistic content (Hagoort, 2005; Hagoort & van Berkum, 2007), but importantly, we argue that it is not uniquely implicated in language processing.

Segmenting streams of information into series of events is another cognitive process known to organize discourse, narrative and film comprehension. According to Event Segmentation Theory (Richmond & Zacks, 2017; Zacks, Speer, & Reynolds, 2009), people parse information streams into units of action, or event models, that guide comprehension. These models allow predictions about future actions by relying on long-term schematic knowledge (Schank & Abelson, 1977), and failures of such predictions form an internal signal for an event boundary. For instance, once people learn that someone filled gas from the pump, long term-knowledge would predict that this person would go on to pay for the gas rather than just drive away, because paying is part of the 'filling gas' event model. In support of this model, behavioral studies have shown that shifting a discourse topic along the temporal or spatial dimensions produces processing costs (Radvansky & Copeland, 2010; Zwaan, Madden, & Whitten, 2000), perhaps due to increased effort involved in updating a new segment of a situation model (Speer & Zacks, 2005). Neurobiological studies have shown that event shifts in narratives are associated with activity in a network consisting of mainly posterior brain regions (Speer, Zacks, & Reynolds, 2007; Whitney et al., 2009). This network overlaps with a more extensive one that tracks event-shifts in movies (Zacks et al., 2001), and there is some evidence (Baldassano et al., 2017) for an overlap between networks involved in segmentation of auditory and audiovisually communicated narratives.

Could such findings can be accounted for via lower level mechanisms unrelated to semantic or episodic memory? Work by Tobia, Iacovella, Davis, and Hasson (2012) suggests this might be the case. In that study, participants listened to tonal series in which transition probabilities between tones were non-stationary and changed over time. Participants indicated when they noticed changes in the series' regularity, and points of cross-participant consensus were labeled as common subjective change-points. Brain activity in temporal windows prior to these change-points was compared to control epochs, defined as epochs that were not followed by change-point indications. This analysis identified a network strongly overlapping with that found in the event-change literature summarized above. These findings are in fact quite consistent with Zacks' Event Segmentation Theory, which is agnostic about whether the operations of working memory access, prediction and the evaluation of prediction are limited to semantic content. However, these findings do suggest there is a basic function that evaluates features of ongoing and prior contexts to identify contextual change, which operates across multiple domains. This activity can account, at least in part, for activation patterns seen in event segmentation studies.

Prediction is another process that is considered integral to language comprehension (see Huettig, 2015 for review).[5] Beyond the well-

established finding that the processing of words appearing at the end of sentences reflects their sentential predictability (e.g, Wlotko & Federmeier, 2012), electrophysiological studies have shown that, when contextually licensed, predictions are generated before word presentation (DeLong, Urbach, & Kutas, 2005; van Berkum, Brown, Zwitserlood, Kooijman, & Hagoort, 2005; but see renewed debate in Nieuwland et al., 2018). It has been suggested that the architecture that supports the construction of predictions in the brain is a general one, based on a hierarchical cascade of predictions made at different levels of granularity, which are evaluated and refined based on bottom-up input (Friston, 2009; Rao & Ballard, 1999).

The neurobiological question pertains to the neural basis of semantic predictions: do these rely on a language-specific predictive architecture or alternatively, on a more basic system? If the latter were the case, it would be sensible to first understand what are the features of the basic system and then (or in tandem) see how this system supports predictions in language. Indeed, recent neurobiological work provides converging evidence for the hypothesis that linguistic predictions are based on a lower-level prediction-generating and evaluation system that also implements sub-lexical predictions. First, certain brain networks appear to have the capacity to generate predictions in an abstract, domain-general manner and may subserve prediction in both linguistic and non-linguistic domains. For instance, activity in the anterior part of the right temporal cortex differentiates between predictable and unpredictable sequences for both auditory and visual inputs (Nastase, Iacovella, & Hasson, 2014). Second, lateral temporal regions implicated in low-level auditory speech processing (left planum temporale, and medial transverse temporal gyrus) have the capacity to engage in low-level predictions and their evaluation. Tremblay, Baroni, and Hasson (2013) found that these regions were sensitive to the predictability of auditory input streams, independently of whether these consisted of speech or non-speech tokens, and importantly, no temporal regions showed sensitivity to predictability in speech series alone, thus suggesting that there is no unique status for predictability in speech stimuli. These same low-level areas also signal surprise, as shown by Mustovic et al. (2003). In that study, unexpected periods of silence embedded within a stimulus stream were associated with increased activity in auditory cortex. To conclude, auditory cortex and nearby association cortices are involved in the construction of predictions in generating surprise (error) signals (sometimes termed a prediction error term'). In all, regions in the vicinity of primary auditory cortex have the capacity to implement non-semantic prediction, when these are licensed by statistical patterns in recently encountered input.

Predictions during language comprehension can be based on numerous types of information that are not unique to the language domain. These include the base-rate frequency of possible continuations, or their sequential transition probabilities. The former reflects the overall context-independent frequency of a token, and the latter its probability given some contextual constraints such as its first-order transition probability. To illustrate, the phrase *there is* is more likely to be followed by *a* than by *an* since *a* is more frequent in the English language. Still, the phrase *he flew* is still more likely to be followed by *an* than by *a* due to transition constraints. Behavioral studies have shown that people are sensitive to both these types of statistical features in language input (Ellis, 2002). But sensitivity to these sources of predictability is not limited to the language domain: Tobia, Iacovella and Hasson (2012) found that while people passively listened to tonal stimuli, lateral-temporal regions were sensitive to both the base-rate (marginal) frequency of the tones and their transition probabilities, though in different subregions. The aforementioned studies are just a sample from a large body of evidence suggesting that predictions for auditory inputs are generally mediated by lateral temporal, inferior frontal, and basal ganglia regions (for review, see Hasson & Tremblay, 2016). These findings support a different level of explanation for neurobiological accounts of predictability or regularity during language comprehension. Rather than assuming that these areas are involved in

---

[5] We note that the terms prediction and predictability are often time used when accounting for any type of differential responses observed to less vs. more-predictable stimuli, without making strong claims about whether this is a result of anticipatory prediction, or post-stimulus backward checking post-diction.

computations that are specifically sensitive to language, such as representation of language statistics, or construction of predictions based on semantic content, such findings point to domain-general functions that may generalize over the specific features of the input stream. To illustrate, effects of syllable-stream regularity on auditory cortex activity has been interpreted as a signature of online-word segmentation processes during language learning which aid cracking the language code (McNealy, Mazziotta, & Dapretto, 2006), but may reflect a basic level computation subserving multiple domains.

To summarize, the capacity to construct an integrated representation of continuous discourse and evaluate it for consistency, segment it into event-like structures, or construct and evaluate predictions may be supported by brain regions that perform related basic computations outside the context of language input. Interpreting such effects in terms of computations that are dedicated to processing language (or inferring that deficits in these abilities relate to a language system) should therefore be approached with caution. In the next section we target a different reason for careful consideration of language-essential explanations for activity patterns during NLC, which is that activity fluctuations during comprehension reflect other processes that occur within the same networks thought to mediate language comprehension.

### 2.2. On the neurobiological separability of interpretation, memory encoding and contextual re-instantiation

The relation between comprehension of and memory for discourse level content is complicated, and has received considerable treatment in the behavioral literature (e.g., Gurevich, Johnson, & Goldberg, 2010). Memory for linguistically communicated information is largely mediated by construction of higher-level situation models that are abstracted from the linguistic input, and it is these models that are encoded to memory rather than the sensory signal itself (Bransford, Barclay, & Franks, 1972). Some models of discourse memory argue it is represented via conceptual networks where nodes stand for events mentioned or implied by the text (Zwaan, Langston, & Graesser, 1995). The strength of the connections between these nodes reflects whether the events in question share dimensions (e.g., a shared character, location, etc.). The Landscape Model of comprehension (e.g., Tzeng, van den Broek, Kendeou, & Lee, 2005) is a computational model in which comprehension consists of recurring update-cycles with each new statement. In each cycle, prior concepts are strengthened or weakened. This allows modeling dynamic changes in concept accessibility as well as the final episodic representation. The model accurately predicts human recall, order of recall and the activation of textual concepts. Thus, when modeling memory for discourse content, it is natural to merge the processes of comprehension and the construction of episodic representations, as the latter reflect connection strength between elements in the discourse model (e.g., agents, goals, events, or more general concepts).

Given the opaque boundary between comprehension and memory encoding in cognitive processing and computational modeling, it is natural to inquire to what extent there exists a separation, at the neurobiological level, between systems related to discourse-level comprehension and systems linked to encoding. In addressing this issue, we evaluate two possibilities. The first is that discourse-level comprehension is implemented in systems mediating semantic processing, but that its encoding to memory is mediated by separate systems (the "segregation hypothesis"). The second possibility, which is consistent with the findings we mentioned above, is that memory for discourse content *just is* a reflection of activity in the areas mediating interpretation (the "single-process" hypothesis). As we review below, there is substantial support for the single-process hypothesis, in addition to some partial support for the segregation hypothesis, both in terms of the timeline of memory formation for semantic content, and the brain regions mediating encoding.

EEG studies show similar timelines for memory formation and

semantic integration: activity patterns within 400–700 ms following word presentation are indicative of whether people will subsequently correctly recognize or recall that word, suggesting a strong temporal proximity to the semantic integration timeline (see Paller & Wagner, 2002 for review). Neuroimaging studies have found that brain regions whose activity predicts successful memory for language content strongly overlap with those typically taken to mediate discourse integration per se. Memory encoding for discourse content involve bilateral temporal and left inferior frontal regions (Hasson, Nusbaum, & Small, 2007). In these regions, higher activity during story comprehension is associated with better subsequent memory for story content (see also, Hasson, Furman, Clark, Dudai, & Davachi, 2008). From a neurolinguistic perspective, it might appear surprising to find that higher activation in language areas is associated with better memory, as these regions are not typically reported in studies examining memory encoding, and are not considered (by neurobiologists of language) to mediate memory functions. Crucially, the same study (Hasson et al., 2007) also found that beyond the relation of lateral temporal regions to subsequent memory, these regions also responded more strongly to contextually inconsistent vs. consistent story endings. Beyond supporting the single-process hypothesis, such findings are merely one of many examples showing that successful memory formation does not depend solely on a general (hippocampal, medial-temporal) memory system, but also on brain regions specifically involved in processing the task stimuli. In a pivotal study, Otten and Rugg (2001) showed that when written words were presented for a phonological-judgment task, activity in one group of brain regions predicted subsequent memory, whereas when the exact same words were presented for a semantic task, activity in a different group of regions predicted subsequent memory. This and many subsequent studies show that the networks implicated in memory encoding differ on the basis of the processing performed during task (see also Duarte, Henson, & Graham, 2011; Gottlieb & Rugg, 2011; Gottlieb, Uncapher, & Rugg, 2010).[6]

A similar argument for a single memory/comprehension process can be made for the position that comprehension and memory-related access to prior discourse content are subserved by the same neurobiological system. It has been shown that during comprehension, the processing of incoming information triggers an automatic re-instantiation of entire chunks of prior context related to this information (e.g., Gerrig & McKoon, 2001). Currently, the brain regions that mediate access to recently encountered content or the re-instantiation of prior episodes during discourse comprehension are not known. However, two research paradigms, that of contextual reinstatement and that of repetition of language content, suggest that lateral temporal and inferior frontal regions may be involved in these processes.

Studies of contextual reinstatement use pattern classifiers to determine whether processing a stimulus produces a brain activity pattern that is indicative of the episodic context in which this stimulus was previously encountered. That is, can the specific context in which an item was initially presented be 'read off' from activity prompted by recall (see Danker & Anderson, 2010, for review of various contexts manipulated)?[7] Using fMRI, Johnson, McDuff, Rugg, and Norman

---

[6] This should not be taken to imply that memory for verbal information is independent of brain regions (hippocampal and ventromedial prefrontal cortex) that mediate short-term encoding and longer-term consolidation. The hippocampus (particularly the anterior section) is sensitive to semantic manipulations (Daselaar, Fleck, & Cabeza, 2006), and structural studies have shown that individuals with greater verbal memory ability have particular thickness profile in the region (Pohlack et al., 2014).

[7] In this paradigm, individuals' brain activity is first recorded with fMRI while they encode items to memory via different strategies. For example, they can be asked how an artist would draw the object, or alternatively, how the name of the object would be pronounced. Then, a pattern classifier is trained to distinguish between fMRI brain-activity patterns during these encoding contexts. Later, fMRI data are collected while the same individuals make old/new

(2009) showed that recognition activity in several left hemisphere regions often linked to language comprehension contained information about the encoding context (inferior frontal gyrus, posterior superior temporal sulcus, posterior middle temporal gyrus, but also superior frontal gyrus). Thus, upon re-encountering an item, brain activity contained signatures of the context in which it was initially encoded, indicating access to a memory trace of prior context. This conclusion is consistent with neuroimaging studies investigating repeated comprehension of sentence content. Repetition of a sentence, even when its content is repeated non-verbatim, results in decreased activation in lateral temporal regions and left inferior frontal regions (e.g., Dehaene-Lambertz et al., 2006; Devauchelle, Oppenheim, Rizzi, Dehaene, & Pallier, 2009; Hasson, Nusbaum, & Small, 2006). The magnitude of this activity reduction in lateral temporal cortex is inversely related to the temporal interval between the initial and repeated sentence presentations (Hasson et al., 2006), suggesting these regions have access to memory traces of recently encountered content, which decay over time.

To summarize, lateral temporal and frontal regions, which are usually associated with linguistic rather than memory functions have an important role in both memory encoding and contextual re-instatement during discourse comprehension.

## 3. Taking context seriously

Broadening the notion of context is necessary for explaining the neurobiology of naturalistic language comprehension. We first review studies suggesting there exists a domain-general mechanism responsible for integrating incoming language content against multiple contextual aspects. We then argue that studying comprehension in context is essential for obtaining valid neurobiological accounts of language, and is not a mere enhancement. Asking individuals to comprehend sentences outside context requires them to explicitly ignore those communicative aspects of language that are core to everyday comprehension. This could lead to underutilization of brain networks important for comprehension. Finally, neurobiological accounts must deal with a different form of context – neural context – through which prior brain states as well as contemporaneous activity in certain brain networks affect ongoing comprehension (Section 3.4).

### 3.1. Context rapidly impacts comprehension

Language comprehension is affected by different types of contexts. We describe three: semantic context, social context, and personal context, and first address neurobiological data showing that they are integrated within the same time window. We draw on evidence from an electrophysiological measure that has been shown to be extremely useful for studying the timeline of semantic processing: the N400 (already mentioned in Section 2; for reviews see Hagoort & van Berkum, 2007; Kutas and Federmeier, 2011; van Berkum et al., 2005).

*Semantic context* is any prior knowledge evoked by the incoming text. This prior knowledge can be part of long-term world knowledge, or knowledge conveyed by earlier narrative or discourse content. Relevant world knowledge has been shown to evoke a stronger N400 when it is inconsistent with the meaning of a sentence. For example, among Dutch readers, who know that Dutch trains are yellow, the final word in the sentence *Dutch trains are white* evokes a greater N400 than the final word in the sentence *Dutch trains are yellow* (Hagoort, Hald, Bastiaansen, & Petersson, 2004). Thus, world knowledge is quickly brought to bear on current sentence integration (see Hagoort & van

___
recognition judgments for these items. The crucial test is applying the classifier (constructed from the encoding data) to the data collected during recognition. This determines if there are brain areas whose activity during recognition contains information diagnostic of the encoding context.

Berkum, 2007; van Berkum et al., 2005 for reviews). Prior narrative context that is related to incoming information is also rapidly integrated. For example, narratives that predicate a certain attribute about a character (e.g., being slow) modulate the N400 amplitude when this character is later referred to in congruent or incongruent terms (slow/quick; van Berkum, Hagoort, & Brown, 1999; van Berkum, Zwitserlood, Hagoort & Brown, 2003).

Interestingly, when narrative context and world knowledge collide, narratives can have priority in integration. For example, Nieuwland and van Berkum (2006) studied narratives in which a critical word was congruent with world knowledge but incongruent with narrative context. By studying the N400 component, they concluded that such words were more difficult to integrate than words congruent with narrative context. To illustrate, in a story about an animated peanut, understanding that the peanut was in love was easier than understanding that the peanut was salted, as evident by a greater N400 for the latter. There is also analogous evidence from cartoon narratives, where seemingly impossible actions appear to be understood fluently (e.g., Filik & Leuthold, 2008). Thus, context, narrowly defined even within limited scopes, strongly impacts online integration.

*Social context* also influences semantic processing as measured by the N400. It includes features of the linguistic exchange that go beyond the linguistic structure and content and that are part of everyday language use in social settings. These include co-speech gestures, emotional intonation, and the speaker's identity or social status. For example, iconic gestures mismatching the meaning of a critical word in a sentence are associated with a greater N400 than matching gestures (e.g., mimicking walking vs. rolling for a sentence describing falling off a roof; Özyürek, Willems, Kita, & Hagoort, 2007). A child's voice uttering a sentence fit for an adult (e.g., about drinking wine) elicits a greater N400 effect than an adult's voice uttering the same sentence (van Berkum et al., 2008). Other work shows that being embedded in a social context *per se* can also impact processing, for instance the specific identity of the speaker or the presence of a fellow addressee rapidly impact language comprehension, within the time frame previously linked to semantic integration (Bornkessel-Schlesewsky, Krauspenhaar, & Schlesewsky, 2013; Rueschemeyer, Gardner, & Stoner, 2014).

*Personal context* includes factors that are independent of any linguistic exchange studied and are related to the comprehender's value system or self-perception. For example, statements that clash with personal values (e.g., positive or negative judgments about abortion) elicit a greater N400 than value-matching statements (van Berkum, Holleman, Nieuwland, Otten, & Murre, 2009). Similarly, self-referential words inconsistent with one's self-view elicit greater a N400 than consistent words (Watson, Dritschel, Obonsawin, & Jentzsch, 2007). In addition, whether negative or positive information is conveyed in a way that is more or less relevant to the reader (e.g. a man knocking on your or someone else's door with a gift or a gun) also affects the N400 response (Fields & Kuperberg, 2015).

Among these personal contexts, particular attention has been given to comprehenders' mood. Happy and sad moods influence a host of cognitive processes (for reviews see Clore & Huntsinger, 2007; Forgas & Koch, 2013). In general, they promote the processing of information congruent with their valence (e.g., Fiedler, 2001; Forgas & Locke, 2005). Happy mood is also thought to promote more comprehensive and top-down processing, whereas sad mood a more analytic and bottom-up processing style (e.g., Clore & Huntsinger, 2007; Fiedler, 2001).

Mood has been shown to impact comprehension of single sentences or sentence pairs (Chwilla, Virgillito, & Vissers, 2011; Federmeier, Kirson, Moreno, & Kutas, 2001; van Berkum, De Goede, Van Alphen, Mulder, & Kerstholt, 2013). In broader discourse, listeners induced into a happy or sad mood show increased N400 amplitude for mood-incongruent story endings, even when these endings are consistent with prior linguistic context (Chung et al., 1996; Egidi & Nusbaum, 2012). Chung et al. (1996) found that positive and negative story endings

elicited a greater N400 when they mismatched the valence of the mood, but the absence of a neutral mood control condition made it impossible to specify the impact of each mood on comprehension. This issue was addressed by Egidi and Nusbaum (2012) who showed that happy and sad moods induce different processing of negative information than neutral mood. Compared to the neutral group, N400 amplitudes to negative information were increased in the happy group and reduced in the sad group.

In all, these wide-ranging studies on the impact of semantic, social and personal context point to an exciting and non-intuitive conclusion, which is that multiple sorts of contextual features affect comprehension *within the same time frame*. This poses difficulty for a functional separation between semantic, social and personal contexts. Yet, both social and personal contexts are often viewed as extra-linguistic, despite the increasing evidence from the studies cited above (and others) which show that such contexts influence comprehension in a similar manner and timing to that of sentence-level cloze probability or semantic-association strength.

### 3.2. Syntax in context

While it may appear that taking context seriously is of most importance to neurobiological theories of *semantic* processing, in this section we argue that taking context seriously is also important for neurobiological accounts of syntactic processing. Behavioral studies have shown that syntactic operations are influenced by contextual factors and, to some extent, the debate on whether there exist encapsulated syntactic processes that are context independent is one that has been settled in the behavioral literature. The main neurobiological interest is in whether there are brain regions that code for syntactic features alone (in a context independent way), and if not, what would be the best way to investigate the neurobiology of syntactic processing. As we review below, it is difficult if not nonsensical to draw conclusions about purported syntactic processing in the brain in a way that is independent of context. Beyond the specific implications for neurolinguistic theories of syntax, this raises the issue of how to best address this neurobiological question, and highlights the fact that some conclusions drawn to date from context-independent studies of syntax may lack explanatory power for understanding syntactic processing as it occurs in natural, contextualized language.

The history of this debate in the behavioral literature is well known and discussed extensively elsewhere (see Altmann, 2013). However, as there are parallels to the debate in the neurobiology of language, we briefly recap it here. In their landmark study, Ferreira and Clifton (1986) concluded that syntactic comprehension is informationally encapsulated, in the sense that providing context, e.g. semantic information supportive to a certain reading of the sentence, did not alter early processing of sentences containing non-minimal attachment. However, it was later shown that the difficulty of these sentences is reduced by disambiguating contexts, as contextual information implemented via discourse influences the resolution of such syntactic ambiguities (Altmann & Steedman, 1988; Spivey-Knowlton & Sedivy, 1995). For instance, relative clauses might have a preferred reading when presented in isolation but another reading when used in context (see Spivey-Knowlton, Trueswell, & Tanenhaus, 1993 for related discussion). The very large literature on the Visual World Paradigm shows just how rapidly and effectively contextual features impact language comprehension, and some of these constraints operate on the discourse level indicating that individuals quickly refine senses of verbs and prepositions based on contextual dependencies (see Huettig, Rommers, & Meyer, 2011, for review). That said, some researchers (e.g., Friederici, 2017) suggest that meaning construction consists of a 'bottom-up', syntax-driven phrase structure construction (captured by an early left anterior negativity component; ELAN) that is thought to be automatic and to precede semantic analysis. For both psychologists and neurobiologists, the consideration of context is not merely an experimental

technicality, but an issue that calls into question the type of conclusions that can be drawn from decontextualized studies. If understanding such sentences out of context omits a natural factor that impacts comprehension, what is the status of conclusions drawn from brain responses to those isolated sentences? This concern is oftentimes overlooked in cognitive neuroscience of language where findings obtained from studying isolated sentences are assumed to hold for language comprehension in everyday life.

An example that shows the problem with this assumption and that specifically targets the influence of prior context on neural activation is an fMRI study of syntax by Kristensen, Engberg-Pedersen, and Wallentin (2014). In that study, participants were presented with subject-initial sentences (*He noticed her*) and object-initial sentences (*Her, he noticed*). Both are grammatically correct in Danish, the language in which the study was performed. However, according to syntactic theories, object initial sentences are more complex and are more difficult to process. These sentences were presented in two conditions: one absent of context, and one where a brief preceding context rendered the (more complex) object initial sentence as more acceptable. For instance, when preceded by the sentence *Peter overlooked all the shoplifters except Anne*, it is the object initial continuation *Her, he noticed* that is more natural. The authors found that activity in left IFG was strongly reduced when context licensed the object initial continuation. These results illustrate that findings from single-sentence, out-of-context studies of language, do not necessarily extend to findings obtained with language materials that are contextualized, even if the context is rather minimal (see also Mak, Vonk, & Schriefers, 2008). Another study suggesting that syntax is processed differently when comprehenders encounter more mundane language is fMRI work by Brennan and colleagues (Brennan et al., 2012), in which the left IFG showed no sensitivity to syntactic distance when participants listened to stories. Rather, the anterior temporal pole was sensitive to this established marker of syntactic difficulty.

Ongoing (e.g., Leiken & Pylkkanen, 2014) and future neurobiological work examining phenomena considered at the semantic-syntax interface is very likely to identify to what extent syntactic manipulations interact with semantic features of the stimuli. In any case, existing findings already show how difficult it is to account for brain activation patterns solely via syntactic-based explanations (e.g., Keller, Carpenter, & Just, 2001; Leiken & Pylkkanen, 2014). For these reasons, the study of de-contextualized, complicated sentences cannot be treated as a valid small-scale model of natural comprehension, as it may simply be a wrong model of semantic and syntactic processing in natural contexts.

The need to isolate syntactic processes has produced a field of study often examining very artificial stimuli. In the extreme, this produces comparisons between well-formed sentences and ill-formed sentences that contain syntactic violations. The underlying assumption is that processing a syntactic violation taxes the syntactic processing system just quantitatively more than syntactically well-formed ones rather than instantiating a qualitatively different mode of processing. But even when outright syntactic violations are avoided, participants may still be asked to deal with complex sentences high on syntactic load, such as those with long distance dependencies. One can spend some time understanding who pinched whom and who loves whom in *Kate loves the woman who the mailman and the mother of Jim pinched* (Santi & Grodzinsky, 2007). Garden path sentences such as *The horse raced past the barn fell* (Bever, 1970) are similarly taxing due to the need to revise an initial syntactic interpretation. In neuroimaging studies, such complex sentences invoke greater activity in left IFG (for review, see Kaan & Swaab, 2002). This has been taken to suggest a role for this area in syntactic processing. However, it has been known from relatively early neuroimaging work (Caplan et al., 2008) that responses to more vs. less syntactically complex sentences are themselves modulated by semantic constraints.

In short, the two strategies of ignoring context in the study of syntax and of pushing the system to the limits are backfiring. Their

generalizability is called into question, as it is not clear to what extent findings obtained with these strategies reflect processes that take place during everyday language comprehension. This of course does not mean that it is impossible for findings from traditionally controlled studies to generalize to more naturalistic settings (e.g., see Demberg & Keller, 2008). But which will or will not remains an open question until more research is done using the naturalistic approaches advocated here.

Another limitation of constructing manipulations along dimensions derived from traditional generative linguistic approaches is that it does not adequately consider either approaches that are usage-based (Goldberg, 2003; Ibbotson, 2013) or theoretical approaches for explaining language that are essentially syntax free, of the sort we describe in Section 5. As recently reviewed by Frank and Christiansen (2018), their is substantial recent work suggesting that the presence of hierarchical syntactic structures should not be assumed a priori, and that sequential explanations may suffice (for recent neurobiological work, see Lopopolo, Frank, Van den Bosch, & Willems, 2017). Furthermore, as suggested recently (Frégnac, 2017) rather than choosing stimuli that can produce strong firing rates, presenting a larger spectrum of input statistics and examining the system's response to those may be more appropriate.

### 3.3. Context bridges the gap between the processing of language and communicative intent

While language can be used for many purposes, perhaps its most important function is for communication. It has been argued that ontogenetically and phylogenetically, the urge to communicate precedes linguistic abilities (Liszkowski et al., 2006; Sperber & Wilson, 1995; Tomasello, 2003). The common theme is that language has evolved as a highly efficient code for sharing information, driven by the need to communicate. Research shows that producing and recognizing *communicative intent* at least partially relies on brain areas outside of the traditional language network.

Despite the importance of successful communication as a core function of language, studies in the cognitive neuroscience of language tend to focus on the core properties of language as traditionally defined in linguistics: phonology, semantics, and syntax; what has sometimes been referred to as the "faculty of language in the narrow sense" (Fitch, Hauser, & Chomsky, 2005). However, from a broader perspective, it is important to consider that successful communication only takes place when someone, by means of acting in a certain manner, conveys his or her intention to someone else, and when that second person understands the intention (Levinson, 2006; Sperber & Wilson, 1995). Even when no speaker is present, such as in written texts, people attempt to model the state of mind of characters or of the author (Gerrig, 1993). Typically, the study of this deciphering/coding of communicative intent is delegated to the domain of pragmatics, which is not considered at the heart of the cognitive neuroscience of language (but see Bara, 2010; Galantucci et al., 2011; van Berkum, 2009, for counterexamples). Indeed, in a typical study, language content is presented under the assumption that participants can process only what has been called the 'descriptive aspect' of a sentence (roughly corresponding to the state of affairs referred to), while ignoring the other aspects that make sentences effective ('speech acts'). Namely, it is assumed that participants can put aside what has been called the *purpose* of the speech act, which entails recognizing the intention of the speaker, his/her attitude towards the listener, and the implied request to respond to the statement. The way in which core linguistic processes lead to actual communication is something that has been largely overlooked in the cognitive neuroscience of language. Recently however, several studies have examined how the generation or understanding of communicative intentions is processed in the brain.

From relatively early on, neuroimaging work provided clues that the areas that are involved in deciphering intentional abilities are different from those typically associated with linguistic computations narrowly

defined. This line of work took as its starting point the study of 'Theory of Mind', which pertains to an individual's ability to understand someone else's state of mind. While initial studies used linguistic stimuli, areas identified in those studies are activated by non-verbal paradigms, such as watching video animations of simple geometrical shapes (see Schurz, Radua, Aichhorn, Richlan, & Perner, 2014, for review) or viewing comic strips, suggesting they perform common functions for both linguistic and non-linguistic contexts. These studies show that reflecting upon, deciphering, or making inferences about another person's belief activates areas outside of the classical language network, but the studies do not speak to the issue of whether core linguistic and communicative abilities are separable or not.

In an fMRI study addressing this issue (Willems et al., 2010), participants engaged in a variant of the 'Taboo game', in which one person describes a concept to another person, but without using certain words. For instance, one tries to describe 'beard' without using the words: man, face, hair or shave. The experimental design had two factors. First, communicative intent was manipulated by having participants construct a description when the other person already did or did not know the answer. When the interlocutor did not know the answer, the participant had to produce a description for this specific concept. When the interlocutor did know the answer in advance, no targeted description was necessary. The second experimental factor manipulated the semantic distance between the target word and the taboo words, with increased distance necessitating managing lexical competition. Planning and speaking durations were indeed influenced by the latter factor, with greater lexical similarity producing slower behavior, and this was accompanied by differential activation in areas classically related to core language processes, such as the left inferior frontal cortex. The communicative manipulation, on the other hand, did not impact behavior, but produced different activations in the MPFC, part of the mentalizing network. The main conclusion is that linguistic and communicative functions rely on distinct neurocognitive substrates (see Willems and Varley, 2010, for review).

From the perspective of comprehension, deciphering intent (within and outside linguistic communication) may similarly rely on systems not core to language comprehension. This was evaluated in a study (Walter et al., 2004) that distinguished between interpretations of private intentions and communicative intentions. Participants viewed short cartoon narratives where protagonists performed actions motivated by private, non-communicative intentions (e.g., changing a broken light bulb to allow reading) or narratives where protagonists acted with a clear communicative intent (e.g., pointing to a bottle to request it). They found greater activity in MPFC for the communicative than for the non-communicative stories (see also Ciaramidaro et al., 2007). Kuhlen, Bogler, Brennan, and Haynes (2017) examined brain activity patterns in two speech-production conditions: one where participants thought a conversational partner would hear them, and another introduced as a technical calibration in which a conversational partner was absent. Analyzing activity in the seconds that preceded the speech itself, they found that activity in MPFC and ventral prefrontal cortex (bilaterally) differentiated the two conditions. We note however that this is a newly explored issue and some work has implicated perisylvian regions in such interpretive functions (Egorova, Shtyrov, & Pulvermuller, 2016).

What is important for present purposes is the neurocognitive separability of understanding linguistic information and (implanting or recognizing) communicative intent. One could argue that this neurobiological fact necessitates separating the study of language from the study of communication-related aspects. However, we argue that given that communication is a core function of language, a full account of the neural basis of language should address systems involved in understanding communicative intent, and their potential interactions with systems mediating traditional language functions, rather than limit itself to the study of core linguistic processes using studies that essentially strip them of communicative purpose.

### 3.4. Neural contexts

In prior sections we emphasized that diverse types of external contexts should be taken seriously as a source of constraint on brain activity during language. Here we argue that, independently of interpretive concerns pertaining to online activity, understanding the neural basis of ongoing comprehension would strongly benefit from considering not only activity evoked by linguistic input, but also the fact that comprehension occurs in a *neural* context. That is, in our brains, activity in the 'here and now' is not determined solely by input features and computations triggered by it, but also by the prior brain state and by activity in areas unrelated to processing input features *per se*. There are at least two mechanisms by which this context can affect comprehension: (*i*) non-stationary activity fluctuations prior to stimulus presentation impact sensory and higher-level processes, and (*ii*) several functional brain networks implicated in the generation of internal content interact with comprehension in complicated ways.

The emerging view in cognitive neuroscience is that brain activity during stimulus processing reflects an interaction between stimulus-triggered processing and the endogenous neural state prior to stimulus appearance (for review, see Sadaghiani & Kleinschmidt, 2013). To illustrate, sensitivity to auditory stimuli is linked to intrinsic activity levels in lower-level auditory regions *prior* to stimulus appearance (e.g., Hesselmann, Sadaghiani, Friston, & Kleinschmidt, 2010), and phonemic restoration effects in noise are impacted by the state of the system prior to presentation of the ambiguous stimulus (Leonard, Baud, Sjerps, & Chang, 2016). Similar pre-stimulus activation patterns (as measured by EEG or fMRI) have been linked to stimulus memory (e.g., Otten, Quayle, Akram, Ditewig, & Rugg, 2006) and attentional performance (e.g., Li, Yan, Bergquist, & Sinha, 2007). The implication of such findings for neurobiological models of language – as for other higher-level cognitive functions – is that comprehension is likely not determined solely by computations triggered by the linguistic input.

More specifically, pre-stimulus fluctuations can impact linguistic computations because the functional networks implicated in language comprehension are not simply brought online in response to language input but maintain connectivity even during resting states (e.g., Hampson et al., 2006; Muller & Meyer, 2014). Endogenous fluctuations in cortical networks, which are rapid and highly non-stationary (Hutchison et al., 2013), could affect comprehension by altering the connectivity of these networks.[8] Thus, these networks' state prior to comprehension could impact their state during stimulus processing. The off-line maintenance of functional networks during rest may be due to the need to preserve the brain's capacity to react in a coordinated manner to frequently encountered types of stimuli, or could be related to the internal generation of semantic content; i.e., rumination or mind-wandering (Mason et al., 2007).

In summary, brain activity at any given point during comprehension likely reflects a combination of stimulus-triggered information processing, as well as prior constraints on comprehension instantiated in the form of activity in different brain networks.

## 4. Reconsidering the utility of a neurobiological language network as an explanatory construct

### 4.1. Neurobiological networks for comprehension

From a neurobiological perspective, several brain networks are implicated in contextual integration. These include one network typically considered fundamental to central language functions (left lateral temporal regions and left inferior frontal gyrus [IFG]), and a second network consisting of posterior and anterior midline regions implicated in the generation of associations, memory retrieval and situation models.

The IFG and lateral temporal regions have been associated, across different studies, with processing of semantic, syntactic, social, and personal information across contexts. For example, semantic predictability in discourse influences activity in left IFG, lateral temporal regions and right IFG (Ferstl et al., 2008; Hasson et al., 2007; Menenti, Petersson, Scheeringa, & Hagoort, 2009). Perisylvian regions are also associated with drawing inferences (Kuperberg, Lakshmanan, Caplan, & Holcomb, 2006; Mason et al., 2004) and detecting inconsistencies with world knowledge (Hagoort et al., 2004). Inconsistency in social signals such as between iconic gestures and speech (e.g., Dick, Mok, Raja Beharelle, Goldin-Meadow, & Small, 2014; Willems et al., 2008), or the emotional prosody of speech (Kotz et al., 2003) also impacts these regions. Finally, comprehenders' mood (Egidi & Caramazza, 2014) or the types of information people look for in a text (Cooper, Hasson, & Small, 2011) also impacts activation in these regions.[9]

Importantly, some contextual manipulations impact activity in brain areas outside the core language network, highlighting their importance for discourse integration. This has been shown in several neuroimaging studies that have employed a well-established behavioral paradigm (Bransford & Johnson, 1972) to identify brain systems associated with the construction of a coherent representation of narrative content. In this paradigm, an ambiguous text may or may not be accompanied by a disambiguating title or picture. When the title or picture is absent, the comprehension of the text is very difficult, to the point that even linking one sentence to the next becomes impossible (e.g., *Once you are settled, your thumbs should be pointing up. Sometimes there is no security but the animal's hair*, which is only clear when understood as horse-riding instructions; e.g., Martin-Loeches, Casado, Hernandez-Tamames, & Alvarez-Linera, 2008; St. George, Kutas, Martinez, & Sereno, 1999).

Brain regions showing greater involvement when discourse is preceded by a disambiguating title are of particular interest, as they are related to the normal process of linking incoming with prior information, and several studies have identified a specific network using this contrast. Martin-Loeches et al. (2008) documented greater activity in a set of posterior midline regions in the disambiguated condition (e.g., precuneus, posterior cingulate cortex, anterior cingulate cortex, but also angular gyrus, dorsolateral prefrontal cortex and insula) and associated those to "linking incoming information with a repository of activated knowledge" (Martin-Loeches et al., 2008, p. 620). Other work (Smirnov et al., 2014) did not replicate these patterns, but found that connectivity of several of those regions (posterior cingulate cortex, angular gyrus, superior frontal gyrus) with the mid- and posterior-left IFG was stronger during the disambiguating condition. The importance of these regions was also documented by research showing that, across individuals, temporal patterns of brain activation were more similar in the presence of the disambiguating title (specifically: anterior- and posterior-cingulate and medial prefrontal cortex; Ames, Honey, Chow, Todorov, & Hasson, 2014).

The networks used during integration also depend on semantic factors such as the relation between an incoming sentence and more- or less-recently introduced narrative context. In a study by Egidi and Caramazza (2013), participants heard stories whose final sentences were either consistent or inconsistent with the immediately preceding context (a local consistency relation). Independently, the earlier, initial

---

[8] This appears to be the general pattern for task-related networks: similar findings have been found for networks involved in social cognitive processes and tool use (Simmons & Martin, 2012) and action observation (Molinari et al., 2013).

[9] We note that although studies often mention the same region label, such references across different studies could refer to functionally dissimilar distributions. That is, normalization of fMRI data and the commonly employed practice of spatially smoothing the data renders it difficult to say with certainty that exactly the same region was activated in two separate experiments.

sentences in each story were also relevant or irrelevant to the final sentence (distal relevance relation). The study demonstrated that while people listened to the final sentence, three different networks were involved in comprehension. One signaled the relevance of the distal context, a second signaled the consistency of the ending with its local context, and a third was sensitive to the interplay of both the relevance of the distal context and the consistency of the ending with the local context. Other work (Egidi & Caramazza, 2014) showed that the processing of consistent vs. inconsistent discourse content was performed in very different networks depending on comprehenders' mood. This implies the existence of alternative routes to the basic language network for comprehension.

It is important to note that it has been clear for more than a decade at this point that language comprehension engages large-scale brain systems beyond those associated with "classical" perisylvian schemes. Ferstl et al. (2008) termed these an "extended language network", encompassing the anterior lateral temporal regions bilaterally, posterior superior temporal sulcus bilaterally, ventral and dorsal medial prefrontal cortices, anterior left IFG, and the precuneus. An even broader network was documented by Wilson, Molnar-Szakacs, and Iacoboni (2008) who identified brain regions that co-varied in a similar manner across individuals during auditory language comprehension. Beyond the regions identified by Ferstl et al., this study implicated regions such as posterior cingulate gyrus, right fusiform and parahippocampal gyrus that are involved in visual processing and memory.

Of course, discourse-level manipulations may impact activity outside a 'core' network via more generic processes that are unrelated to language processing. For instance, inter-subject correlations may reflect variations in "global levels of engagement", rather than language-related computations, as noted by Wilson et al. (2008). Some findings are consistent with these objections: Wide-spread activity changes in lower level auditory and visual regions is triggered by memory recall (Azulay, Striem, & Amedi, 2009), and large-scale changes in interactions between well-defined brain networks are linked to increased memory load (Liang, Zou, He, & Yang, 2015). Nonetheless, as our examples above show, there is extensive literature linking regions outside the core language network to specific semantic processes at the core of linguistic competence. As a case in point, in the next section we focus on the lingual gyrus, a region in the occipital cortex, which is traditionally associated with visual processing.

### 4.2. Outside the language network: The lingual gyrus as an example

The lingual gyrus is an occipital region, adjacent to the calcarine sulcus. It has a unique anatomical feature in that it is one of few non-perisylvian regions whose cortical thickness covaries with that of a cluster of lateral temporal, inferior parietal and inferior frontal regions often implicated in language functions (Chen, He, Rosa-Neto, Germann, & Evans, 2008). That is, individuals with greater cortical thickness in the left lingual gyrus also show greater thickness in many regions linked to language processing. Such correlations have been interpreted as indicating that these regions experience common activation (Chen et al., 2008; Zielinski, Gennatas, Zhou, & Seeley, 2010). In several neuroimaging studies, it is the only region or one of few outside perisylvian regions to show speech-related activity (Rodd, Davis, & Johnsrude, 2005; Zekveld, Heslenfeld, Festen, & Schoonhoven, 2006) or involvement in evaluation of grammaticality (Hoen et al., 2006).

This region also tracks rapidly changing semantic and structural features of language input. Brennan et al. (2012) used fMRI to identify brain areas that showed similar temporal fluctuation patterns across participants while listening to auditory stories. Many perisylvian regions showed this inter-subject correlation, but also the lingual gyrus and nearby regions. Further analyses showed this region's activity tracked lexical frequency of the auditory inputs. This does not however mean that the lingual gyrus mediates an essentially linguistic function, and related work (Tobia, Iacovella & Hasson, 2012) using non-linguistic

stimuli suggests otherwise. In that study, listening to simple tonal series was associated with inter-subject correlations in the lingual gyrus and several posterior midline regions (posterior cingulate cortex, precuneus). Furthermore, in these regions, the time course of common activity profiles, across participants, tracked changes in tonal transition structure. Thus, the involvement of this region in processing of a temporally extended stimulus could be generally related to integration of incoming and recently encountered information. The region might be mediating lower level memory processes as seen in the fact that it shows repetition suppression for words but not non-words (Fiebach, Gruber, & Supp, 2005), and for simple sentences but not complex ones (Hasson et al., 2006).

The mechanisms underlying the involvement of this region in language are yet to be fleshed out. What is important is that it is possible to identify brain regions outside the language network that can support important language related functions. To date, identifying such regions has been left as an exercise to those readers inclined to comb through the literature, while the main focus has been maintained on those results found within regions considered core to the language network. In this way, a large knowledge base has been accumulated for perisylvian regions, but remarkably less is understood about the role of other networks or regions.

### 4.3. Core language networks: stable, unstable, or non-existent?

In previous sections we have questioned the putative centrality of a 'language network' for understanding NLC. Here we ask whether it is at all reasonable to assume that discourse-level comprehension relies on a stable set of regions (a stable functional network) that implements discourse-level functions beyond those needed for lower-level speech perception. There are two reasons for the assumption that a stable functional language network exists, and we evaluate both in the following sections. First, lateral temporal and inferior frontal regions are strongly connected via white matter fiber tracts (Dick & Tremblay, 2012). Some have suggested that these structural pathways can be assigned language functions (Friederici & Gierhan, 2013), and that they form a general constraint on functional connectivity patterns (Honey, Thivierge, & Sporns, 2010). Second, several studies have documented synchronized activity between lateral temporal and frontal regions, both during the resting state and during language comprehension (e.g., Hampson et al., 2006).

Cognitive neuroscience has made great progress by identifying functional networks in the human brain. These, non-technically speaking, consist of brain regions whose time series are correlated during wakeful rest or during perception. This discovery has constituted a significant advance, with many tools devised for characterizing these networks. However, several implicit suppositions in this field have come under increasing scrutiny, with important implications for the study of NLC. Specifically, as we detail below, several studies have questioned:

1. Whether functional networks do reflect structural connectivity,
2. Whether these networks are stationary over time,
3. Whether the temporal correlations that define these networks can be taken to reflect continuously synchronized activity within the networks (as opposed to, e.g., rare coordinated spiking).

#### 4.3.1. Does structural connectivity constrain functional connectivity?

Initial studies emphasized that structural (anatomical) connectivity between brain regions can predict patterns of resting-state functional connectivity (e.g., Honey et al., 2010), and that resting-state connectivity patterns themselves may serve as a substrate for organization of different sorts of computations (e.g., Smith et al., 2009). But later studies showed that both these constraints are moderate. Specifically, structural connectivity accounts for 30–60% of the variance in functional connectivity (e.g., Messe, Rudrauf, Benali, & Marrelec, 2014),

and different tasks evoke markedly different organization of whole-brain coordinated activity (e.g., Mennes, Kelly, Colcombe, Castellanos, & Milham, 2012). Thus, the structural white-matter pathways linking lateral temporal, inferior parietal and inferior frontal regions do not necessarily support a stationary functional network, as brain structure is a constraint on functional activity, but far from a deterministic one.

Instead, these moderate structural constraints allow for the possibility that during comprehension, the 'core' functional network (typically implicated in lexical and sentence level processing) alters its connectivity with additional regions/networks depending on context. Regions supporting language-related working-memory operations (IFG, intraparietal sulcus) change their strength and directionality of connectivity with perisylvian regions depending on the syntactic complexity of sentences (Makuuchi & Friederici, 2013). Syntactic complexity is accompanied by stronger connectivity between inferior frontal and lateral temporal regions (den Ouden et al., 2012). In a study of discourse comprehension (Chow et al., 2013), participants heard stories that contained, in separate sections, strong action-, emotion- or visually-related content. Functional connectivity between temporal regions and other brain regions changed dynamically depending on the content heard. Similar findings are found for music perception: fMRI and MEG studies have shown that functional connectivity between auditory cortex and the medial temporal lobe (a region traditionally linked to memory encoding and retrieval) varies as function of familiarity with a musical piece (Müller et al., 2013; Wilkins, Hodges, Laurienti, Steen, & Burdette, 2014).

In fact, context-sensitive changes to functional connectivity appear to be the norm, rather than an exception. A large literature shows that even simple changes to cognitive state produce changes in functional connectivity networks. For instance, increasing working memory demands induces fundamental changes to functional connectivity in specific brain networks (e.g., Fransson, 2006; Fransson & Marrelec, 2008), and different types of discourse content impact connectivity of nodes in the default mode network (Hasson, Nusbaum, & Small, 2009). In a study of movie perception (Raz et al., 2014), connectivity of the limbic system with other brain network changed dynamically and tracked fluctuations in autonomic activity on timescales of around 30sec. Thus, rapid, content-driven changes in arousal covary with network dynamics. Altogether, the findings we have mentioned suggest that dynamic network reconfiguration is a basic feature of brain activity during comprehension.

### 4.3.2. The gap between functional connectivity and continuous information exchange

Functional connectivity is most frequently quantified via the correlation between time-series sampled from pairs of brain regions. The initial studies that identified large-scale functional connectivity networks in the brain (Fox et al., 2005) emphasized that connectivity is driven by low-frequency oscillations, accompanying those data by figures depicting similar time series across different brain regions. However, researchers have begun revising this implicit conceptualization of connectivity as reflecting continuous synchronization over time, and instead suggest that time-series correlations reflect infrequent bouts of joint activity. In a surprising demonstration that has been repeatedly replicated, it has been shown that the same functional connectivity networks identified via correlation procedures can be identified by a procedure that discards more than 90% of the data in the time series, and that instead is based on finding infrequent but spatially synchronized activity spikes across brain regions (Tagliazucchi, Balenzuela, Fraiman, & Chialvo, 2012). Conversely, removing the variance attributable to such spikes results in reduced correlations in well-defined networks (Petridou, Gaudes, Dryden, Francis, & Gowland, 2013). The second point has already been touched on when we mentioned studies (reviewed in Hutchison et al., 2013) showing that even within well-defined networks connectivity is non-stationary, with profiles changing at a rapid pace. This reconfiguration is not random, but can be

systematically described via a set of spatio-temporal activation patterns cycled through over time (Rashid, Damaraju, Pearlson, & Calhoun, 2014; Robinson, Atlas, & Wager, 2015).

Thus, the most appropriate level of explanation for functional connectivity is not that of correlations derived from temporally extended periods (typically on the scale of several minutes), but a much shorter scale. Technically, this is completely feasible and studies have shown that it is possible to quantify correlated activity in the brain on scales of 30 s or so (e.g., Shirer, Ryali, Rykhlevskaia, Menon, & Greicius, 2012). To summarize, while studies of correlated activity will probably continue to be of major interest in future work on NLC, it is important to reconsider whether quantifying functional connectivity within restricted networks as currently performed holds the key to understanding the information processing in brain networks during language comprehension. Alternatively, it may be that relatively infrequent co-occurring spikes driven by very particular input features drive such correlations. Identifying what are these features is an interesting direction for future work.

### 4.3.3. Questioning the concept of a functional language network

In the prior two subsections we argued that a better understanding of NLC could be gained by examining high frequency patterns of non-stationary correlated activity. Still, that discussion was premised on the idea that there is a set of regions that can be called a "language network", which supposes that, formally defined, such a functional network can be meaningfully said to exist. There is little doubt that perisylvian regions are more involved in processing language inputs than some other brain systems (see, e.g., Blank & Fedorenko, 2017), and may be particularly involved in language tasks (e.g., Fedorenko, Behr, & Kanwisher, 2011). This however is not sufficient for treating these regions as a functional network. This would be challenged if it were shown that core features of connectivity between brain regions linked to this network change with context. For example, if it were shown that depending on textual features, core topological features[10] of functional connectivity change in a way that leads to re-defining the network, it would be formally inaccurate to state that a fixed network implements comprehension. Instead, it would be accurate to say that different networks are established depending on the goals of comprehension.

To our knowledge, this argument has not been evaluated to date, or at least not reported in the literature on language comprehension. But outside the domain of language, there is an ongoing debate on whether functional connectivity networks do, or can, fundamentally change depending on context. Arguing for the relative fixedness of such networks, are reports showing that the modular partitioning of these networks, as quantified from fMRI connectivity, is not altered in sleep as compared to wakefulness (Uehara et al., 2014) and remains unchanged during processing of sensory inputs in different modalities (Moussa et al., 2011). Similarly, MEG work documented immutability of global network features during task vs. during rest (Bassett, Meyer-Lindenberg, Achard, Duke, & Bullmore, 2006). On the other hand, very simple manipulations of information content in input streams have been shown to impact a range of core topological features of functional networks, including their modularity, module number, partition structure, and node degree distribution (Andric & Hasson, 2015), and other work has shown that simple manipulations such as repeated presentation of narrative content has the *strongest* effect on perisylvian regions (Andric, Goldin-Meadow, Small, & Hasson, 2016). A productive direction for future research would be examining how connectivity features (either of a restricted set of regions, or on the whole-brain level) change with language content and language task, or even naturally with the progress of discourse (see e.g., Chai, Mattar, Blank, Fedorenko, &

---

[10] These include features such as partition modularity, partition structure, and node degree distributions and various quantifications of small world features.

Bassett, 2016). More generally, constructing research paradigms and analysis methods that question the notion of a stable functional language network can provide findings that will, at minimum, expose the relative flexibility of functional (re)organization within perisylvian regions.

## 5. Statistical models as an explanatory framework for processing costs

To this point, we have shown that different types of contextual constraints (or knowledge) share similarities with respect to the timeline by which they are integrated and the brain systems they impact. From a neurobiological perspective, treating these disparate information sources, that is contextual constraints, as a single category would be supported if it were possible to specify computational architectures that could accommodate long-term knowledge, situational constraints, and textual constraints within the same architecture. In recent years, the advancement of convolutional neural networks for the processing of language and co-occurring information suggests viable computational implementations for acquiring and representing such contextual effects. For example, from co-presented visual and language information, such systems can automatically produce image captions for novel images (Vinyals, Toshev, Bengio, & Erhan, 2015), or align book and video contents (Zhu et al., 2015). These architectures work on unlabeled data and do not require any sort of mid- or high-level labeled information. These architectures, as well as simpler ones based on recurrent neural networks, can account for the contextual effects we described using mechanisms that code for corpus-based proximity and statistical co-occurrence. In presenting this issue we draw on several advances in computational linguistics, artificial intelligence, and psycholinguistics (for related discussions, see, Andrews et al., 2014; Armeni et al., 2017). Such advances have allowed modeling contextual effects on learning and processing in ways that were not conceivable even a few years ago.

### 5.1. Distributional data, lexical predictability and semantics

Recent years have seen a resurgence of computational and experimental studies addressing how statistical models may account for online incremental language comprehension. These models show that computational mechanisms exploiting lexical, part-of-speech, or other corpus-based-distance metrics perform well in accounting for online processing during comprehension, and can account for what are traditionally considered as semantic or syntactic effects. For instance, estimating the processing cost of upcoming information is the chief goal of models based on statistical constraints between words or groups of words, such as the ones developed by Frank and Bod (2011). These architectures, comprising Markov models and recurrent neural networks (RNNs), can predict the processing cost of an upcoming word based on the previous sentence context. Their predictions of processing costs are estimated using information-theoretic measures such as entropy and surprisal, and have been found to correlate with the amplitude of the N400 as well as with reading times (Frank, 2013;Frank, Otten, Galli, & Vigliocco, 2015; Delogu, Crocker, & Drenhaus, 2017; Murphy, Wehbe, & Fyshe, 2018). Processing costs (as reflected in an ERP N400) are also accounted for by a model that scales cost with the dissimilarity between a words meaning and that of the preceding context, when those are quantified via multidimensional vector representations (Broderick et al., 2018). Neuroimaging work shows that these models' estimations of processing cost predict neural activity fluctuations in temporal regions during language comprehension (e.g., Fruchter & Marantz, 2015; Willems, Frank, Nijhof, Hagoort, & Van den Bosch, 2016). An EEG study (Frank & Willems, 2017) has further shown that lexical predictability (contextual surprise) and semantic similarity (in terms of cosine distance between the vector of the target word and the combined vector of the preceding words) elicit distinct patterns of neural activity in participants that are processing naturalistic language. Importantly, these models do not require any explicit syntactic structure or rule-based system to obtain their good performance.

The models described so focused on predicting how likely a word is in a given sentence, but not how plausible a certain meaning will be in a given context. However, the very same statistical metrics can be used to model word meanings by means of multidimensional spaces, as in distributional semantic models. At the basis of distributional semantics is the *distributional hypothesis* stating that words occurring in the same contexts tend to have similar meanings (Harris, 1954). This principle has been used to develop successful models of the semantic system, such as Latent Semantic Analysis (Landauer & Dumais, 1997) and Topic Models (Griffiths, Steyvers, & Tenenbaum, 2007), in which word meanings are represented as their occurrence vectors across large collections of text documents (i.e., corpora).

These computational models correspond with a family of psychological models of semantic memory (conceptual knowledge) that do not assume a conceptual core for concepts (words) but hold that concepts are inextricably linked to the multiple contexts in which they appear (e.g., Kemmerer, 2015; Yee & Thompson-Schill, 2016). These hold that access to conceptual knowledge during language comprehension depends on multiple factors such as comprehension context and the long term co-occurrence patterns. Kemmerer (2015) notes that contextual features might account for which features are activated, and shows this can account for findings reported in the study of motor verbs, where activation of motor features is not deterministic but depends on task and context features.

### 5.2. Integrating linguistic and non-linguistic information

This brief introduction to distribution-based models highlights the typical focus on language data that characterize these approaches. However, for the purpose of our argument, it is important to realize that the tendency to limit these types of analyses to linguistic contexts (i.e., linguistic co-occurrences) was mainly driven by availability of large text databases, rather than by theoretical constraints. In fact, any kind of data for which it is possible to draw associations between information units can be integrated into distributional representations, thus making the boundary between linguistic and extra-linguistic information less rigid than researchers usually assume. In other words, the co-occurrence of elements, irrespective of the "cognitive category" one assigns them (lexeme, phoneme, visual element), could serve as foundation for a comprehensive statistical model of language representations. Although there are clearly levels of representation where these are distinct units, for purposes of *neurobiological* computations underlying language comprehension, such multi-modal elements may be bound together.

There are several computational frameworks that share the capacity to integrate both linguistic and non-linguistic aspects of incoming stimuli, and that in this way can elegantly account for contextual effects in language. For the sake of clarity in exposing these models, we will use as an example the interesting finding reported in van Berkum et al.'s (2008) study in which a sentence such as *Every evening I drink some wine before I go to sleep* evoked a stronger EEG response when pronounced with a child's voice than an adult voice.

Given that a child's voice is unlikely to state a preference for wine, that kind of voice is an unlikely context for the "drinking wine" concept. The incongruence between these pieces of information can be captured by a distributional model based on co-occurrences between acoustic and semantic features, in which the physical features of a "young voice" (vs. any other age of voice) will not be a prominent dimension in the "drinking wine" vector. Visual-semantic distributional models (Bruni, Tran, & Baroni, 2014) construct semantic representations that combine text-based linguistic properties (as defined, e.g., by Turney & Pantel, 2010) with low-level visual features obtained from image databases and show how merging co-occurrence of linguistic and

non-linguistic features boosts modeling of language comprehension. Moreover, models combining text- and image-based information, as opposed to their textually-specific counterparts, provide the best fit to brain data (Anderson, Bruni, Lopopolo, Poesio, & Baroni, 2015) and to behavioral performance in novel-word learning (Lazaridou, Marelli, & Baroni, 2017). These architectures can potentially integrate other kinds of information, including conversational elements, speaker properties, and world-knowledge aspects. Indeed, distributional models based on acoustic information has already been proposed (Lopopolo & van Miltenburg, 2015), as well as approaches combining textual an experiential statistical data in the same distributional architecture (Andrews, Vigliocco, & Vinson, 2009). The main challenge is hence a practical one, as training data would need to capture co-occurrences between and within different levels of description (linguistic, visual, discourse, contextual, speaker-associated, etc.) in complex scenarios.

### 5.3. A learning-centered perspective

Co-occurrence-based systems capture not only the end-state of learning but also the learning process by which the distributional structure is established. Indeed, the way co-occurrences are re-weighted in distributional semantic models is quite close to the outcome of a Rescorla-Wagner learning procedure (Rescorla & Wagner, 1972), even more so in recent neural-network-based systems (see Mandera, Keuleers, & Brysbaert, 2017). In the Rescorla-Wagner equations, context plays a crucial role in determining the final link weights between the model units. During learning, a cue (e.g., a phoneme) appearing along with an outcome (e.g., a lexeme) will lead to a stronger link between the two corresponding representations. However, when many cues appear together, they share the entire amount of discriminative power: the link between each cue and the outcome will still be strengthened, but to a lesser degree than if the outcome were associated with a single cue. Consequently, the final state of the model is not only influenced by the cue-outcome associations, but also by the cue co-occurrence patterns themselves (for a description in probabilistic terms, see Danks, 2003). In this way, the Rescorla-Wagner equations account for the influence of complex contexts in a general learning framework.

The Rescorla-Wagner equations have been applied to language, within the general approach of Nave Discriminative Learning (NDL, Baayen, Milin, Durđević, Hendrix, & Marelli, 2011). In NDL, word meanings are considered outcomes, whereas simpler aspects/features (e.g., graphemes representing meaning-devoid language units, raw phonetic features, visual elements included in a learning scenario, etc.) are considered cues. The 2-layer NDL model is based only on a feature-based input layer (typically coding for sublexical features such as phonemic or orthographic features) and a target semantic layer, and is trained without feedback. The model provides a unified account for a wide range of phenomena in online word recognition (Baayen et al., 2011). NDL also permits modeling predictive processes in language comprehension, with NDL weights determining how a given word restricts the uncertainty of upcoming information. Indeed, model simulations in this sense can account for results of priming effects, ERPs, and fixation times in reading (Hendrix, Nick, & Baayen, 2014). Multiple-layer NDL networks have been shown to simulate semantic effects in priming studies (Milin, Feldman, Ramscar, Hendrix, & Baayen, 2017). These cases provide a clear demonstration of how a powerful and rudimentary learning mechanism can explain phenomena that have been typically addressed by linguistic (language-essential) formalisms.

NDL naturally accommodates the impact of non-linguistic features on comprehension. Considering the example of the wine-enthusiastic child, from an NDL perspective, vocal features can be represented as (negative or positive) cues for concepts or semantic features. A child's voice is likely a positive cue for the concept "drinking milk". In contrast, it is likely a negative cue for the concept of "drinking wine", leading to inhibition of the corresponding representation units, and hence increased cognitive cost when said unit are de facto activated by

the stimulus content. This result, captured in the adult state of the system, would emerge from the vocal/semantic co-occurrences across repeated learning events. Computational developments within the NDL framework show that meaning access can indeed be informed by simple acoustic features, providing an architecture in which the association between sound and meaning is learned through statistics from natural conversation data (Arnold, Tomaschek, Sering, Lopez, & Baayen, 2017). In conclusion, NDL and similar approaches position language comprehension within a very general learning perspective, in which the focus is on the elements of the context, with no need for distinction between linguistic and non-linguistic aspects.

### 5.4. From words to sentences

Clearly, simple models of this sort cannot capture structural dependencies, which has been a longstanding argument against their viability as models of linguistic competence (Chomsky, 1959). Still, the importance of findings like the ones reviewed here is in showing just how much can be accounted for by such (non-hierarchical-based) systems. In fact, Christiansen and Chater (2015a) have shown how a simple recurrent network can produce signatures matching sensitivity to recursion in absence of any grammar model. Whether linear approaches offer a sufficient account (for most of naturally occurring sentences) or are hierarchical models necessary is an ongoing debate, which is addressed, for example, in Frank and Bod's (2011) work comparing the latent knowledge coded in recurrent neural-network vs. phrase-structure systems.

As discussed above, feature co-occurrence, at different levels, offers a powerful account for online processing. The models mentioned here and similar ones are typically trained on a large text corpus to derive a representation for word meanings or lexical items (either as a set of values in multidimensional space, or as a set of links to input features), and it is this representation that is used to model behavioral data. Beyond that, the last several years have introduced breakthroughs in understanding how such distributional semantic representations can be compositionally combined, thus affording an even better account for how the large scale statistical knowledge acquired over the life span can be further used for comprehension in a specific context. Mitchell and Lapata (2010) have proposed algorithms that can generate vector representations for high-order word combinations (phrases and sentences) as function of the features of their constituents. Extensions of these models (Baroni & Zamparelli, 2010) can account for the fact that corpus unattested combinations such as *legendary province* are judged as sensible, whereas equally unattested combinations such as *empty fungus* are perceived as non-sensible. Even when applied at the morphological level, such models account for core behavioral effects that are considered gold standards for semantic theories. Marelli and Baroni (2015) have shown that such a model can predict that nonce forms such as *re-browse*, *re-provoke* and *re-matter* have graded acceptability and also account for semantic transparency effects. Marelli, Gagné, and Spalding (2017) have further shown that compositional distributional models can account for chronometric data, such as response times in priming paradigm and word comprehension, in the processing of novel noun combination (e.g., *snow shovel* or *mountain magazine*). In these tasks, the model simulations are able to replicate the speaker sensitivity to the (unexpressed) semantic relations binding the constituent nouns (e.g., *shovel_FOR_snow* or *magazine_ABOUT_mountain*).

Crucially, distributional models based on input statistics can explain and predict whether a new expression would be judged as sensible, which has been the main challenge posed to these approaches (Chomsky, 1959). This has traditionally been argued to be a core weakness of co-occurrence based systems (albeit the argument was made against the simpler models that existed in the mid-twentieth century). In sum, vector-based models capture essential properties of terms, which can be further related to how sensible they are when combined. Predictions of these models have been examined empirically

(Marelli & Baroni, 2015; Vecchi, Marelli, Zamparelli, & Baroni, 2017), showing that quantitative properties of the resulting vectors (e.g., their entropy, the density of their semantic neighborhoods) correlate with human intuitions concerning the plausibility of the associated word combinations. We emphasize that these are aspects of sensibility that are *not* accounted for by semantic models based on simple co-occurrence frequency or information-theoretic metrics such as Mutual Information.[11]

### 5.5. Statistical models as an interpretative framework for language comprehension

The models we have reviewed offer a parsimonious yet powerful and general explanation for the wine-drinking child result (as a reminder: a child's voice uttering a sentence about drinking wine elicits a greater N400 effect than an adult's voice uttering the same sentence van Berkum et al., 2008). In explaining their findings, van Berkum et al. (2008) persuasively argued against the position that message level content is initially computed independently from vocal features (speaker identity and emotional tone), and only then evaluated against those features. Instead, they suggested that social dimensions are appreciated so quickly and that "Language users very rapidly model the speaker to help determine what is being said" (p. 586). Statistical models of the sort we described above offer a more general explanation for this finding, as large-scale statistical knowledge, based on co-occurrence between different types of features (linguistic and non-linguistic), can in principle account for this effect. At the same time, these models offer an explanation that is based on a general computational framework, does not assume that modeling the speaker's features (e.g., age, gender, social class) is obligatory and does not necessitate a dedicated linguistic apparatus (see Frank and Bod, 2011, for a detailed discussion).[12]

In conclusion, cognitive and brain scientists now have a collection of computational approaches for modeling non-linguistic influences on language comprehension. We have focused the discussion on distributional semantics, naive discriminative learning, Markov models, and recurrent neural networks. The list is far from exhaustive. Rather, it is meant to highlight a few approaches that share desirable properties. First, the context in which information is acquired plays a central role in all these models. Second, they are not systems specifically designed for language, but can be adapted to other kinds of information sources. Third, they do not assume ad hoc structures or precompiled categories, but rely on statistical patterns in the input data. Fourth, they are not aimed at establishing the grammaticality of a given combination, but rather measure to what extent that combination appears plausible or surprising to the speaker and can account for the relative sensibility of completely new content. To this end, model performances are assessed on the psycholinguistic ground, with human-generated data (explicit judgments, behavioral responses, brain activation) taken as gold standards. All these aspects are strictly associated to each other, and provide a framework that is consistent with the programmatic points of the present paper.

## 6. Oscillations as a neurobiological framework for brain activity and connectivity

Though the rhythmic nature of human electro-encephalogram (EEG) data is perhaps their defining characteristics, the possibility that oscillations are associated with specific high-level language operations is a much more recent development (see Meyer, 2017 for review). Notwithstanding, a substantial number of studies already suggest links between theoretically-specified computations and EEG oscillatory features such as power, coherence, and cross-frequency coupling. Approaching language comprehension from the perspective of neural oscillations is theoretically advantageous: it allows grounding temporal and spatial patterns of brain activity during language comprehension in general computations, and also naturally emphasizes consideration of large-scale networks rather than small isolated regions.

Decades of research have identified general processes linked to major EEG frequency bands, including gamma, alpha, theta, and delta. This vast knowledge both constrains interpretation and provides a natural theoretical interface between the computations carried out during language to more general ones. This produces a better grounding of novel results, and opens new directions for theoretical development. Because our purpose here is to speak to how this development bridges between the interpretive principles we have presented and neurobiology, rather than provide a general review, we refer the reader to Meyer's (2017) extensive review, which outlines the following possible links (the following is necessarily a coarse summary). Delta-band activity is related to chunking words into syntactic phrases, possibly related to basic chunking mechanisms evident outside language. Alpha is linked to verbal memory demands and storage of phrases in verbal working memory, and its well-studied link to inhibition could indicate that processing syntactic dependencies engages more general mechanisms. Theta relates to memory retrieval during language, potentially in virtue of its more general role in retrieving elements from sequences. Finally, Beta and Gamma are interpreted in relation to predictions and their evaluation, with Beta linked to forward looking operations (or maintenance of status quo; Lewis & Bastiaansen, 2015) and gamma to backward-oriented integration processes, with higher gamma indicating more successful integration (see also Bastiaansen & Hagoort, 2015). While these are only provided as examples, such high-level mapping statements can ground interpretations of brain activity during language comprehension in basic principles.

Focusing on oscillations allows examining issues related to coupled activity across brain regions and thus naturally lends itself to a network-based rather than region-based perspective. Because the same brain areas may be involved in networks with different frequency characteristics this forms a logical model for linking single regions to multiple, parallel computations. Finally, different sorts of stimuli or tasks can interact with ongoing oscillatory patterns in complex ways, for example, either by impacting the power of ongoing oscillations, or by inducing phrase reset of the sort captured by averaged ERP components. This provides an insight into how language comprehension interacts with background oscillatory activity, and how phase-reset or modulated changes are linked to successful comprehension.

---

[11] Several features of these models make them psychologically appealing. First, a word's meaning is obtained from its distribution in a corpus, but is not captured by a fuzzy distribution-like representation (e.g., by a Dirichlet distribution with $n$ hyper-parameters). Instead, each term is represented via a uni-dimensional vector, with no uncertainty attached to that specific representation. These models also offer much flexibility for accounting for how a sense of a word might change in context, as compositionality is reflected by vector addition or multiplication (this applies both at the sub-lexical and lexical level). This obviates the need to assume a singular meaning for each word that is stable across usages (see Lebois, Wilson-Mendenhall, & Barsalou, 2015, and references therein for discussion), and is consistent with Cruse's (1986, p. 50) observation that the meaning of any word form is in some sense different in every context in which it occurs.

[12] We note that the models we have presented are well suited to explain match/mismatch effects between any two communicative cues, computations that have also been linked to activity in lateral temporal regions. For example, activity in these regions is greater when a tone of voice does not match a speaker's facial expression. However, the mechanism behind these incongruence effects is unclear, and has been suggested to be related to mismatch detection or implementation of executive control (Watson et al., 2013). On a statistical approach, two rarely co-occurring stimuli (e.g., happy voice + sad face) can be associated with increased processing due to either (1) increased competition, as each cue activates alternative representations that typically co-occur with it i.e., a happy face or sad voice respectively, or, (2) increased processing due to the fact that the mutual inhibition between the cue and target results in a longer time necessary to instantiate target representation.

Oscillatory activity can provide information regarding integration on multiple timescales relevant to language comprehension, from few milliseconds (gamma) to much slower processes. A straightforward example is the finding (Bastiaansen, Magyari, & Hagoort, 2010) that reading syntactically correct sentences, but not random word sequences, produces a linear increase in beta power, which is disrupted when encountering a minor syntactic violation. Given that these patterns developed gradually over a relatively long interval (here, 5 s), the authors suggested they might be related to syntactic unification operations. In addition, reading syntactically structured sentences, either with or without violations, produced a gradual increase in theta power. The linear trend found for theta, jointly with its insensitivity to syntactic violation is compatible with its relation to construction of working-memory traces for language.

Using MEG, it is possible to localize brain areas where oscillatory changes are induced. This allows linking frequency bands to their spatial sources. For instance, Kielar, Panamsky, Links, and Meltzer (2015) found that both semantic and syntactic violations induce weaker activity in the alpha and Beta (8–30 Hz) bands for a duration of around 1 s. A striking result is that the spatial origins of these effects was widespread, including left perisylvian regions but also large areas in occipital cortex and across the superior frontal gyrus. In addition, a left-lateralized N400 response was time locked to the violations (i.e., linked to a phase reset of oscillatory activity). As discussed by the authors drawing on diverse literature, the occipital cortex effects may suggest natural feedback interactions, whereby alpha and beta power regulates information flow in the brain by deactivation of task irrelevant regions. Thus, changes in these oscillations, and the localization of these changes may point to coordination of activity in functionally coordinated areas.

Electrocorticography data can be merged to contribute to understanding these questions. Recently, Nelson et al. (2017) obtained data from intracranial recordings of sentence comprehension. They documented a gradual increase in high-gamma power during single word reading of sentences, which decreased at points where words could be merged into a constituent phrase. While word-yoked gamma power was explained using computational models that estimated forward-entropy and word-surprisal, (similar to Willems et al., 2016), grammatically motivated predictions derived from a phrase-structure parser provided the best fit to gamma fluctuations. Addressing a similar question, Ding, Melloni, Zhang, Tian, and Poeppel (2016) showed that it was possible to identify ECoG signal in the high gamma band that tracked sentential, phrasal, or syllabic level rate fluctuations. The mechanisms producing these signatures are under debate, with some (Ding et al., 2016) suggesting they are a product of processes sensitive to hierarchical syntactic structure, but other work suggesting that such neuroelectric patterns can be derived from purely linear statistical co-occurrence history (Frank & Yang, 2018). In related work, Rommers, Dickson, Norton, Wlotko, and Federmeier (2017) evaluated EEG oscillatory activity during comprehension of sentences that strongly or weakly constrained the terminal word in the sentence and found that prior to onset of the critical word, strongly constraining contexts were associated with greater alpha decrease in the alpha band.

A network/oscillatory perspective raises multiple novel questions. The potential link between Beta and Gamma naturally leads to questions regarding the temporal relationships between regions that code for predictions and those that evaluate them, and well-studied memory principles related to gamma/theta precession provide an analytic approach for investigating correlates of successful encoding of verbal content to memory and later access to those memory traces. It also suggests that the medial temporal lobe plays a more central role in online comprehension than typically considered. The fact that the degree to which prior context constrains future complements appears to mediate both Gamma (Nelson et al., 2017) and alpha (Rommers et al., 2017) activity suggests their interplay is crucial, and that coupled changes in alpha and gamma power may be a signature of efficient

forward-looking predictions. In fact, this very issue was recently addressed (Wang, Hagoort, & Jensen, 2017), in a study whose departure point was that the relationship between different frequency bands active in different brain regions may act as a way of neural information exchange between brain areas. Wang et al. examined this sort of 'cross-frequency coupling' in the domain of sentence comprehension using sentences in which a critical word was more or less predictable given the earlier part of the sentence. They observed greater alpha power suppression for more predictable words in a set of areas typically though to be involved in language comprehension. Most interestingly, they found that the alpha power suppression in temporal areas was negatively correlated with gamma band in left inferior frontal cortex. The authors suggest that there is anticipatory influence from left inferior frontal cortex onto temporal regions, which is instantiated in a cross-frequency coupling between frequencies that these brain regions operate on when performing language comprehension. While future work may refine these conclusions, a major importance of the study is showing how grounding explanations in oscillatory principles contributes to making specific predictions about interactions between different brain regions.

Schoffelen et al. (2017) extended this approach by showing that it is possible to determine the direction of information flow across brain regions during language comprehension. Participants read sentences or scrambled word lists and Granger causality (a measure of causal influence here applied for different frequency bands) was used to compute interactions between a large set of regions linked to language processing. It was found that left inferior frontal and left anterior temporal cortices mainly received input from other areas, whereas middle temporal cortex mainly sent out information to other regions. They observed efferent information from temporal or parietal areas to peak around 12 Hz (beta frequency range), whereas directed influences from frontal to temporal areas was most prominent at a higher frequency (~27 Hz). What is interesting about this approach is that it suggests the existence of frequency-specific subnetworks, which can operate dynamically to instantiate language comprehension.

With advances in multimodal brain imaging, joint studies of oscillatory activity and BOLD signals can be carried out in parallel, which can offer significant insights into the computations underlying language. Such investigations can code for points of cross-frequency coupling, oscillatory changes and coherence across brain regions in the EEG band, and identify brain areas whose activity tracks these changes during language comprehension.

Considering language from this perspective is consistent with the position developed in prior sections which argued against assuming that a given network subserves language functions, and in favor of considering that specific context-dependent computations may shift network configuration. While extant data do not allow strong conclusions about which networks may be identified in such analyses, it is already possible to think of how to go about constructing such descriptions, as these will have different structures than current ones. Rather than linking a function to a region or to a set of regions, these descriptions will link sets of functions to a region (or to sets of regions) in a quantitative and continuous way. In a conceptual review of functional networks, Pessoa (2014) outlined the foundations for such a vocabulary. These include *functional fingerprints* that capture the many-to-many relation between regions/networks and functions, and *diversity maps* that quantify the relative heterogeneity of functions that any putative functional network is involved in. Clearly this is complicated endeavor, but the very utility of such descriptions and the information they capture belies the usefulness of assigning cognitive functions to single brain regions (see also Anderson, 2010 for a similar argument). While some anatomical structures will undoubtedly be linked to certain computations (e.g., lower level sensory regions), and some networks could be strongly constrained by anatomical structure (e.g., resting state connectivity of cingulate gyrus follows cytoarchitectonic boundaries, Beckmann, Johansen-Berg, & Rushworth, 2009), context

dependent connectivity will nonetheless be a dominant factor (see Spunt & Adolphs, 2017 for a related discussion).

## 7. Where from here?

As evident from our evaluation to this point, understanding how the brain implements language comprehension in naturalistic everyday circumstances is not a simple matter. Moreover, answering this question in a non-myopic or language-centric manner is crucial not only to cognitive neuroscientists that are interested in the brain *per se*, but to diverse disciplines including clinical psychology, neuropsychology, computational modeling of language, and of course experimental psychology for which brain data often provide useful constraints in theory construction.

Our aim here was not to sketch a neurobiological model of language comprehension. If anything, our review shows that there is insufficient data to address this issue. Instead, we aimed to present a constructive conceptual framework or research approach toward delineating such models. Our core arguments address the following issues: (1) the difficulty in extending language-centric explanations to discourse, as such explanations ignore the potential role of basic information integration, segmentation and predictive functions and, for various reasons, assume separation of language from other, related functions; (2) the necessity of taking context as a serious topic of study, modeling it formally, and acknowledging the limitations on external validity when studying language comprehension outside context; (3) the fact that neural computations at any given point are strongly related to the state of brain function prior to that point; and (4) the status of the language network as an explanatory construct in future work on NLC. In this final section we identify the challenges introduced by this approach, and then describe which subfields of science will benefit most from it.

### 7.1. Challenges

First, one unpalatable point is that conclusions derived from prior work within the neurobiology of linguistics (neurolinguistics) or the neurobiology of psycholinguistics may not hold at the level of NLC. For instance, the assumption that measures of syntactic complexity can explain brain activity patterns during natural comprehension hinges on the (often implicit) proviso that there are brain regions whose activity reflects a pure effect of syntactic complexity, immune to contextual interaction. We have reviewed current work in neurobiology and prior work in psycholinguistics that shows how tenuous this assumption is. Compounding this difficulty, the fact that much of the neurobiological knowledge to date has been obtained from studies devoid of communicative intent may mean that at this point, the field holds a very detailed and comprehensive knowledge of a highly particular and non-default mode of comprehension, which may be associated with a unique pattern of brain activity. The chiasm between comprehension of simple, experimentally-derived texts and natural texts/discourse has been a lingering concern (see Graesser, Millis, & Zwaan, 1997 who criticized the use of experimenter-constructed "textoids"). The first challenge is abandoning a restrictive view of studying language with overly artificial or simplified texts and devoid of context.

A second, related challenge is that, even in domains where current knowledge can arguably be extended, via induction, to naturalistic comprehension, it may be *insufficient* for this purpose: NLC cannot be assumed to engage the same brain areas, reflect similar network interactions, or operate over similar time scales as those found for single-sentence processing. Consequently, prior work within the neurobiology of language, while making important contributions for lexical and single-sentence processing, cannot be expected to provide a complete account of NLC, or even account for how single words or sentence are understood in ongoing context. Instead, the impact of concurrent activity in other brain networks, the interaction between networks implicated in semantic processing and other networks, the impact of

contextual information (both textual and para-textual), and the effect of the recent neural context, will jointly determine the process and outcome of comprehension. This has significant implications for the *practice* of science, because training within psycholinguistics or neurobiology of language as currently defined may be insufficient for tackling these questions. Instead, it will be important to rely more extensively on collaborations with other domains in the cognitive neurosciences. In particular, we foresee that insights from computational linguistics (Frank, 2013; Frank and Bod, 2011; Wehbe et al., 2014; Willems et al., 2016), the interface of mood and cognition (Egidi & Caramazza, 2014), and from research in network theory will be essential for characterizing NLC. What this means (and this elephant in the room should be acknowledged) is that the neurobiology of NLC, while sharing the word "language" with neurolinguistics will constitute a substantially different research topic. NLC will be less committed to the goal of mapping of activity patterns to well defined, pre-established cognitive categories, and instead will formulate basic (neurobiological) principles that organize activity during comprehension.

A technical challenge will be devising experimental methods that allow researchers to study how natural language is understood in the brain (Andric & Small, 2015). One option is to adapt the standard modeling approaches in which certain events are treated as conditions, while participants listen to continuous speech (Hsu, Jacobs, Citron, & Conrad, 2015; Nijhof & Willems, 2015; Yarkoni, Speer, Balota, McAvoy, & Zacks, 2008). Another set of tools constitutes model-free techniques such as independent component analysis, graph-construction and classification methods (Takerkart, Auzias, Thirion, & Ralaivola, 2014), intra-participant correlations (Levin & Uftring, 2001) or inter-participant correlations (Hasson et al., 2008). These methods can document, with minimum functional assumptions, which brain regions are associated with language processing. For example, Honey, Thompson, Lerner, and Hasson (2012) used inter-participant correlation to identify brain areas that response similarity to discourse segments presented in different languages thus identifying processing at a certain level of abstraction. However, such paradigms are accompanied by interpretive difficulties, as brain activity may track discourse for several reasons, including not only linguistic computations, but due to content-related fluctuations in attention, sensitivity to lexical valence, lexical or word-cluster (N-gram) frequency, changes to breathing or cardiac activity, and several other factors. Nonetheless, such methods are useful for generating more constrained hypotheses about functional networks and their potential information processing roles. These hypotheses in turn can be tested with creative adaptations of existing methods, which are more suited towards hypothesis-testing experiments. The study of neuroelectric oscillatory activity in these paradigms may turn out to provide further constraints on data interpretation. Though at present, our understanding of various oscillatory dynamics might preclude relying on them as explanatory foundations for naturalistic language comprehension, that domain of study is rapidly developing, and it would be important for language research to interface with it both in terms of developing integrative theories and adoption of suitable methods.

One method that is already proving very useful is modeling brain activity during language comprehension using distributional models, or dependency structures (e.g, Alday, Schlesewsky, & Bornkessel-Schlesewsky, 2017; Cibelli, Leonard, Johnson, & Chang, 2015; Ettinger, Linzen, & Marantz, 2014; Frank et al., 2015; Frucheter et al., 2015; Leonard, Bouchard, Tang, & Chang, 2015; Nelson et al., 2017; Wehbe et al., 2014; Willems et al., 2016). Such studies allow modeling EEG, MEG, electrical recording and fMRI data using quantities derived from corpora, which include transition probabilities, base-rate frequencies, or forward entropy measures. This same approach has been used in several behavioral studies as well, notably where eye tracking data during reading were modeled using constructs derived from different linguistic theories (e.g., Demberg & Keller, 2008; Frank & Bod, 2011).

## 7.2. How neurobiology impacts functional theories and produces new hypothesis

Cognitive scientists interested in language belong to either of two groups: (i) those who believe that behavioral studies and sturdy logic are sufficient for obtaining an understanding of language function, or (ii) those who think that neurobiological data can additionally inform this issue. The first group may find our preoccupation with valid neurobiological models of language to be of little relevance. However, most practicing cognitive scientists likely fall in the second category (as evident, for example, in the extensive report of brain data in the journal Trends in Cognitive Sciences in the recent 5 years or more recently in Nature Human Behavior). And members of this group, by definition, have some idea of what data are consistent or inconsistent with different cognitive accounts (see Chatham & Badre, 2014, for discussion of these two positions). The data we have reviewed indicate that this group is already faced with findings that urge shifts in functional models of language.

First, as reviewed, recent neurobiological data do not support the often-espoused distinctions between prior linguistic/discourse context (co-text), situational context (e.g., speaker features), long-term knowledge, and intrinsic beliefs/mood as constraints on language comprehension. Instead, it appears that all these types of knowledge or constraints impact online processing within a similar time frame and in the same brain regions. This suggests there is no initial "context independent" reading, nor a priority for co-text constraints, and that proper modeling of language-in-context needs to take such factors into account.

Second, emerging and influential models of language attribute a major role to the capacity to acquire distributional linguistic information (e.g., transition probability, entropy of potential word-completions or sentence completions, Frank & Bod, 2011), and treat this as a fundamental linguistic competence (Baroni & Zamparelli, 2010). In support of this view, an emerging body of neurobiological work suggests that such statistical features of the language input are neurally represented, and rapidly impact comprehension (e.g., Cibelli et al., 2015; Ettinger et al., 2014; Frank & Bod, 2011; Fruchter & Marantz, 2015; Leonard et al., 2015; Lopopolo et al., 2017; Willems et al., 2016).

Third, the fact that so-called "language regions" are implicated in distributional learning, segmentation and prediction for non-linguistic stimuli (as reviewed in Section 2) suggests that the statistics of language are acquired via inherently non-linguistic functions. Arguing from first principles, it makes little sense to just assume that there exist distributional-learning or predictive processes unique to language. Instead, it is important to actively test, empirically, whether there are simpler functions that language may build on, and think of how to ask neurobiological questions about those. It is therefore the workings of these more basic functions that should move to the forefront of behavioral and neurobiological investigation as this may shed light on why language is processed the way it does. Unexpected insights can develop from such work: for instance, current models of statistical learning of language do not separate sensitivity to marginal frequencies (base rate) from sensitivity to transition probabilities. Brain data, however, suggest that these features are tracked by different systems (e.g., Tobia, Iacovella & Hasson, 2012). Neurobiological research that contrasts information processing in linguistic and non-linguistic domains within the same study can provide useful insights into these issues (e.g., Christiansen, Conway, & Onnis, 2012). Our conclusion on this issue is consistent with an increased emphasis in cognitive science on explaining language – both acquisition and online parsing – via much more basic processes that are based on chunking, local-replay, prediction, or abstraction (compression, gist extraction) as opposed to assuming sophisticated higher-level linguistic functions

(Christiansen & Chater, 2015a, 2015b).[13] These developments align with our argument for studying basic neurobiological computations that support both language acquisition and other forms of online parsing.

Finally, neurobiological studies suggest a rethinking of the nature of semantic operations. As we have reviewed, cognitive operations considered at the core of the psychology of semantics (e.g., integration of input against schematic knowledge, Bransford & Johnson, 1972) are subserved by brain systems that are not typically considered as a core part of a language processing system (narrowly defined; Bein, Reggev, & Maril, 2014; van Kesteren, Fernández, Norris, & Hermans, 2010). Furthermore, integration of incoming information locally against the immediate preceding text vs. globally against more remote text engages different brain systems, rather than engaging a single system in different ways (Egidi & Caramazza, 2013). These latter findings argue against the putative equivalence of semantic operations at the sentence and discourse level, and cannot be easily brushed aside as being "consistent with" or "complementary to" theories that hold on to the notion of an encapsulated language system. This is because such neurobiological findings are in fact tearing away at the set of functions purportedly assigned to the classic language areas (or language module/system). In addition, such findings could have clinical implications. For instance, deficits in paragraph-level comprehension or schema integration, as opposed to sentence-level integration, may not be related to linguistic operations or work of 'language' systems and have little in common with deficits linked to single-sentence comprehension.

In this way, neurobiological data are already suggesting important revisions to how we should think of language. This means acknowledging the strong similarity between processes previously thought as separate (e.g., the impact of different contexts), while at the same time breaking up what were thought to be single constructs (e.g., semantic integration).

Modifications to functional models of language, which are informed by neurobiological findings, will in turn influence the sorts of hypotheses made about brain systems. As mentioned in the Introduction, influential approaches to language still aim to link brain regions to particular language functions. The neurolinguistic approach has been described as follows (Grodzinsky & Friederici, 2006): *each subpart of the linguistic system – whether it is phonology, syntax or semantics – can be neurologically decomposed into subcomponents*, and Friederici (2011) notes that a neurolinguistic model, …*relates a particular function to a particular brain region within the* <u>language system</u>, *leaving the option open that this same brain region serves another function in another domain than language* (underlines mark our emphasis). In this sense, neurolinguistics encapsulates positions that see an inherent value in the ontology suggested by (typically generative) linguistics for purposes of understanding how the brain processes language, and assume it should be possible to identify the "precise anatomical locations of higher level syntactic computations" (see recent review by Skeide & Friederici, 2017). This work continues, for example, by attempting to identify brain areas that mediate a syntactic merge operation (Zaccarella, Meyer, Makuuchi, & Friederici, 2017). There is little concern with discourse level processes. The idea of a language/linguistic system is treated as a core organizational principle, which guides the construction of a functional ontology and, consequently, a search for its neurobiological realization. These ideas permeate even recent studies, as seen in a strong focus on perisylvian regions for modeling language.

---

[13] Christiansen and Chater (2015a, p. 41) have argued that even the assumption that the operation of recursion is strictly linguistic (i.e., "not emerging from general-purpose cognitive mechanisms or constraints") is one that should be abandoned. Instead, they take the position that *language processing, implemented by domain-general mechanism – not recursive grammars – is what endows language with its hallmark productivity*.

Several examples make the point. For instance, a recent study of language function recovery after stroke (Del Gaizo et al., 2017) used Connectome-Dynamics Lesion Symptom Mapping to identify changes to network structure that are caused by lesions and their correlation with language function. Probabilistic tracking was applied at the whole-brain level, but the connections evaluated were only those between "eight language-network nodes". In a study of long-term plasticity of language organization in the blind (Lane et al., 2017), the regions chosen for investigation were five frontotemporal language areas consisting of pars orbitalis and triangularis and anterior, middle and posterior lateral temporal cortex. And from a theoretical perspective, a recent review (Hagoort, 2017) still shows the field's focus on temporal and inferior frontal regions.

As we have argued throughout, there are good reasons to critically assess whether presenting the research question this way is a viable model for understanding how the brain organizes naturalistic language comprehension. A serious challenge is that cognitive models of language processing would need to be inspired by neurobiological data rather than attempting to fit the data within current linguistics-influenced frameworks. As we showed in Section 2, research is showing that our brains don't divide their processing repertoire according to these pre-defined partitions. For example, memory encoding of linguistic situation models may not be a result of an 'encoding' operation but *just be* a reflection of the activation state at the end of comprehension, which can be read off from lateral temporal regions and reflects the strength of a memory trace (consistent with several functional theories, e.g., Tzeng et al., 2005). A separate challenge to neurobiological theories is to make contact with multiple relevant linguistic theories and relevant pragmatic theories. These include increased focus on probabilistic information in Artificial Intelligence NLP models, and usage based accounts of language acquisition and use.

Feedback from brain studies to cognitive theory is already taking place in other domains of language research. To illustrate, in a highly detailed analysis (Pickering & Garrod, 2013), neurobiological data have been used to support a model in which there is no clear dichotomy between the functional machinery used for speech production and comprehension. Drawing on examples from sublexical and single word processing, Price (2018) argues that neuroimaging data can hint to processes not previously considered in cognitive models, and "redefine old cognitive functions". It is argued that such re-definitions can provide more general explanations for involvement of a region in a cognitive task, and Price argues that observing common components for different cognitive tasks will provide unique insights into the biological organization of cognitive function. An openness to revisions of functional ontologies on the basis of neurobiological data could lead to more precise cognitive models and a better understanding of brain data. Of course, distinctions such as those between phonetic, semantic and syntactic processing will be accompanied by activity in different brain systems (see Vigneau et al., 2006; Vigneau et al., 2011 for meta-analyses) especially if those are studied independently rather than synergistically. However, the fact that similar statistical computations can explain information-integration at the phonetic, morphemic and lexical level (mainly in lateral temporal cortex and inferior frontal gyri, e.g., Cibelli et al., 2015; Ettinger et al., 2014; Frucheter et al., 2015; Leonard et al., 2015; Willems et al., 2016), points to the utility in studying common computations across these levels. In the future, it will also be important to focus more strongly on the role of subcortical regions including the basal ganglia, hippocampus and cerebellum in these processes, as there is a large body of work implicating these regions in related operations (i.e., chunking, prediction, evaluation of prediction; see Hasson & Tremblay, 2016, for review).

While our position has a neurobiological focus, it is not a call for giving prominence to neuroscience findings at the cost of insights obtained from cognitive theory. Some scholars argue for a marked distinction and a specific relationship between neural and cognitive theories, for example, by proposing that "cognitive theory guides neuroscience" (Frank & Badre, 2015). On such an approach (e.g., Chatham & Badre, 2014) behavioral researchers should embellish their theories with testable neuroscientific predictions/correlates. These in turn can be supported or weakened by neuroscientific data. However, on this view, neurobiological data have little role in generating full-fledged cognitive approaches, and are insufficient for revising cognitive ontologies. While we agree with many elements in this perspective, we argue it cannot reflect the entirety of the relationship between cognitive theory and neuroscience data (see also Price, 2018, for a similar argument). There are multiple ways in which neurobiological data can impact functional accounts. These include providing general, qualitative support for one functional account or another, formally adjudicating between cognitive accounts (see, e.g., Kragel, Morton, & Polyn, 2015), suggesting new functions and redefining older functions (see Price, 2018). In fact, as our review shows, a synthesis of neurobiological findings from language research already suggests revisions to how we conceive of language functions (e.g., semantic operations at the discourse level, encoding, and recall) and the ways these functions are inter-related. It will be impossible to understand how the brain supports language comprehension if we were to only study the responses of a "language network" to linguistic stimuli. Clearly, one of our premises is that brain data can generate novel cognitive insights in an emergent, bottom-up manner, rather than just be used to test extant cognitive notions or help adjudicate between cognitive theories (see Spunt & Adolphs, 2017, for a similar argument in the study of domain specificity). Recent advances in convolutional neural networks offer an analogue for these impacts: in some domains, these networks are so strikingly effective in predicting human behavior (e.g., work by Kümmerer, Wallis, & Bethge, 2016 on modeling eye fixation patterns) that scientists are studying these networks' representations (in what is called "explainable AI") to understand what coding principles may underlie this performance. All this shows how progress on the neuro/cognitive fronts can be made in parallel (see Embick & Poeppel, 2015 for a similar conclusion but from a different argument). For this reason, it strikes us as uncontroversial that neural and cognitive insights for research should go hand in hand. This does not entail dominance of one over the other, but the openness to have insights from the one field of enquiry to constrain the other (see Fitch, 2014, for a similar development).

The study of NLC opens an interesting age of scientific research and opportunity. The challenges we have outlined are substantial but surmountable and have the potential for producing massively positive developments in neurobiology, cognitive psychology and cognitive science. We expect that such developments will catalyze dialogues between linguists, psycholinguists and cognitive neuroscientists and produce new research directions at the nexus of those fields.

## Acknowledgements

## References

Alday, P. M., Schlesewsky, M., & Bornkessel-Schlesewsky, I. (2017). Electrophysiology reveals the neural dynamics of naturalistic auditory language processing: Event-related potentials reflect continuous model updates. eneuro, doi:https://doi.org/10.1523/eneuro.0311-16.2017.

Altmann, G. T. M. (2013). Anticipating the garden path: The horse raced past the barn ate the cake. In M. Sanz, I. Laka, & M. Tanenhaus (Eds.). *Language down the garden path: The cognitive and biological basis for linguistic structure* (pp. 111–130). Oxford, UK: Oxford University Press.

Altmann, G., & Steedman, M. (1988). Interaction with context during human sentence processing. *Cognition, 30*(3), 191–238.

Ames, D. L., Honey, C. J., Chow, M., Todorov, A., & Hasson, U. (2014). Contextual alignment of cognitive and neural dynamics. *Journal of Cognitive Neurosciences,* 1–10. https://doi.org/10.1162/jocn_a_00728.

Andric, M., & Small, S. L. (2015). *fMRI methods for studying the neurobiology of language under naturalistic conditions.* Cambridge University Press8.

Anderson, M. L. (2010). Neural reuse: A fundamental organizational principle of the brain. *Behavioral and brain Sciences, 33*(4), 245–266.

Anderson, A. J., Bruni, E., Lopopolo, A., Poesio, M., & Baroni, M. (2015). Reading visually embodied meaning from the brain: Visually grounded computational models decode visual-object mental imagery induced by written text. *Neuroimage, 120*, 309–322. https://doi.org/10.1016/j.neuroimage.2015.06.093.

Andrews, M., Frank, S., & Vigliocco, G. (2014). Reconciling embodied and distributional accounts of meaning in language. *Topics in Cognitive Science, 6*(3), 359–370.

Andrews, M., Vigliocco, G., & Vinson, D. (2009). Integrating experiential and distributional data to learn semantic representations. *Psychological Review, 116*(3), 463.

Andric, M., Goldin-Meadow, S., Small, S. L., & Hasson, U. (2016). Repeated movie viewings produce similar local activity patterns but different network configurations. *NeuroImage, 142*, 613–627.

Andric, M., & Hasson, U. (2015). Global features of functional brain networks change with contextual disorder. *Neuroimage.* https://doi.org/10.1016/j.neuroimage.2015.05.025.

Armeni, K., Willems, R. M., & Frank, S. (2017). Probabilistic language models in cognitive neuroscience: Promises and pitfalls. *Neuroscience & Biobehavioral Reviews*.

Arnold, D., Tomaschek, F., Sering, K., Lopez, F., & Baayen, R. H. (2017). Words from spontaneous conversational speech can be recognized with human-like accuracy by an error-driven learning algorithm that discriminates between meanings straight from smart acoustic features, bypassing the phoneme as recognition unit. *PloS One, 12*(4), e0174623.

Azulay, H., Striem, E., & Amedi, A. (2009). Negative BOLD in sensory cortices during verbal memory: A component in generating internal representations? *Brain Topogr, 21*(3–4), 221–231. https://doi.org/10.1007/s10548-009-0089-2.

Baayen, R. H., Milin, P., Đurđević, D. F., Hendrix, P., & Marelli, M. (2011). An amorphous model for morphological processing in visual comprehension based on naive discriminative learning. *Psychological Review, 118*(3), 438.

Baldassano, C., Chen, J., Zadbood, A., Pillow, J. W., Hasson, U., & Norman, K. A. (2017). Discovering event structure in continuous narrative perception and memory. *Neuron, 95*(3), 709–721.

Bara, B. G. (2010). *Cognitive pragmatics.* Cambridge, MA: MIT Press.

Baroni, M., & Zamparelli, R. (2010). Nouns are vectors, adjectives are matrices: Representing adjective-noun constructions in semantic space. In *Paper presented at the empirical methods in natural language processing (EMNLP 2010)*. East Stroudsburg, PA.

Bassett, D. S., Meyer-Lindenberg, A., Achard, S., Duke, T., & Bullmore, E. (2006). Adaptive reconfiguration of fractal small-world human brain functional networks. *Proceedings of the National Academy od Sciences USA, 103*(51), 19518–19523. https://doi.org/10.1073/pnas.0606005103.

Bastiaansen, M., & Hagoort, P. (2015). Frequency-based segregation of syntactic and semantic unification during online sentence level language comprehension. *Journal of Cognitive Neurosciences, 27*(11), 2095–2107. https://doi.org/10.1162/jocn_a_00829.

Bastiaansen, M., Magyari, L., & Hagoort, P. (2010). Syntactic unification operations are reflected in oscillatory dynamics during on-line sentence comprehension. *Journal of Cognitive Neurosciences, 22*(7), 1333–1347. https://doi.org/10.1162/jocn.2009.21283.

Beckmann, M., Johansen-Berg, H., & Rushworth, M. F. (2009). Connectivity-based parcellation of human cingulate cortex and its relation to functional specialization. *Journal of Neurosciences, 29*(4), 1175–1190. https://doi.org/10.1523/JNEUROSCI.3328-08.2009.

Bein, O., Reggev, N., & Maril, A. (2014). Prior knowledge influences on hippocampus and medial prefrontal cortex interactions in subsequent memory. *Neuropsychologia, 64*, 320–330.

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.

Bever, T. G. (1970). The cognitive basis for linguistic structures. In R. Hayes (Ed.). *Cognition and language development* (pp. 279–362). New York: Wiley.

Binder, J. R., & Desai, R. H. (2011). The neurobiology of semantic memory. *Trends in Cognitive Sciences, 15*(11), 527–536. https://doi.org/10.1016/j.tics.2011.10.001.

Blank, I. A., & Fedorenko, E. (2017). Domain-general brain regions do not track linguistic input as closely as language-selective regions. *Journal of Neuroscience, 37*(41), 9999–10011.

Bornkessel-Schlesewsky, I., Krauspenhaar, S., & Schlesewsky, M. (2013). Yes, you can? A speaker's potency to act upon his words orchestrates early neural responses to message-level meaning. *PLoS One, 8*(7), e69173.

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2013). Reconciling time, space and function: A new dorsal-ventral stream model of sentence comprehension. *Brain and Language, 125*(1), 60–76. https://doi.org/10.1016/j.bandl.2013.01.010.

Bransford, J. D., Barclay, J. R., & Franks, J. J. (1972). Sentence memory: A constructive versus interpretive approach. *Cognitive Psychology, 3*(2), 193–209.

Bransford, J. D., & Johnson, Marcia K. (1972). Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior, 11*(6), 717–726. https://doi.org/10.1016/S0022-5371(72)80006-9.

Brennan, J., Nir, Y., Hasson, U., Malach, R., Heeger, D. J., & Pylkkanen, L. (2012). Syntactic structure building in the anterior temporal lobe during natural story listening. *Brain and Language, 120*(2), 163–173. https://doi.org/10.1016/j.bandl.2010.04.002.

Broderick, M. P., Anderson, A. J., Di Liberto, G. M., Crosse, M. J., & Lalor, E. C. (2018). Electrophysiological correlates of semantic dissimilarity reflect the comprehension of natural, narrative speech. *Current Biology, 28*, 803809.

Bruni, E., Tran, N. K., & Baroni, M. (2014). Multimodal distributional semantics. *Journal of Artificial Intelligence Research, 49*(1), 1–47.

Caplan, D., Stanczak, L., & Waters, G. (2008). Syntactic and thematic constraint effects on blood oxygenation level dependent signal correlates of comprehension of relative clauses. *Journal of Cognitive Neuroscience, 20*(4), 643656. https://doi.org/10.1162/jocn.2008.20044.

Caramazza, A., Capasso, R., Capitani, E., & Miceli, G. (2005). Patterns of comprehension performance in agrammatic Broca's aphasia: A test of the Trace Deletion Hypothesis. *Brain and Language, 94*(1), 43–53.

Chai, L. R., Mattar, M. G., Blank, I. A., Fedorenko, E., & Bassett, D. S. (2016). Functional network dynamics of the language system. *Cerebral Cortex, 26*(11), 4148–4159.

Chatham, C. H., & Badre, D. (2014). How to test cognitive theory with fMRI. arXiv preprint arXiv:1404.2917.

Chen, Z. J., He, Y., Rosa-Neto, P., Germann, J., & Evans, A. C. (2008). Revealing modular architecture of human brain structural networks by using cortical thickness from MRI. *Cereb Cortex, 18*(10), 2374–2381. https://doi.org/10.1093/cercor/bhn003.

Chomsky, N. (1959). Review of Skinner's verbal behavior. *Language, 35*, 2658.

Chow, H. M., Mar, R. A., Xu, Y., Liu, S., Wagage, S., & Braun, A. R. (2013). Embodied comprehension of stories: Interactions between language regions and modality-specific neural systems. *Journal of Cognitive Neurosciences.* https://doi.org/10.1162/jocn_a_00487.

Christiansen, M. H., & Chater, N. (2015a). The language faculty that wasn't: A usage-based account of natural language recursion. *Frontiers in Psychology, 6*, 1182. https://doi.org/10.3389/fpsyg.2015.01182.

Christiansen, M. H., & Chater, N. (2015b). The now-or-never bottleneck: A fundamental constraint on language. *Behaviours in Brain Sciences*, 1–52. https://doi.org/10.1017/S0140525X1500031X.

Christiansen, M. H., Conway, C. M., & Onnis, L. (2012). Similar neural correlates for language and sequential learning: Evidence from event-related brain potentials. *Language and Cognitive Process, 27*(2), 231–256. https://doi.org/10.1080/01690965.2011.606666.

Chung, W., Tucker, D. M., West, P., Potts, G. F., Liotti, M., Luu, P., & Hartry, A. L. (1996). Emotional expectancy: Brain electrical activity associated with an emotional bias in interpreting life events. *Psychophysiology, 33*(3), 218–233.

Chwilla, D. J., Virgillito, D., & Vissers, C. T. W. M. (2011). The relationship of language and emotion: N400 support for an embodied view of language comprehension. *Journal of Cognitive Neurosciences, 23*(9), 2400–2414.

Ciaramidaro, A., Adenzato, M., Enrici, I., Erk, S., Pia, L., Bara, B. G., & Walter, H. (2007). The intentional network: How the brain reads varieties of intentions. *Neuropsychologia, 45*(13), 3105–3113. https://doi.org/10.1016/j.neuropsychologia.2007.05.011.

Cibelli, E. S., Leonard, M. K., Johnson, K., & Chang, E. F. (2015). The influence of lexical statistics on temporal lobe cortical dynamics during spoken word listening. *Brain and Language, 147*, 66–75.

Clore, G. L., & Huntsinger, J. R. (2007). How emotions inform judgment and regulate thought. *Trends in Cognitive Sciences, 11*(9), 393–399.

Cooper, E. A., Hasson, U., & Small, S. L. (2011). Interpretation-mediated changes in neural activity during language comprehension. *Neuroimage, 55*(3), 1314–1323. https://doi.org/10.1016/j.neuroimage.2011.01.003.

Cruse, D. A. (1986). *Lexical semantics.* Cambridge Cambridgeshire; New York: Cambridge University Press.

Daltrozzo, J., & Schon, D. (2009). Conceptual processing in music as revealed by N400 effects on words and musical targets. *Journal of Cognitive Neurosciences, 21*(10), 1882–1892. https://doi.org/10.1162/jocn.2009.21113.

Danker, J. F., & Anderson, J. R. (2010). The ghosts of brain states past: Remembering reactivates the brain regions engaged during encoding. *Psychological Bulletin, 136*(1), 87102. https://doi.org/10.1037/a0017937.

Danks, D. (2003). Equilibria of the Rescorla-Wagner model. *Journal of Mathematical Psychology, 47*(2), 109–121.

Daselaar, S. M., Fleck, M. S., & Cabeza, R. (2006). Triple dissociation in the medial temporal lobes: Recollection, familiarity, and novelty. *Journal of Neurophysiology, 96*(4), 1902–1911. https://doi.org/10.1152/jn.01029.2005.

Dehaene-Lambertz, G., Dehaene, S., Anton, J. L., Campagne, A., Ciuciu, P., Dehaene, G. P., ... Dehaene, J. B. (2006). Functional segregation of cortical language areas by sentence repetition. *Human Brain Mapping, 27*(5), 360–371. https://doi.org/10.1002/hbm.20250.

Del Gaizo, J., Fridriksson, J., Yourganov, G., Hillis, A. E., Hickok, G., Misic, B., ... Bonilha, L. (2017). Mapping language networks using the structural and dynamic brain connectomes. *eNeuro, 4*(5) ENEURO-0204.

Delogu, F., Crocker, M. W., & Drenhaus, H. (2017). Teasing apart coercion and surprisal: Evidence from eye-movements and ERPs. *Cognition, 161*, 46–59.

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neurosciences, 8*(8), 1117–1121. https://doi.org/10.1038/nn1504.

Demberg, V., & Keller, F. (2008). Data from eye-tracking corpora as evidence for theories of syntactic processing complexity. *Cognition, 109*(2), 193–210.

den Ouden, D. B., Saur, D., Mader, W., Schelter, B., Lukic, S., Wali, E., ... Thompson, C. K. (2012). Network modulation during complex syntactic processing. *Neuroimage, 59*(1), 815–823. https://doi.org/10.1016/j.neuroimage.2011.07.057.

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. https://doi.org/10.1162/jocn.2009.21070.

Dick, A. S., Mok, E. H., Raja Beharelle, A., Goldin-Meadow, S., & Small, S. L. (2014). Frontal and temporal contributions to understanding the iconic co-speech gestures that accompany speech. *Human Brain Mapping, 35*(3), 900–917. https://doi.org/10.1002/hbm.22222.

Dick, A. S., & Tremblay, P. (2012). Beyond the arcuate fasciculus: Consensus and controversy in the connectional anatomy of language. *Brain, 135*(12), 3529–3550.

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

Duarte, A., Henson, R. N., & Graham, K. S. (2011). Stimulus content and the neural correlates of source memory. *Brain Research, 1373*, 110–123. https://doi.org/10.1016/j.brainres.2010.11.086.

Egidi, G., & Caramazza, A. (2013). Cortical systems for local and global integration in discourse comprehension. *Neuroimage, 71*, 59–74. https://doi.org/10.1016/j.neuroimage.2013.01.003.

Egidi, G., & Caramazza, A. (2014). Mood-dependent integration in discourse comprehension: Happy and sad moods affect consistency processing via different brain networks. *Neuroimage, 103*, 20–32. https://doi.org/10.1016/j.neuroimage.2014.09.008.

Egidi, G., & Nusbaum, H. C. (2012). Emotional language processing: How mood affects integration processes during discourse comprehension. *Brain Lang, 122*(3), 199–210. https://doi.org/10.1016/j.bandl.2011.12.008.

Egorova, N., Shtyrov, Y., & Pulvermuller, F. (2016). Brain basis of communicative actions in language. *Neuroimage, 125*, 857–867. https://doi.org/10.1016/j.neuroimage.2015.10.055.

Ellis, Nick C. (2002). Frequency effects in language processing. *Studies in Second Language Acquisition, 24*(02), 143–188.

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

Ettinger, A., Linzen, T., & Marantz, A. (2014). The role of morphology in phoneme prediction: Evidence from MEG. *Brain and Language, 129*, 14–23.

Federmeier, K. D., Kirson, D. A., Moreno, E. M., & Kutas, M. (2001). Effects of transient, mild mood states on semantic memory organization and use: An event-related potential investigation in humans. *Neuroscience Letters, 305*(3), 149–152.

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, E., & Thompson-Schill, S. L. (2014). Reworking the language network. *Trends in Cognitive Sciences, 18*(3), 120–126. https://doi.org/10.1016/j.tics.2013.12.006.

Ferreira, F., & Clifton, C. (1986). The independence of syntactic processing. *Journal of Memory and Language, 25*(3), 348–368. https://doi.org/10.1016/0749-596X(86)90006-9.

Ferstl, E. C., Neumann, J., Bogler, C., & von Cramon, D. Y. (2008). The extended language network: A meta-analysis of neuroimaging studies on text comprehension. *Human Brain Mapping, 29*(5), 581–593. https://doi.org/10.1002/hbm.20422.

Fiebach, C. J., Gruber, T., & Supp, G. G. (2005). Neuronal mechanisms of repetition priming in occipitotemporal cortex: Spatiotemporal evidence from functional magnetic resonance imaging and electroencephalography. *Journal of Neurosciences, 25*(13), 3414–3422. https://doi.org/10.1523/JNEUROSCI.4107-04.2005.

Fiebach, C. J., & Schubotz, R. I. (2006). Dynamic anticipatory processing of hierarchical sequential events: A common role for Broca's area and ventral premotor cortex across domains? *Cortex, 42*(4), 499–502.

Fiedler, K. (2001). Affective states trigger processes of assimilation and accommodation. In L. L. Martin, & G. L. Clore (Eds.). *Theories of mood and cognition: A user's guidebook* (pp. 85–98). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Fields, E. C., & Kuperberg, G. R. (2015). Loving yourself more than your neighbor: ERPs reveal online effects of a self-positivity bias. *Social Cognitive and Affective Neuroscience*. https://doi.org/10.1093/scan/nsv004.

Filik, R., & Leuthold, H. (2008). Processing local pragmatic anomalies in fictional contexts: Evidence from the N400. *Psychophysiology, 45*(4), 554–558. https://doi.org/10.1111/j.1469-8986.2008.00656.x.

Fitch, W. T. (2014). Toward a computational framework for cognitive biology: Unifying approaches from cognitive neuroscience and comparative cognition. *Physics of Life Reviews, 11*(3), 329–364. https://doi.org/10.1016/j.plrev.2014.04.005.

Fitch, W. T., Hauser, M. D., & Chomsky, N. (2005). The evolution of the language faculty: clarifications and implications. *Cognition, 97*(2), 179–210 discussion 211-125. doi:S0010-0277(05)00086-7 [pii] 10.1016/j.cognition.2005.02.005.

Forgas, J. P., & Koch, A. (2013). Mood effects on cognition. In M. D. Robinson, E. R. Watkins, & E. Harmon-Jones (Eds.). *Handbook of emotion and cognition*. New York: Guilford.

Forgas, J. P., & Locke, Judith. (2005). Affective influences on causal inferences: The effects of mood on attributions for positive and negative interpersonal episodes. *Cognition and Emotion, 19*(7), 1071–1081.

Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., & Raichle, M. E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences USA, 102*(27), 9673–9678. https://doi.org/10.1073/pnas.0504136102.

Frank, S. L. (2013). Uncertainty reduction as a measure of cognitive load in sentence comprehension. *Topics in Cognitive Science, 5*(3), 475–494.

Frank, M. J., & Badre, D. (2015). How cognitive theory guides neuroscience. *Cognition, 135*, 14–20. https://doi.org/10.1016/j.cognition.2014.11.009.

Frank, S. L., & Bod, R. (2011). Insensitivity of the human sentence-processing system to hierarchical structure. *Psychological Science, 22*(6), 829–834.

Frank, S. L., & Christiansen (2018). Hierarchical and sequential processing of language. *Language, Cognition, & Neuroscience*. https://doi.org/10.1080/23273798.2018.1424347.

Frank, S. L., Otten, L. J., Galli, G., & Vigliocco, G. (2015). The ERP response to the amount of information conveyed by words in sentences. *Brain Language, 140,* 1–11.

Frank, S. L., & Willems, R. M. (2017). Word predictability and semantic similarity show distinct patterns of brain activity during language comprehension. *Language,*

Cognition and Neuroscience, 32*(9), 1192–1203.

Frank, S. L., & Yang, J. (2018). Lexical representation explains cortical entrainment during speech comprehension. *PloS One, 13*(5), e0197304.

Fransson, P. (2006). How default is the default mode of brain function? Further evidence from intrinsic BOLD signal fluctuations. *Neuropsychologia, 44*(14), 2836–2845. https://doi.org/10.1016/j.neuropsychologia.2006.06.017.

Fransson, P., & Marrelec, G. (2008). The precuneus/posterior cingulate cortex plays a pivotal role in the default mode network: Evidence from a partial correlation network analysis. *Neuroimage, 42*(3), 1178–1184. https://doi.org/10.1016/j.neuroimage.2008.05.059.

Frégnac, Y. (2017). Big data and the industrialization of neuroscience: A safe roadmap for understanding the brain? *Science, 358*(6362), 470–477.

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

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

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

Friederici, A. D., & Gierhan, S. ME. (2013). The language network. *Current Opinion in Neurobiology, 23*(2), 250–254.

Friston, K. (2009). The free-energy principle: A rough guide to the brain? *Trends in Cognitive Science, 13*(7), 293–301. https://doi.org/10.1016/j.tics.2009.04.005.

Fruchter, J., & Marantz, A. (2015). Decomposition, lookup, and recombination: MEG evidence for the full decomposition model of complex visual word recognition. *Brain and Language, 143*, 81–96.

Galantucci, B., & Garrod, S. (2011). Experimental semiotics: A review. *Frontiers in Human Neuroscience, 5*. https://doi.org/10.3389/fnhum.2011.00011.

Gerrig, R. J. (1993). *Experiencing narrative worlds: On the psychological activities of reading.* New Haven: Yale University Press.

Gerrig, R. J., & McKoon, G. (2001). Memory processes and experiential continuity. *Psychological Sciences, 12*(1), 81–85. https://doi.org/10.1111/1467-9280.00314.

Goldberg, A. E. (2003). Constructions: A new theoretical approach to language. *Trends in Cognitive Sciences, 7*(5), 219–224.

Gottlieb, L. J., & Rugg, M. D. (2011). Effects of modality on the neural correlates of encoding processes supporting recollection and familiarity. *Learning Memory, 18*(9), 565–573. https://doi.org/10.1101/lm.2197211.

Gottlieb, L. J., Uncapher, M. R., & Rugg, M. D. (2010). Dissociation of the neural correlates of visual and auditory contextual encoding. *Neuropsychologia, 48*(1), 137–144. https://doi.org/10.1016/j.neuropsychologia.2009.08.019.

Graesser, A. C., Millis, K. K., & Zwaan, R. A. (1997). Discourse comprehension. *Annual Review of Psychology, 48*, 163–189. https://doi.org/10.1146/annurev.psych.48.1.163.

Griffiths, T. L., Steyvers, M., & Tenenbaum, J. B. T. (2007). Topics in semantic representation. *Psychological Review, 114*(2), 211–214.

Grodzinsky, Y., & Friederici, A. D. (2006). Neuroimaging of syntax and syntactic processing. *Current Opinion in Neurobiology, 16*(2), 240–246. https://doi.org/10.1016/j.conb.2006.03.007.

Grunwald, T., Lehnertz, K., Heinze, H. J., Helmstaedter, C., & Elger, C. E. (1998). Verbal novelty detection within the human hippocampus proper. *Proceedings of the National Academy of Sciences USA, 95*(6), 3193–3197.

Gurevich, O., Johnson, M. A., & Goldberg, A. E. (2010). Incidental verbatim memory for language. *Language and Cognition, 2*(01), 45–78. https://doi.org/10.1515/langcog.2010.003.

Hagoort, P. (2005). On Broca, brain, and binding: A new framework. *Trends in Cognitive Science, 9*(9), 416–423.

Hagoort, P. (2017). The core and beyond in the language-ready brain. *Neuroscience & Biobehavioral Reviews.*

Hagoort, P., Hald, L., Bastiaansen, M., & Petersson, K. M. (2004). Integration of word meaning and world knowledge in language comprehension. *Science, 304*(5669), 438–441. https://doi.org/10.1126/science.1095455.

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

Hagoort, P., & van Berkum, J. (2007). Beyond the sentence given. *Philosophical Transacations on Royal Society of London B: Biological Sciences, 362*(1481), 801–811. https://doi.org/10.1098/rstb.2007.2089.

Hampson, M., Tokoglu, F., Sun, Z., Schafer, R. J., Skudlarski, P., Gore, J. C., & Constable, R. T. (2006). Connectivity-behavior analysis reveals that functional connectivity between left BA39 and Broca's area varies with reading ability. *Neuroimage, 31*(2), 513–519. https://doi.org/10.1016/j.neuroimage.2005.12.040.

Harman, G. (1965). The inference to the best explanation. *Philosophical Review, 74*(8895).

Harris, Z. S. (1954). Distributional structure. *Word, 10*(23).

Hasson, U., Furman, O., Clark, D., Dudai, Y., & Davachi, L. (2008). Enhanced intersubject correlations during movie viewing correlate with successful episodic encoding. *Neuron, 57*(3), 452–462. https://doi.org/10.1016/j.neuron.2007.12.009.

Hasson, U., Nusbaum, H. C., & Small, S. L. (2006). Repetition suppression for spoken sentences and the effect of task demands. *Journal of Cognitive Neuroscience, 18*(12), 2013–2029. https://doi.org/10.1162/jocn.2006.18.12.2013.

Hasson, U., Nusbaum, H. C., & Small, S. L. (2007). Brain networks subserving the extraction of sentence information and its encoding to memory. *Cerebral Cortex, 17*(12), 2899–2913. https://doi.org/10.1093/cercor/bhm016.

Hasson, U., Nusbaum, H. C., & Small, S. L. (2009). Task-dependent organization of brain regions active during rest. *Proceedings of the National Academy of Science USA, 106*(26), 10841–10846. https://doi.org/10.1073/pnas.0903253106.

Hasson, U., & Tremblay, P. (2016). Neurobiology of statistical information processing in the auditory domain. In L. S. Steven, & G. Hickok (Eds.). *Neurobiology of language* (pp. 527–537). San Diego: Academic Press.

Hendrix, P., Nick, J., & Baayen, R. H. (2014). Compound reading in natural discourse contexts. In *Paper presented at the The Mental Lexicon conference*. Niagara-on-the-Lake, Canada.

Hesselmann, G., Sadaghiani, S., Friston, K. J., & Kleinschmidt, A. (2010). Predictive coding or evidence accumulation? False inference and neuronal fluctuations. *PLoS One, 5*(3), e9926. https://doi.org/10.1371/journal.pone.0009926.

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews. Neuroscience, 8*(5), 393–402. https://doi.org/10.1038/nrn2113.

Hoen, M., Pachot-Clouard, M., Segebarth, C., & Dominey, P. F. (2006). When Broca experiences the Janus syndrome: An ER-fMRI study comparing sentence comprehension and cognitive sequence processing. *Cortex, 42*(4), 605–623.

Honey, C. J., Thivierge, J.-P., & Sporns, O. (2010). Can structure predict function in the human brain? *Neuroimage, 52*(3), 766–776.

Honey, C. J., Thompson, C. R., Lerner, Y., & Hasson, U. (2012). Not lost in translation: Neural responses shared across languages. *Journal of Neuroscience, 32*(44), 15277–15283.

Hsu, C.-T., Jacobs, A. M., Citron, F. M. M., & Conrad, M. (2015). The emotion potential of words and passages in reading Harry Potter: An fMRI study. *Brain Language, 142*, 96–114. https://doi.org/10.1016/j.bandl.2015.01.011.

Huettig, F. (2015). Four central questions about prediction in language processing. *Brain Research, 64*. https://doi.org/10.1016/j.brainres.2015.02.014.

Huettig, F., Rommers, J., & Meyer, A. S. (2011). Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta Psychologica, 137*(2), 151–171.

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.

Hutchison, R. M., Womelsdorf, T., Allen, E. A., Bandettini, P. A., Calhoun, V. D., Corbetta, M., ... Chang, C. (2013). Dynamic functional connectivity: Promise, issues, and interpretations. *Neuroimage, 80*, 360–378. https://doi.org/10.1016/j.neuroimage.2013.05.079.

Ibbotson, P. (2013). The scope of usage-based theory. *Frontiers in Psychology, 4*, 255.

Johnson, J. D., McDuff, S. G., Rugg, M. D., & Norman, K. A. (2009). Recollection, familiarity, and cortical reinstatement: A multivoxel pattern analysis. *Neuron, 63*(5), 697–708. https://doi.org/10.1016/j.neuron.2009.08.011.

Kaan, E., & Swaab, T. Y. (2002). The brain circuitry of syntactic comprehension. *Trends in Cognitive Science, 6*(8), 350–356.

Keller, T. A., Carpenter, P. A., & Just, M. A. (2001). The neural bases of sentence comprehension: A fMRI examination of syntactic and lexical processing. *Cereb Cortex, 11*(3), 223–237.

Kemmerer, D. (2015). Are the motor features of verb meanings represented in the precentral motor cortices? Yes, but within the context of a flexible, multilevel architecture for conceptual knowledge. *Psychonomic Bulletin & Review, 22*(4), 1068–1075.

Kielar, A., Panamsky, L., Links, K. A., & Meltzer, J. A. (2015). Localization of electrophysiological responses to semantic and syntactic anomalies in language comprehension with MEG. *Neuroimage, 105*, 507–524. https://doi.org/10.1016/j.neuroimage.2014.11.016.

Koelsch, S., Kasper, E., Sammler, D., Schulze, K., Gunter, T., & Friederici, A. D. (2004). Music, language and meaning: Brain signatures of semantic processing. *Nature Neurosciences, 7*(3), 302–307. https://doi.org/10.1038/nn1197.

Kotz, S. A., Meyer, M., Alter, K., Besson, M., von Cramon, D. Y., & Friederici, A. D. (2003). On the lateralization of emotional prosody: an event-related functional MR investigation. *Brain Language, 86*(3), 366–376.

Kragel, J. E., Morton, N. W., & Polyn, S. M. (2015). Neural activity in the medial temporal lobe reveals the fidelity of mental time travel. *Journal of Neuroscience, 35*(7), 2914–2926.

Kristensen, L. B., Engberg-Pedersen, E., & Wallentin, M. (2014). Context predicts word order processing in Broca's region. *Journal of Cognitive Neuroscience, 26*(12), 2762–2777. https://doi.org/10.1162/jocn_a_00681.

Kuhlen, A. K., Bogler, C., Brennan, S. E., & Haynes, J. D. (2017). Brains in dialogue: Decoding neural preparation of speaking to a conversational partner. *Social Cognitive and Affective Neuroscience, 64*. https://doi.org/10.1093/scan/nsx018.

Kümmerer, M., Wallis, T. S. A. & Bethge, M. (2016). DeepGaze II: Reading fixations from deep features trained on object recognition. arXiv 116.

Kuperberg, G. R., Lakshmanan, B. M., Caplan, D. N., & Holcomb, P. J. (2006). Making sense of discourse: An fMRI study of causal inferencing across sentences. *Neuroimage, 33*(1), 343–361. https://doi.org/10.1016/j.neuroimage.2006.06.001.

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

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The Latent Semantic Analysis theory of acquisition, induction and representation of knowledge. *Psychological Review, 104*(2), 211–240.

Lane, C., Kanjlia, S., Richardson, H., Fulton, A., Omaki, A., & Bedny, M. (2017). Reduced left lateralization of language in congenitally blind individuals. *Journal of Cognitive Neuroscience, 29*(1), 65–78.

Lazaridou, A., Marelli, M., & Baroni, M. (2017). Multimodal word meaning induction from minimal exposure to natural text. *Cognitive Science, 41*(4), 677–705.

Lebois, L. A., Wilson-Mendenhall, C. D., & Barsalou, L. W. (2015). Are automatic conceptual cores the gold standard of semantic processing? The context-dependence of spatial meaning in grounded congruency effects. *Cognitive Science, 39*(8), 1764–1801.

Leiken, K., & Pylkkanen, L. (2014). MEG evidence that the LIFG effect of object extraction requires similarity-based interference. *Language and Cognitive Processes, 29*(3), 381–389. https://doi.org/10.1080/01690965.2013.863369.

Leonard, M. K., Baud, M. O., Sjerps, M. J., & Chang, E. F. (2016). Perceptual restoration of masked speech in human cortex. *Nature Communications, 7*, 13619.

Leonard, M. K., Bouchard, K. E., Tang, C., & Chang, E. F. (2015). Dynamic encoding of speech sequence probability in human temporal cortex. *Journal of Neuroscience, 35*(18), 7203–7214.

Levelt, W. (2012). *A history of psycholinguistics: The pre-Chomskyan era*. Oxford University Press.

Levinson, S. C. (2006). On the human interactional engine. In N. J. Enfield, & S. C. Levinson (Eds.). *Roots of human sociality*. New York: Berg.

Levin, D. N., & Uftring, S. J. (2001). Detecting brain activation in FMRI data without prior knowledge of mental event timing. *Neuroimage, 13*(1), 153–160. https://doi.org/10.1006/nimg.2000.0663.

Lewis, A. G., & Bastiaansen, M. (2015). A predictive coding framework for rapid neural dynamics during sentence-level language comprehension. *Cortex, 68*, 155–168.

Liang, X., Zou, Q., He, Y., & Yang, Y. (2015). Topologically reorganized connectivity architecture of default-mode, executive-control, and salience networks across working memory task loads. *Cerebral Cortex, 64*. https://doi.org/10.1093/cercor/bhu316.

Liszkowski, U. (2006). Infant pointing at 12 months: Communicative goals, motives, and social-cognitive abilities. In N. J. Enfield, & S. C. Levinson (Eds.). *Roots of human sociality*. New York, NY: Berg.

Li, C. S., Yan, P., Bergquist, K. L., & Sinha, R. (2007). Greater activation of the "default" brain regions predicts stop signal errors. *Neuroimage, 38*(3), 640–648. https://doi.org/10.1016/j.neuroimage.2007.07.021.

Lopopolo, A., Frank, S. L., Van den Bosch, A., & Willems, R. M. (2017). Using stochastic language models (SLM) to map lexical, syntactic, and phonological information processing in the brain. *Plos ONE, 12*(5), e0177794.

Lopopolo, A., & van Miltenburg, E. (2015). Sound-based distributional models. *IWCS, 2015*, 70.

Makuuchi, M., & Friederici, A. D. (2013). Hierarchical functional connectivity between the core language system and the working memory system. *Cortex, 49*(9), 2416–2423. https://doi.org/10.1016/j.cortex.2013.01.007.

Mak, W. M., Vonk, W., & Schriefers, H. (2008). Discourse structure and relative clause processing. *Memory and Cognition, 36*(1), 170–181.

Mandera, P., Keuleers, E., & Brysbaert, M. (2017). Explaining human performance in psycholinguistic tasks with models of semantic similarity based on prediction and counting: A review and empirical validation. *Journal of Memory and Language, 92*, 57–78.

Marelli, M., & Baroni, M. (2015). Affixation in semantic space: Modeling morpheme meanings with compositional distributional semantics. *Psychological Review, 122*(3), 485–515.

Marelli, M., Gagné, C., & Spalding, T. (2017). Compounding as abstract operation in semantic space: Investigating relational effects through a large-scale, data-driven computational model. *Cognition, 166*, 207–224.

Martin-Loeches, M., Casado, P., Hernandez-Tamames, J. A., & Alvarez-Linera, J. (2008). Brain activation in discourse comprehension: A 3t fMRI study. *Neuroimage, 41*(2), 614–622. https://doi.org/10.1016/j.neuroimage.2008.02.047.

Mason, R. A., & Just, M. A. (2004). How the brain processes causal inferences in text a theoretical account of generation and integration component processes utilizing both cerebral hemispheres. *Psychological Science, 15*(1), 1–7.

Mason, M. F., Norton, M. I., Van Horn, J. D., Wegner, D. M., Grafton, S. T., & Macrae, C. N. (2007). Wandering minds: The default network and stimulus-independent thought. *Science, 315*(5810), 393–395. https://doi.org/10.1126/science.1131295.

McNealy, K., Mazziotta, J. C., & Dapretto, M. (2006). Cracking the language code: Neural mechanisms underlying speech parsing. *Journal of Neurosciences, 26*(29), 7629–7639. https://doi.org/10.1523/JNEUROSCI.5501-05.2006.

Menenti, L., Petersson, K. M., Scheeringa, R., & Hagoort, P. (2009). When elephants fly: Differential sensitivity of right and left inferior frontal gyri to discourse and world knowledge. *Journal of Cognitive Neuroscience, 21*(12), 2358–2368. https://doi.org/10.1162/jocn.2008.21163.

Mennes, M., Kelly, C., Colcombe, S., Castellanos, F. X., & Milham, M. P. (2012). The extrinsic and intrinsic functional architectures of the human brain are not equivalent. *Cerebral Cortex, 23*(1), 223–229.

Messe, A., Rudrauf, D., Benali, H., & Marrelec, G. (2014). Relating structure and function in the human brain: Relative contributions of anatomy, stationary dynamics, and non-stationarities. *PLoS Computational Biology, 10*(3), e1003530. https://doi.org/10.1371/journal.pcbi.1003530.

Meyer, L. (2017). The neural oscillations of speech processing and language comprehension: State of the art and emerging mechanisms. *European Journal of Neuroscience, 64*. https://doi.org/10.1111/ejn.13748.

Milin, P., Feldman, L. B., Ramscar, M., Hendrix, P., & Baayen, R. H. (2017). Discrimination in lexical decision. *PLoS One, 12*(2), e0171935.

Mitchell, J., & Lapata, M. (2010). Composition in distributional models of semantics. *Cognitive Science, 34*(8), 1388–1429.

Molinari, E., Baraldi, P., Campanella, M., Duzzi, D., Nocetti, L., Pagnoni, G., & Porro, C. A. (2013). Human parietofrontal networks related to action observation detected at rest. *Cerebral Cortex, 23*(1), 178–186. https://doi.org/10.1093/cercor/bhr393.

Moussa, M. N., Vechlekar, C. D., Burdette, J. H., Steen, M. R., Hugenschmidt, C. E., & Laurienti, P. J. (2011). Changes in cognitive state alter human functional brain networks. *Frontiers in Human Neuroscience, 5*, 83. https://doi.org/10.3389/fnhum.2011.00083.

Müller, N., Keil, J., Obleser, J., Schulz, H., Grunwald, T., Bernays, R.-L., ... Weisz, N. (2013). You can't stop the music: Reduced auditory alpha power and coupling between auditory and memory regions facilitate the illusory perception of music during noise. *Neuroimage, 79*, 383393. https://doi.org/10.1016/j.neuroimage.2013.05.001.

Muller, A. M., & Meyer, M. (2014). Language in the brain at rest: new insights from resting state data and graph theoretical analysis. *Frontiers in Human Neuroscience, 8*, 228. https://doi.org/10.3389/fnhum.2014.00228.

Murphy, B., Wehbe, L., & Fyshe, A. (2018). Decoding language from the brain. In T. Poibeau, & A. Villavicencio (Eds.). *Language, cognition, and computational models* (pp. 53–80). Cambridge: Cambridge University Press.

Mustovic, H., Scheffler, K., Di Salle, F., Esposito, F., Neuhoff, J. G., Hennig, J., & Seifritz, E. (2003). Temporal integration of sequential auditory events: Silent period in sound pattern activates human planum temporale. *Neuroimage, 20*(1), 429–434.

Nastase, S., Iacovella, V., & Hasson, U. (2014). Uncertainty in visual and auditory series is coded by modality-general and modality-specific neural systems. *Human Brain Mapping, 35*(4), 1111–1128. https://doi.org/10.1002/hbm.22238.

Nelson, M. J., El Karoui, I., Giber, K., Yang, X., Cohen, L., Koopman, H., ... Dehaene, S. (2017). Neurophysiological dynamics of phrase-structure building during sentence processing. *Proceedings of the National Academy of Sciences, 114*(18), E3669–E3678. https://doi.org/10.1073/pnas.1701590114.

Nieuwland, M. S., Politzer-Ahles, S., Heyselaar, E., Segaert, K., Darley, E., Kazanina, N., ... Mézière, D. (2018). Large-scale replication study reveals a limit on probabilistic prediction in language comprehension. *eLife, 7*, e33468.

Nieuwland, M. S., & van Berkum, J. J. (2006). When peanuts fall in love: N400 evidence for the power of discourse. *Journal of Cognitive Neuroscience, 18*(7), 1098–1111. https://doi.org/10.1162/jocn.2006.18.7.1098.

Nijhof, A. D., & Willems, R. M. (2015). Simulating fiction: Individual differences in literature comprehension revealed with FMRI. *PLoS One, 10*(2), https://doi.org/10.1371/journal.pone.0116492.

Novick, J. M., Trueswell, J. C., & Thompson-Schill, S. L. (2005). Cognitive control and parsing: Reexamining the role of Broca's area in sentence comprehension. *Cognitive, Affective, & Behavioral Neuroscience, 5*(3), 263–281.

Otten, L. J., Quayle, A. H., Akram, S., Ditewig, T. A., & Rugg, M. D. (2006). Brain activity before an event predicts later recollection. *Nature Neuroscience, 9*(4), 489–491. https://doi.org/10.1038/nn1663.

Otten, L. J., & Rugg, M. D. (2001). Task-dependency of the neural correlates of episodic encoding as measured by fMRI. *Cerebral Cortex, 11*(12), 1150–1160.

Özyürek, A., Willems, R. M., Kita, S., & Hagoort, P. (2007). On-line integration of semantic information from speech and gesture: Insights from event-related brain potentials. *Journal of Cognitive Neuroscience, 19*(4), 605–616.

Paller, K. A., & Wagner, A. D. (2002). Observing the transformation of experience into memory. *Trends in Cognitive Science, 6*(2), 93–102.

Pessoa, L. (2014). Understanding brain networks and brain organization. *Physics of Life Reviews, 11*(3), 400–435.

Petridou, N., Gaudes, C. C., Dryden, I. L., Francis, S. T., & Gowland, P. A. (2013). Periods of rest in fMRI contain individual spontaneous events which are related to slowly fluctuating spontaneous activity. *Human Brain Mapping, 34*(6), 1319–1329. https://doi.org/10.1002/hbm.21513.

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences, 36*(4), 329–347. https://doi.org/10.1017/S0140525X12001495.

Pohlack, S. T., Meyer, P., Cacciaglia, R., Liebscher, C., Ridder, S., & Flor, H. (2014). Bigger is better! Hippocampal volume and declarative memory performance in healthy young men. *Brain Structure and Function, 219*(1), 255–267. https://doi.org/10.1007/s00429-012-0497-z.

Poldrack, R. A. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Science, 10*(2), 59–63. https://doi.org/10.1016/j.tics.2005.12.004.

Price, C. J. (2018). The evolution of cognitive models: From neuropsychology to neuroimaging and back. *Cortex, 64*, 64.

Radvansky, G. A., & Copeland, D. E. (2010). Reading times and the detection of event shift processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*(1), 210–216. https://doi.org/10.1037/a0017258.

Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience, 2*(1), 79–87. https://doi.org/10.1038/4580.

Rashid, B., Damaraju, E., Pearlson, G., & Calhoun, V. D. (2014). Dynamic connectivity states estimated from resting fMRI Identify differences among Schizophrenia, bipolar disorder, and healthy control subjects. *Frontires in Human Neuroscience, 8*, 897. https://doi.org/10.3389/fnhum.2014.00897.

Raz, G., Jacob, Y., Gonen, T., Winetraub, Y., Flash, T., Soreq, E., & Hendler, T. (2014). Cry for her or cry with her: Context-dependent dissociation of two modes of cinematic empathy reflected in network cohesion dynamics. *Social Cognitive and Affective Neuroscience, 9*(1), 30–38. https://doi.org/10.1093/scan/nst052.

Rescorla, R. A., & Wagner, A. W. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Classical Conditioning II: Current Research and Theory*, 64–99.

Richmond, L. L., & Zacks, J. M. (2017). Constructing experience: Event models from perception to action. *Trends in Cognitive Sciences, 21*, 962–980.

Robinson, L. F., Atlas, L. Y., & Wager, T. D. (2015). Dynamic functional connectivity using state-based dynamic community structure: Method and application to opioid analgesia. *Neuroimage, 108*, 274–291. https://doi.org/10.1016/j.neuroimage.2014.12.034.

Rodd, J. M., Davis, M. H., & Johnsrude, I. S. (2005). The neural mechanisms of speech comprehension: FMRI studies of semantic ambiguity. *Cerebral Cortex, 15*(8), 1261–1269. https://doi.org/10.1093/cercor/bhi009.

Rogalsky, C., & Hickok, G. (2010). The role of Broca's area in sentence comprehension. *Journal of Cognitive Neuroscience, 23*(7), 1664–1680. https://doi.org/10.1162/jocn.2010.21530.

Rommers, J., Dickson, D. S., Norton, J. J. S., Wlotko, E. W., & Federmeier, K. D. (2017). Alpha and theta band dynamics related to sentential constraint and word expectancy.

*Language, Cognition and Neuroscience, 32*(5), 576–589.

Rueschemeyer, S. A., Gardner, T., & Stoner, C. (2014). The social N400 effect: How the presence of other listeners affects language comprehension. *Psychonomical Bulletin Review, 64*. https://doi.org/10.3758/s13423-014-0654-x.

Sadaghiani, S., & Kleinschmidt, A. (2013). Functional interactions between intrinsic brain activity and behavior. *Neuroimage, 80*, 379–386. https://doi.org/10.1016/j.neuroimage.2013.04.100.

Santi, A., & Grodzinsky, Y. (2007). Working memory and syntax interact in Broca's area. *Neuroimage, 37*(1), 8–17. https://doi.org/10.1016/j.neuroimage.2007.04.047.

Schank, R. C., & Abelson, R. (1977). *Scripts, plans, goals, and understanding.* Hillsdale, NJ: Earlbaum.

Schoffelen, J. M., Hultn, A., Lam, N., Marquand, A. F., Uddn, J., & Hagoort, P. (2017). Frequency-specific directed interactions in the human brain network for language. *Proceedings of the National Academy of Science USA, 114*(30), 8083–8088. https://doi.org/10.1073/pnas.1703155114.

Schurz, M., Radua, J., Aichhorn, M., Richlan, F., & Perner, J. (2014). Fractionating theory of mind: A meta-analysis of functional brain imaging studies. *Neuroscience & Biobehavioral Reviews, 42*, 9–34. https://doi.org/10.1016/j.neubiorev.2014.01.009.

Shirer, W. R., Ryali, S., Rykhlevskaia, E., Menon, V., & Greicius, M. D. (2012). Decoding subject-driven cognitive states with whole-brain connectivity patterns. *Cerebral Cortex, 22*(1), 158–165. https://doi.org/10.1093/cercor/bhr099.

Simmons, W. K., & Martin, A. (2012). Spontaneous resting-state BOLD fluctuations reveal persistent domain-specific neural networks. *Social Cognitive and Affective Neuroscience, 7*(4), 467–475. https://doi.org/10.1093/scan/nsr018.

Sitnikova, T., Holcomb, P. J., Kiyonaga, K. A., & Kuperberg, G. R. (2008). Two neurocognitive mechanisms of semantic integration during the comprehension of visual real-world events. *Journal of Cognitive Neuroscience, 20*(11), 2037–2057. https://doi.org/10.1162/jocn.2008.20143.

Skeide, M. A., & Friederici, A. D. (2017). Neurolinguistic 19 studies of sentence comprehension. *The Handbook of Psycholinguistics, 438*.

Smirnov, D., Glerean, E., Lahnakoski, J. M., Salmi, J., Jaaskelainen, I. P., Sams, M., & Nummenmaa, L. (2014). Fronto-parietal network supports context-dependent speech comprehension. *Neuropsychologia, 63*, 293–303. https://doi.org/10.1016/j.neuropsychologia.2014.09.007.

Smith, S. M., Fox, P. T., Miller, K. L., Glahn, D. C., Fox, P. M., Mackay, C. E., ... Beckmann, C. F. (2009). Correspondence of the brain's functional architecture during activation and rest. *Proceedings of the National Academy of Sciences, 106*(31), 13040–13045.

Speer, N. K., & Zacks, J. M. (2005). Temporal changes as event boundaries: Processing and memory consequences of narrative time shifts. *Journal of Memory and Language, 53*(1), 125–140.

Speer, N. K., Zacks, J. M., & Reynolds, J. R. (2007). Human brain activity time-locked to narrative event boundaries. *Psychological Sciences, 18*(5), 449–455. https://doi.org/10.1111/j.1467-9280.2007.01920.x.

Sperber, D., & Wilson, D. (1995). *Relevance; communication and cognition, Vol. 2nd.* Oxford, UK: Basil Blackwell.

Spivey-Knowlton, M. J., & Sedivy, J. C. (1995). Resolving attachment ambiguities with multiple constraints. *Cognition, 55*(3), 227–267.

Spivey-Knowlton, M. J., Trueswell, J. C., & Tanenhaus, M. K. (1993). Context effects in syntactic ambiguity resolution: Discourse and semantic influences in parsing reduced relative clauses. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie exp rimentale, 47*(2), 276.

Spunt, R. P., & Adolphs, R. (2017). A new look at domain specificity: Insights from social neuroscience. *Nature Reviews Neuroscience, 18*(9), 559.

St. George, M., Kutas, M., Martinez, A., & Sereno, M. I. (1999). Semantic integration in reading: Engagement of the right hemisphere during discourse processing. *Brain, 122*(Pt 7), 1317–1325.

Tagliazucchi, E., Balenzuela, P., Fraiman, D., & Chialvo, D. R. (2012). Criticality in large-scale brain fMRI dynamics unveiled by a novel point process analysis. *Frontiers in Physiology, 3*, 15. https://doi.org/10.3389/fphys.2012.00015.

Takerkart, S., Auzias, G., Thirion, B., & Ralaivola, L. (2014). Graph-based inter-subject pattern analysis of FMRI data. *PLoS One, 9*(8), e104586. https://doi.org/10.1371/journal.pone.0104586.

Tobia, M., Iacovella, V., Davis, B., & Hasson, U. (2012). Neural systems mediating recognition of changes in statistical regularities. *NeuroImage, 63*, 1730–1742.

Tobia, M. J., Iacovella, V., & Hasson, U. (2012). Multiple sensitivity profiles to diversity and transition structure in non-stationary input. *Neuroimage, 60*(2), 991–1005. https://doi.org/10.1016/j.neuroimage.2012.01.041.

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition.* Cambridge, MA: Harvard University Press.

Tremblay, P., Baroni, M., & Hasson, U. (2013). Processing of speech and non-speech sounds in the supratemporal plane: Auditory input preference does not predict sensitivity to statistical structure. *Neuroimage, 66*, 318–332. https://doi.org/10.1016/j.neuroimage.2012.10.055.

Turney, P. D., & Pantel, P. (2010). From frequency to meaning: Vector space models of semantics. *Journal of Artificial Intelligence Research, 37*, 141–188.

Tzeng, Y., van den Broek, P., Kendeou, P., & Lee, C. (2005). The computational implementation of the landscape model: Modeling inferential processes and memory representations of text comprehension. *Behavior Research Methods, 37*(2), 277–286.

Uehara, T., Yamasaki, T., Okamoto, T., Koike, T., Kan, S., Miyauchi, S., ... Tobimatsu, S. (2014). Efficiency of a "small-world" brain network depends on consciousness level: A resting-state FMRI study. *Cerebral Cortex, 24*(6), 1529–1539. https://doi.org/10.1093/cercor/bht004.

van Berkum, J. J. (2008). Understanding sentences in context: What brain waves can tell us. *Current Directions in Psychological Science, 17*(6), 376–380. https://doi.org/10.1111/j.1467-8721.2008.00609.x.

van Berkum, J. J. (2009). The neuropragmatics of 'simple' utterance comprehension: An

ERP review. In U. Sauerland, & K. Yatsushiro (Eds.). *Semantics and pragmatics: From experiment to theory* (pp. 276–316). Basingstoke: Palgrave Macmillan.

van Berkum, J. J., Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(3), 443–467. https://doi.org/10.1037/0278-7393.31.3.443.

van Berkum, J. J., De Goede, D., Van Alphen, P. M., Mulder, E. R., & Kerstholt, J. H. (2013). How robust is the language architecture? The case of mood. *Frontiers in Psychology, 4*, 64.

van Berkum, J. J., Hagoort, P., & Brown, C. M. (1999). Semantic integration in sentences and discourse: Evidence from the N400. *Journal of Cognitive Neurosciences, 11*(6), 657–671.

van Berkum, J. J., Holleman, B., Nieuwland, M., Otten, M., & Murre, J. (2009). Right or wrong?: The brain's fast response to morally objectionable statements. *Psychological Science, 20*(9), 1092–1099. https://doi.org/10.1111/j.1467-9280.2009.02411.x.

van Berkum, J. J., van den Brink, D., Tesink, C. M., Kos, M., & Hagoort, P. (2008). The neural integration of speaker and message. *Journal of Cognitive Neuroscience, 20*(4), 580–591. https://doi.org/10.1162/jocn.2008.20054.

van Berkum, J. J., Zwitserlood, P., Hagoort, P., & Brown, C. M. (2003). When and how do listeners relate a sentence to the wider discourse? Evidence from the N400 effect. *Cognitive Brain Research, 17*(3), 701–718.

van Kesteren, M. T., Fernández, G., Norris, D. G., & Hermans, E. J. (2010). Persistent schema-dependent hippocampal-neocortical connectivity during memory encoding and postencoding rest in humans. *Proceedings of the National Academy of Sciences, 107*(16), 7550–7555.

Vecchi, E. M., Marelli, M., Zamparelli, R., & Baroni, M. (2017). Spicy adjectives and nominal donkeys: Capturing semantic deviance using compositionality in distributional spaces. *Cognitive Science, 41*(1), 102–136.

Vigneau, M., Beaucousin, V., Herve, P. Y., Duffau, H., Crivello, F., Houde, O., ... Tzourio-Mazoyer, N. (2006). Meta-analyzing left hemisphere language areas: Phonology, semantics, and sentence processing. *Neuroimage, 30*(4), 1414–1432. https://doi.org/10.1016/j.neuroimage.2005.11.002.

Vigneau, M., Beaucousin, V., Herve, P. Y., Jobard, G., Petit, L., Crivello, F., ... Tzourio-Mazoyer, N. (2011). What is right-hemisphere contribution to phonological, lexicosemantic, and sentence processing? Insights from a meta-analysis. *Neuroimage, 54*(1), 577–593. https://doi.org/10.1016/j.neuroimage.2010.07.036.

Vinyals, O., Toshev, A., Bengio, S., & Erhan, D. (2015, June). Show and tell: A neural image caption generator. In *2015 IEEE conference on Computer vision and pattern recognition (CVPR)* (pp. 3156-3164). IEEE.

Walter, H., Adenzato, M., Ciaramidaro, A., Enrici, I., Pia, L., & Bara, B. G. (2004). Understanding intentions in social interaction: The role of the anterior paracingulate cortex. *Journal of Cognitive Neuroscience, 16*(10), 1854–1863. https://doi.org/10.1162/0898929042947838.

Wang, L., Hagoort, P., & Jensen, O. (2017). Language prediction is reflected by coupling between frontal gamma and posterior alpha oscillations. *Journal of Cognitive Neuroscience, 64*, 1–16. https://doi.org/10.1162/jocn_a_01190.

Watson, L. A., Dritschel, B., Obonsawin, M. C., & Jentzsch, I. (2007). Seeing yourself in a positive light: Brain correlates of the self-positivity bias. *Brain Research, 1152*, 106–110.

Watson, R., Latinus, M., Noguchi, T., Garrod, O., Crabbe, F., & Belin, P. (2013). Dissociating task difficulty from incongruence in face-voice emotion integration. *Frontiers in Human Neurosciences, 7*, 744. https://doi.org/10.3389/fnhum.2013.00744.

Wehbe, L., Murphy, B., Talukdar, P., Fyshe, A., Ramdas, A., & Mitchell, T. (2014). Simultaneously uncovering the patterns of brain regions involved in different story reading subprocesses. *PLoS One, 9*(11), e112575. https://doi.org/10.1371/journal.pone.0112575.

Whitney, C., Huber, W., Klann, J., Weis, S., Krach, S., & Kircher, T. (2009). Neural correlates of narrative shifts during auditory story comprehension. *Neuroimage, 47*(1), 360–366. https://doi.org/10.1016/j.neuroimage.2009.04.037.

Wilkins, R. W., Hodges, D. A., Laurienti, P. J., Steen, M., & Burdette, J. H. (2014). Network science and the rffects of music preference on functional brain connectivity: from Beethoven to Eminem. *Scientific Reports, 4*, 6130. https://doi.org/10.1038/srep06130.

Willems, R. M., de Boer, M., de Ruiter, J. P., Noordzij, M. L., Hagoort, P., & Toni, I. (2010). A cerebral dissociation between linguistic and communicative abilities in humans. *Psychological Science, 21*(1), 8–14.

Willems, R. M., Frank, S. L., Nijhof, A. D., Hagoort, P., & Van den Bosch, A. (2016). Prediction during natural language comprehension. *Cerebral Cortex, 26*(6), 2506–2516.

Willems, R. M., Ozyurek, A., & Hagoort, P. (2008). Seeing and hearing meaning: ERP and fMRI evidence of word versus picture integration into a sentence context. *Journal of Cognitive Neuroscience, 20*(7), 1235–1249. https://doi.org/10.1162/jocn.2008.20085.

Willems, R. M., & Varley, R. A. (2010). Neural insights into the relation between language and communication. *Frontiers in Human Neuroscience, 4*, 203.

Wilson, S. M., Molnar-Szakacs, I., & Iacoboni, M. (2008). Beyond superior temporal cortex: Intersubject correlations in narrative speech comprehension. *Cerebral Cortex, 18*(1), 230–242. https://doi.org/10.1093/cercor/bhm049.

Wlotko, E. W., & Federmeier, K. D. (2012). So that's what you meant! Event-related potentials reveal multiple aspects of context use during construction of message-level meaning. *Neuroimage, 62*(1), 356–366. https://doi.org/10.1016/j.neuroimage.2012.04.054.

Yarkoni, T., Speer, N. K., Balota, D. A., McAvoy, M. P., & Zacks, J. M. (2008). Pictures of a thousand words: Investigating the neural mechanisms of reading with extremely rapid event-related fMRI. *Neuroimage, 42*(2), 973–987. https://doi.org/10.1016/j.neuroimage.2008.04.258.

Yee, E., & Thompson-Schill, S. L. (2016). Putting concepts into context. *Psychonomic Bulletin & Review, 23*(4), 1015–1027.

Zaccarella, E., Meyer, L., Makuuchi, M., & Friederici, A. D. (2017). Building by syntax: The neural basis of minimal linguistic structures. *Cerebral Cortex, 27*(1), 411–421.

Zacks, J. M., Braver, T. S., Sheridan, M. A., Donaldson, D. I., Snyder, A. Z., Ollinger, J. M., ... Raichle, M. E. (2001). Human brain activity time-locked to perceptual event boundaries. *Nature Neuroscience, 4*(6), 651–655. https://doi.org/10.1038/88486.

Zacks, J. M., Speer, N. K., & Reynolds, J. R. (2009). Segmentation in reading and film comprehension. *Journal of Experimental Psychology: General, 138*(2), 307–327. https://doi.org/10.1037/a0015305.

Zekveld, A. A., Heslenfeld, D. J., Festen, J. M., & Schoonhoven, R. (2006). Top-down and bottom-up processes in speech comprehension. *Neuroimage, 32*(4), 1826–1836. https://doi.org/10.1016/j.neuroimage.2006.04.199.

Zhu, Y., Kiros, R., Zemel, R., Salakhutdinov, R., Urtasun, R., Torralba, A., & Fidler, S. (2015). Aligning books and movies: Towards story-like visual explanations by watching movies and reading books. In *Proceedings of the IEEE international conference on computer vision* (pp. 19-27).

Zielinski, B. A., Gennatas, E. D., Zhou, J., & Seeley, W. W. (2010). Network-level structural covariance in the developing brain. *Proceedings of the National Academy of Sciences USA, 107*(42), 18191–18196. https://doi.org/10.1073/pnas.1003109107.

Zwaan, R. A. (2014). Embodiment and language comprehension: Reframing the discussion. *Trends in Cognitive Science, 18*(5), 229–234. https://doi.org/10.1016/j.tics.2014.02.008.

Zwaan, R. A., Langston, M. C., & Graesser, A. C. (1995). The construction of situation models in narrative comprehension: An event-indexing model. *Psychological Science,* 292–297.

Zwaan, R. A., Madden, C. J., & Whitten, S. N. (2000). The presence of an event in the narrated situation affects its availability to the comprehender. *Memory & Cognition, 28*(6), 1022–1028.