Available online at www.sciencedirect.com

ScienceDirect

Journal homepage: www.elsevier.com/locate/cortex

Special Issue: Neurocognitive Perspectives on Discourse and Connected Language.
Research Report

Mapping meaning in the brain's language

Harm Brouwer [a,b]

[a] Research Center for Cognitive Science and Artificial Intelligence, Tilburg University, the Netherlands
[b] Department of Computational Cognitive Science, Tilburg University, the Netherlands

ARTICLE INFO

Article history:
Received 9 April 2025
Revised 11 August 2025
Accepted 31 October 2025
Published online 19 November 2025

Keywords:
Decoding models
Encoding models
Formal semantics
Neurocognition of language

ABSTRACT

Recent advances in neuroscience and artificial intelligence have pushed the state-of-the-art from being able to decode the meaning of individual words from non-invasive brain recordings, to the reconstruction of the meaning of continuous language. Beyond game changing practical implications of such “mind reading” *mapping models*, e.g., brain-computer interfaces that restore lost ability to speak, they also hold the promise to be instrumental in addressing a fundamental question in the cognitive sciences: How does the human brain represent the meaning of concepts, phrases, and sentences? In order to fulfil this promise, however, important methodological and theoretical challenges need to be overcome: (1) extant mapping results are inconsistent and difficult to reconcile with neurocognitive theory, (2) extant neural meaning representations do not model the compositional semantics capturing the meaning of multi-word utterances, and (3) extant mapping models fail to take into account the spatiotemporal dynamics of lexical and compositional semantic representation and computation. I argue that in order to overcome these challenges, we should ground mapping models in linguistic and neurocognitive theory, and develop neurocomputational models that explicate the spatiotemporal dynamics of meaning in the brain's language.

1. Introduction

How does the human brain represent the meaning of concepts, phrases, and sentences? This is a central question in cognitive science that is studied in many of its subdisciplines, including linguistics, philosophy, and neuroscience. Recent advances in neuroscience and artificial intelligence have led to a proliferation of analysis methods that allow for the empirical investigation of neural semantic representations in the brain (Poldrack, 2011; King & Dehaene, 2014): Decoding models seek to accurately infer representations of meaning from neural activity, while encoding models invert this process, and aim to predict neural activity from such meaning representations. While these *mapping models* differ in directionality, both can be used to infer or “decode” neural semantic representations from neuroimaging data (Ivanova et al., 2023; Frisby et al., 2023).

Indeed, using vector-based neural semantic representations—which can be constructed from either lexical co-occurrences, e.g., using language models, or human

This article is part of a special issue entitled: Neurocognitive perspectives on discourse published in Cortex.

E-mail address: h.brouwer@tilburguniversity.edu.

https://doi.org/10.1016/j.cortex.2025.10.012
0010-9452/

ratings—it has previously been shown that it is possible to decode the meaning of words and sentences from non-invasive functional magnetic resonance imaging (fMRI) recordings (e.g., Anderson et al., 2017; Huth et al., 2012; Mitchell et al., 2008; Pereira et al., 2018). More recently, models employing the contextualized embeddings from large language models (LLMs; e.g., Devlin et al., 2018; Radford et al., 2018), trained on self-supervised next word prediction, have been shown to achieve state-of-the-art reconstruction of the meaning of continuous language from fMRI recordings (Tang et al., 2023), as well as to accurately predict the explainable variance in fMRI and invasive electrocorticography (ECoG) recordings (Schrimpf et al., 2021; Goldstein et al., 2022).

Beyond game changing practical implications of such "mind reading" *mapping models*, e.g., brain-computer interfaces that restore lost ability to speak, they also hold the promise to be instrumental in unraveling the organization, representation, and computation of meaning in the brain. However, in order to fulfil this promise, important methodological and theoretical challenges need to be overcome: (1) extant mapping results are inconsistent and difficult to reconcile with neurocognitive theory, (2) extant neural representations do not model the compositional semantics capturing the meaning of multi-word utterances, and (3) extant mapping models fail to take into account the spatiotemporal dynamics of lexical and compositional semantic representation and computation.

In what follows, I will discuss each of these challenges in detail, and argue how they can be overcome going forward by systematically investigating how different mapping models are affected by assumptions about representational structure and the nature of the neuroimaging data they model, by grounding mapping models in linguistic and neurocognitive theory, as well as by complementing mapping models with explicit neurocomputational modeling of the spatiotemporal dynamics of language comprehension.

2. Mapping neural semantic representations

Mapping models typically assume semantic representations for individual lexical items that can be conceptualized as vectors in a high-dimensional vector space. The dimensions of these vectors may be either directly interpretable (e.g., semantic categories, componential features, co-occurrence frequencies) or only bare relative meaning to each other (Frisby et al., 2023). Given a set of words, the aim of decoding models is then to predict the activation level y of each individual semantic dimension d from, for instance, each voxel v_i of fMRI activation associated with a word:

$$y_d = \sum_{i=1}^{N} w_{di} \times v_i \quad (1)$$

where w_{di} weighs the influence of voxel v_i on dimension d. These models thus predict the full semantic representation by predicting the activation level of each semantic dimension as a weighted sum of voxel activations. After training, the fitted weights can be investigated to assess how different voxels contribute to each dimension, and brain responses to novel stimuli can be decoded by finding the most likely semantic representation given a predicted semantic vector (see Fig. 1A).

Encoding models, in turn, invert this process and aim to estimate the neural activation level y of each individual voxel v from each semantic dimension d_i associated with a word:

$$y_v = \sum_{i=1}^{N} w_{vi} \times d_i \quad (2)$$

where w_{vi} weighs the influence of dimension d_i on voxel v. Indeed, these models predict the full neural activity pattern by predicting the activation level of each voxel as a weighted sum of semantic dimension activations (see Fig. 1B). After training, the fitted weights are informative about the relative contribution of each semantic dimension to a given voxel, and the

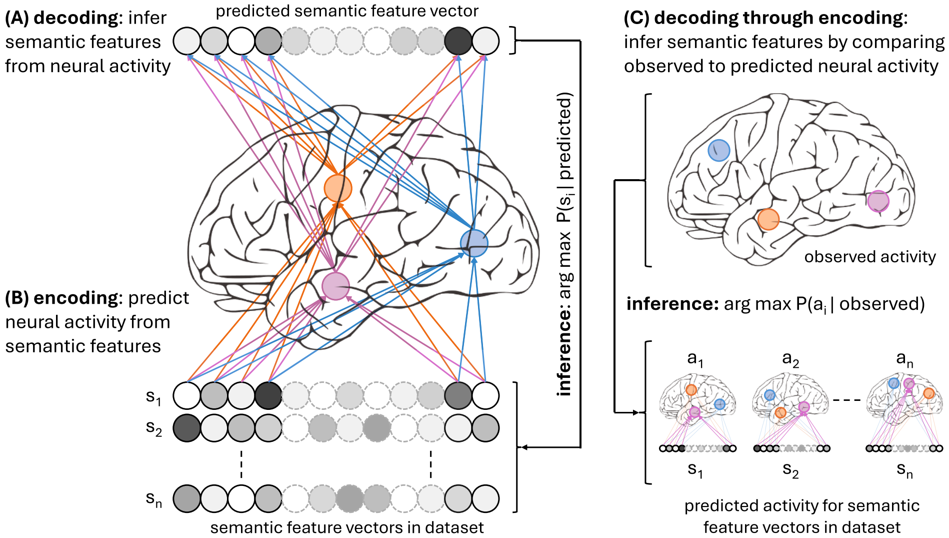
(A) decoding: infer semantic features from neural activity
predicted semantic feature vector
(B) encoding: predict neural activity from semantic features
s_1
s_2
s_n
semantic feature vectors in dataset
inference: arg max P(s_i | predicted)
(C) decoding through encoding: infer semantic features by comparing observed to predicted neural activity
observed activity
inference: arg max P(a_i | observed)
a_1
a_2
a_n
s_1
s_2
s_n
predicted activity for semantic feature vectors in dataset

Fig. 1 – Mapping models. Decoding models aim to infer semantic features from neural activity (A). Encoding models seek to predict neural activity from semantic features (B). Encoding models can also be used for decoding by predicting the neural activity for different semantic feature vectors, and inferring the most likely prediction given a pattern of observed neural activity (C).

model allows for predicting brain responses to unseen words. Decoding brain responses to novel stimuli is less straightforward using encoding models, but can be cast into an informed search problem, e.g., by predicting possible continuations of a sentence and finding the predicted brain response that is most likely given the observed brain response (e.g., Tang et al., 2023, see Fig. 1C).

2.1. *Why mapping model results are inconsistent*

While mapping models provide a powerful method to connect semantic representations to brain activity, and to ask fundamental questions about the representation and computation of meaning in the brain, Frisby et al. (2023) point out that mapping results have led to “sometimes startlingly different conclusions about the nature, structure, and organization of semantic representation” (pg. 258). One explanation for these inconsistencies is that mapping results may in fact depend on data acquisition, analysis and size, experimental design, as well as the combination of representational structure and mapping model, among other things (see Ivanova et al., 2023; Frisby et al., 2023, for discussions).

To illustrate this point, let us turn to the decoding of the *lexical semantic* (LS) representations capturing the meaning of individual words. There are essentially three different theoretical views on the neural basis of LS representations coding for conceptual structure (Frisby et al., 2023): Dimensions of LS representations either (1) encode discrete category membership for concepts (e.g., *flower*, *house*, *bird*), (2) encode the presence/absence of features (e.g., *has leaves*, *made of brick*, *is alive*), or (3) bare only relative meaning to each other. Each of these distinct hypotheses can have different neural realizations: representations may be self-contained or grounded in neural circuitry sub-serving perception and action, their dimensions may independently or conjointly code for a concept, and vary either homogeneously (adopting similar activations to code for the presence of semantic information) or heterogeneously (adopting graded activations). Furthermore, representations may locate to contiguous or dispersed cortical regions, and be consistent or inconsistent across individuals. Different combinations of these properties lead to different hypotheses about LS representations in the brain, and different mapping models may be more or less appropriate to investigate these hypotheses (Frisby et al., 2023). Mapping models harnessing a Region of Interest (ROI) approach (e.g., by focusing on theory-driven voxels of interest), for instance, assume cortically contiguous representations, while mapping models using a “searchlight” approach (e.g., by moving a spherical multivariate searchlight through a volume; Kriegeskorte et al., 2006) can detect dispersed representations. Both approaches do, however, typically assume consistent representations across subjects. Whole-brain mapping models, by contrast, are less constrained, but may consequently be less interpretable.

To complicate matters further, it remains an open question if linear or nonlinear mapping models better connect semantic representations to brain activity (see Ivanova et al., 2023). That is, mapping models typically assume a linear relationship between neural activity and semantic feature spaces that can be estimated using (regularized) linear regression. While such simple linear mapping models are most common, and are favoured over more complex nonlinear mapping models because of their apparent interpretability (of voxel/feature weights), biological plausibility (approximating downstream readout in the brain), and comparability (across feature/data sets), these desired properties may be too limiting, and may not neatly align with a linear/nonlinear model dichotomy (Ivanova et al., 2023). Harnessing the power of complex nonlinear mapping models such as deep neural networks, by contrast, may allow for more accurate modeling of how semantic feature spaces relate to recorded neural activity, for instance, by capturing nonlinear interactions between semantic features and/or accounting for nonlinearities in the recorded brain signal. However, whether such nonlinear mapping models do indeed outperform linear mapping models is subject to ongoing debate as results are inconsistent (e.g., Bertolero et al., 2020; He et al., 2020; Schulz et al., 2020), and may interact with the factors discussed above.

2.2. *Towards robust and interpretable mapping models*

The inconsistency of state-of-the-art mapping results and the entanglement of data, representational structure, and mapping model properties, call for a systematic investigation of mapping models going forward. A starting point for such an investigation, for instance, is to build upon and extend the theoretical work on LS representations by Frisby et al. (2023), and to systematically and empirically investigate if and to what degree different linear (e.g., linear/lasso/ridge regression) and nonlinear (e.g., feed-forward/convolutional deep neural network) mapping models can be used to encode and/or decode competing hypotheses about the structure of LS representations from data within and across studies (e.g., open data from Mitchell et al., 2008; Wang et al., 2022; Kaiser et al., 2022, among others). To this end, a benchmark could be developed for the automatic evaluation and introspection of different combinations of mapping model, representational hypothesis, and data set, in order to directly contrast and compare these combinations. This critically allows for contrasting different hypotheses on the same experimental data, with the same pre-processing procedure, and so forth, but also to examine how mapping models generalize across studies. To obtain a deeper understanding of the models, this comparative evaluation could be complemented with extensive feature analysis aimed at investigating what the different models are sensitive to, and what drives their performance. To achieve this for nonlinear mapping models, techniques can be employed to decompose the models into interpretable modules (Bertolero et al., 2020) or to reduce their dimensionality through multidimensional scaling approaches (Rogers et al., 2021; Venhuizen et al., 2022). Taken together, such a systematic comparison and analysis provides critical insight into how choice of mapping model, hypotheses on representational structure, and data acquisition, processing and analysis interact, and how this affects the robustness and interpretability of the results.

Indeed, arriving at robust and interpretable mapping models has the potential to provide breakthrough answers to fundamental questions on the organization and computation of meaning in the brain. The state of affairs is, however, further complicated by the coexistence of different types of neural semantic representations, and their spatiotemporal representational dynamics.

3. Lexical versus compositional neural semantic representations

Much of the work on decoding neural semantic representations has focused on LS representations: that is, on the meaning of individual words. The hallmark of human language, however, is that we can combine words into a potentially infinite number of novel sentences. A core notion in linguistic theory, therefore, is the *principle of compositionality*. While this principle has different theoretical instantiations, these are all grounded in the assumption that the meaning of a complex phrase (e.g., a sentence) is a function of the meanings of its constituent words and the way they are combined (Baggio, 2018; Martin & Baggio, 2020; Partee, 1995; Pylkkänen, 2019; Venhuizen & Brouwer, 2025b). The resultant *compositional semantic* (CS) representation of an utterance can be conceptualized as a mental model (Johnson-Laird, 1980) or situation model (Zwaan & Radvansky, 1998): a mental representation of a described state of affairs informed by and grounded in world knowledge.

3.1. *Insights from the neurocognition of language*

The idea that CS representations play a crucial role in incremental, word-by-word comprehension is also echoed in prominent neurocognitive models of language processing. While these models differ in their precise architectural assumptions, they converge upon the idea that LS representations are stored and/or activated in temporal areas, whereas other, typically frontal, areas are involved in combining/integrating these LS representations into CS representations spanning multiple words (e.g., Hagoort, 2005; Hagoort et al., 2009; Baggio & Hagoort, 2011; Brouwer & Hoeks, 2013; see Fig. 2A, but also see Lau et al., 2008; Friederici, 2011 for models implicating non-frontal regions in combinatory processing). Critically, these models make concrete predictions about the spatial and temporal organization of LS and CS representation in the brain.

To illustrate this, consider the Retrieval-Integration theory of the electrophysiology of incremental, word-by-word language comprehension (Brouwer et al., 2012, 2017; Brouwer & Hoeks, 2013; Venhuizen & Brouwer, 2025a), in which the spatiotemporal segregation of LS and CS is particularly pronounced (see Fig. 2B). This model is centered around two core processes underlying language comprehension: The retrieval of word-associated LS representations from long-term memory, and the compositional integration of these retrieved LS representations into a CS representation of unfolding utterance meaning. The process of retrieving word-associated LS representations from long-term memory, which is facilitated if these representations are anticipated by the unfolding context, is assumed to be reflected in the N400 component of the event-related potential (ERP) signal—a negative deflection peaking around 400 msec post word onset—such that more effortful retrieval results in a larger N400 (Kutas & Federmeier, 2000; Brouwer et al., 2012; Lau et al., 2008; Van Berkum, 2009). The left posterior Middle Temporal Gyrus is identified as a cortical hub mediating this retrieval process (Brouwer & Hoeks, 2013), while the LS representations are themselves assumed to be stored in a distributed manner across the association cortices (Pulvermüller, 1999; McClelland & Rogers, 2003). Retrieved LS representations are subsequently integrated into an unfolding CS representation spanning the utterance thus far. These compositional processes are assumed to be reflected in the P600 component—a positive deflection that typically reaches maximum within 600–1000 msec post word onset—such that effortful composition results in a larger P600 (Brouwer et al., 2012). The left Inferior Frontal Gyrus (lIFG) is identified as a cortical hub subserving these integrative compositional processes (Brouwer & Hoeks, 2013). Critically, the model is cyclic, in that the updated CS representation serves as context for the retrieval of upcoming word-associated LS representations and their subsequent integration into the further unfolding CS representation (see Brouwer et al., 2017, 2021, for explicit neurocomputational instantiations of the Retrieval-Integration model).

To understand how the brain represents the meaning of multi-word utterances, we should thus not only be concerned with decoding LS representations, but also consider decoding

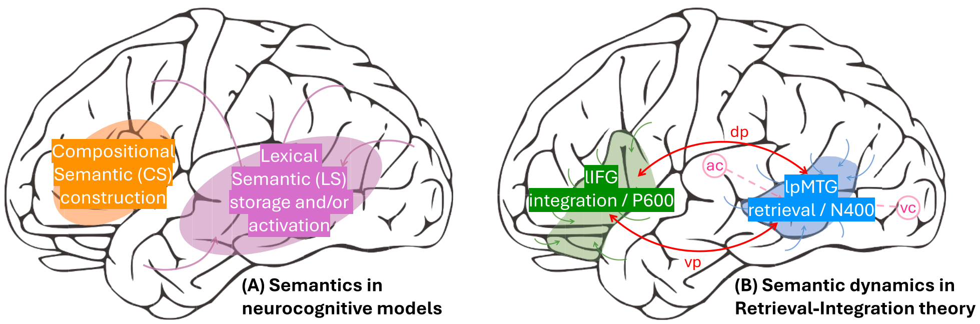
Compositional Semantic (CS) construction
Lexical Semantic (LS) storage and/or activation
(A) Semantics in neurocognitive models
dp
ac
lIFG integration / P600
lpMTG retrieval / N400
vc
vp
(B) Semantic dynamics in Retrieval-Integration theory

Fig. 2 – Neurocognitive models of language processing. Neurocognitive models of language processing assume distinct regions for the storage/activation of lexical semantics and their combination into compositional semantic representations (A). Retrieval-Integration theory makes explicit predictions about the spatiotemporal representational dynamics of lexical and compositional semantic representations (B).

CS representations. This raises the question, however, of how CS representations can be modeled in the first place; that is, while there exist concrete hypotheses about the nature of LS representation in the brain (see Frisby et al., 2023, for discussion), the neural basis of CS representation remains largely uncharted territory.

3.2. *From linguistic theory to neural CS representations*

As many decoding models assume feature-based LS representations, one approach to modeling CS representations is to examine how LS representations can be compositionally combined into the meaning of sentences. In *distributional semantics*, where LS features reflect data-derived linguistic co-occurrences, different approaches have been explored, but the most common approach to compositional sentence meaning is the order-insensitive averaging or multiplication of the constituent word meaning vectors (Mitchell & Lapata, 2010). While this approach has shown some promising results in approximating simple sentence meaning (Pereira et al., 2018), it does not achieve human-like compositional generalization, supporting the adage "good at lexical semantics, bad at composition" (Pavlick, 2022, pg. 464).

In *formal semantics*, by contrast, frameworks such as Discourse Representation Theory (DRT) fare a lot better in modeling the dynamic construction and representation of multi-word utterance meaning (Kamp & Reyle, 1993). DRT formalizes the concept of a mental model (Johnson-Laird, 1980) or situation model (Zwaan & Radvansky, 1998) in terms of Discourse Representation Structure (DRS) representations. A DRS D is a tuple $\langle U, C\rangle$, consisting of a set of referents $U = \{x_1, \ldots, x_n\}$, known as the universe, and a set of conditions on those referents $C = \{c_1, \ldots, c_n\}$, which are either simple relations $R(x_1, \ldots, x_n)$ over referents or logical combinations of sub-DRSs. The representational power of DRT is evidenced by its extensive use in the theoretical modeling of diverse aspects of meaning (Asher & Lascarides, 2003; Muskens, 1996; Van der Sandt, 1992; Venhuizen et al., 2018) as well as the automatic, large-scale derivation of DRSs using both rule-based (Bos, 2003) and neural (Van Noord et al., 2018) approaches. However, in order for these DRS *representations* to obtain an *interpretation*, they need to be embedded relative to a formal model structure M, for instance, through a translation to first-order logic. While this is theoretically well-defined, it is unclear how the representational power of DRSs can be given an interpretation at the neural level, e.g., to account for entailment and inference relations between meanings. Moreover, as DRT is a truth-theoretical theory of meaning, it does not directly capture probabilistic 'world-knowledge'-driven inferences.

Another formalism—building on neurocognitive models of story comprehension (Frank et al., 2009; Golden & Rumelhart, 1993)—that directly models interpretation at the neural level is Distributional Formal Semantics (DFS; Venhuizen et al., 2022). DFS defines meaning relative to a set of formal models $\mathbb{M}_{\mathbb{P}}$. Each model $M \in \mathbb{M}_{\mathbb{P}}$ is a first-order model that is defined relative to a set of propositions $\mathbb{P}$ and can be represented as the set of propositions that it satisfies. The full set of models $\mathbb{M}_{\mathbb{P}}$ effectively represents a set of possible worlds, in which propositions that are related to each other co-occur in many of the same worlds. Propositional and sub-propositional meaning is grounded in this (probabilistic) propositional co-occurrence space, and represented by real-valued vectors that capture meaning in terms of (fuzzy) truth values relative to these models; the meaning of a proposition $p \in \mathbb{P}$, for instance, is defined by a vector $\vec{v}(p)$ that assigns a 1 to all $M \in \mathbb{M}_{\mathbb{P}}$ that satisfy p, and a 0 otherwise. Critically, DFS representations are fully compositional in that the meaning of any logical combination of propositions of arbitrary complexity can be expressed. Beyond being fully compositional, DFS is also fully probabilistic, in that the probability of a (complex) proposition directly derives from its meaning vector (see Venhuizen et al., 2022, for more detail). While DFS has been successfully used to model compositional neural semantic representations in cognitive models of comprehension, accounting for key semantic phenomena (Venhuizen et al., 2019a, 2019b, 2022; Brouwer et al., 2021), these models are small-scale, closed world models, and their representations do not approximate the scale and broad coverage required for usage in mapping models. The main reason for this is that DFS uses propositions as atomic meaning units, and scaling up coverage entails increasing the number of propositions and their co-occurrences in $\mathbb{M}_{\mathbb{P}}$, which quickly becomes practically intractable.

Indeed, while DRT offers wide-coverage representations without neural interpretations that capture entailment and inference between meanings, DFS does offer such entailment-preserving interpretations at the neural level, but its representations have limited coverage. A first step towards deriving neural CS representations, therefore, could be to seek to combine the rich, wide-coverage representational power of DRT with the direct interpretability of neurally-instantiable DFS representations. A concrete starting point is the observation that a DRS D can be embedded relative to $\mathbb{M}_{\mathbb{P}}$ through a translation into first-order logic; that is, to the degree that $\mathbb{P}$ captures the referents and conditions of D, DRS D has an interpretation $\vec{v}(D)$ relative to $\mathbb{M}_{\mathbb{P}}$ that is expressed at the neural level. The wide-coverage of DRT as compared to DFS, however, stems from the fact that DRT allows for abstraction over referents x_i, and by extension events e_i, in the universe. To achieve this coverage in DFS, it is therefore essential to move from a set of fixed propositions $\mathbb{P}$ to a set of propositions that abstract over variables within $\mathbb{P}$, for instance by adopting neo-Davidsonian event semantics, as is common in wide-coverage DRT (Bos, 2003). The resultant, integrated framework could then allow for deriving a meaning space $\mathbb{M}_{\mathbb{P}}$ that yields neural CS representations that scale up to cover the referents and events of the training data of relevant mapping models (e.g., the podcasts used by Tang et al., 2023).

3.3. *From LLMs to neural CS representations*

A complementary approach to neural CS representation is to harness the contextualized representations of LLMs; that is, early decoding models assume LS representations at the *type*-level such that two identical *tokens* have the same meaning even if they occur in different contexts. Rather than using such static representations, more recent mapping models have turned to the use of contextualized representations from

LLMs, in which meaning is assigned to each individual *token* as a function of the context in which it occurs (e.g., Tang et al., 2023; but also see Schrimpf et al., 2021; Goldstein et al., 2022). The impressive comprehension-like behavior of LLMs suggests that these contextualized representations may instantiate a form of CS representation, and has in fact led to the suggestion that LLMs accurately model human comprehension to some degree (e.g., Goldstein et al., 2022; Piantadosi, 2023; Schrimpf et al., 2021). Critically, it has been argued that LLMs "are effective at predicting brain responses because they generally capture a wide variety of linguistic phenomena" (Antonello & Huth, 2024, p. 1). This suggests that we may elucidate CS representation in the brain by obtaining a better understanding of contextualized representations in LLMs (see also Dhar & Søgaard, 2024; Xu et al., 2024, for discussion).

Recent advances in explainability for LLMs offer a plethora of methods aimed at elucidating the behavior of these models as well as their representations (see Belinkov & Glass, 2019; Zhao et al., 2024, for reviews). One such prominent method is probing, which provides a means to assess whether the contextualized representations of LLMs code for certain linguistic properties, typically by predicting the presence or absence of these properties from the representations using a classifier (Belinkov, 2022). Manning et al. (2020), for instance, developed a probing method to show that LLM representations code for syntactic relations, and that using a linear transformation on these representations allows for the approximate reconstruction of sentence tree structures that are common in linguistic theory. Critically, such explainability techniques offer a means to investigate the degree to which the contextualized representations of LLMs code for the deep semantic features assumed by semantic theory in a top-down manner. However, as of yet, these contextualized representations "have not engaged rigorously with semantic theory" (Pavlick, 2022, p. 448), leaving them poorly understood.

Indeed, while probing techniques have been used to investigate LLM representations, as well as other word and sentence embeddings, for the encoding of surface-level features—including sentence length, the presence of a particular word, syntactic tree depth, syntactic relations, and shallow structure-driven semantics such as subject and object number (see Conneau et al., 2018; Manning et al., 2020, among others)—these investigations have yet to be extended to deep semantics. That is, in order to understand to what degree these representations do indeed capture deep semantic properties, inferences, and relationships, representations need to be probed for features that go beyond what is directly present in linguistic surface realizations. To exemplify this, consider the number of entities available for anaphoric reference in a discourse such as "Mary was talking to Bill and John. Bill needed to get up early but John decided to stay for a bit. After a while, Mary told him [...]": Indeed, while this discourse starts off introducing three available entities in the first sentence, the second sentence contains the implicature that "Bill" left, changing the state of affairs, such that only "John" is available as an antecedent for "him" in the last sentence. Human comprehenders readily register the unavailability of "Bill" in such situations (see Nieuwland et al., 2007; Venhuizen & Brouwer, 2025a). To evaluate whether LLM representations capture the unavailability of "Bill" as well, one could contrast the discourse above with one in which "Bill" remains available, and classify the number of available entities. Critically, if the representations do indeed distinguish between these constellations, independent of the order in which the entities are mentioned, this cannot be attributed to shallow surface-structural features, and would suggest that contextualized LLM representations do indeed code for deep semantic features.

Beyond investigating if representations code for such deep semantic features, it is important to understand how these features are encoded. To this end, probes instantiating specific hypotheses about how features are encoded can be contrasted (e.g., whether available antecedents are coded using a numerical *vs* a more general uniqueness feature). Such a systematic, deep semantic investigation of LLM representations will allow us to understand to what degree and how these representations encode critical aspects of CS, and to the extent that they do, ground the claim that LLMs "capture a wide variety of linguistic phenomena" (Antonello & Huth, 2024, p. 1) in linguistic theory. Indeed, this top-down investigation of CS can be seen as complementary to the bottom-up approach towards CS grounded in linguistic theory; that is, while linguistic theory informs on what deep semantic features to look for, the investigation of LLM representations may inform on potential ways in which these features can be neurally coded.

4. Spatiotemporal representational dynamics

The human language comprehension system incrementally combines LS representations into CS representations spanning multi-word utterances. This means that going forward we should be concerned with the nature of both LS and CS representations, but also with how these representations interact in the compositional process (Venhuizen & Brouwer, 2025b). One possibility is that LS and CS representations are inherently intertwined, and an approach to modeling this would be to integrate LS representations as atomic structures into a neural framework for CS (Asher et al., 2016; Beltagy et al., 2016). Alternatively, composition could be an emergent property, such as in neurocomputational models that learn to incrementally map sequences of word-associated LS representations into an unfolding CS representation spanning a multi-word utterance (Brouwer et al., 2017, 2021). Critically, these questions have important implications for mapping models: As LS representations form the building blocks for CS representations, these two types of representation may be entangled in space and time, meaning that the neural activity recorded in a particular cortical region, at a particular point in time, may capture aspects of both LS and CS representations. This implies that beyond being concerned with the nature of both LS and CS representations, we must address the spatiotemporal dynamics of their interaction in the compositional process.

Mappings between semantic feature spaces and cortical feature maps have identified voxels scattered across many cortical areas (Huth et al., 2012, 2016; Pereira et al., 2018; Tang et al., 2023). Tang et al. (2023), for instance, showed that continuous language could be accurately decoded from the classical language network, the parietal-temporal-occipital association network, and the prefrontal network in each hemisphere. It is not clear how this apparent scattered organization of neural semantic representation can be reconciled with neurocognitive models of language comprehension that assume—informed by neuroimaging but also clinical and lesion studies—a more left-lateralized and focused functional organization in terms of a spatiotemporal segregation of LS representation storage/activation and their combination into CS representations spanning multiple words (Hagoort, 2005; Baggio & Hagoort, 2011; Ben Shalom & Poeppel, 2008; Brouwer & Hoeks, 2013; Friederici, 2011; Lau et al., 2008; Pylkkänen, 2019). To address this apparent inconsistency going forward, mapping models should be guided by, and their results grounded in, neurocognitive theory in terms of the spatiotemporal dynamics of LS and CS (see Goldstein et al., 2023, for a different approach using LLMs).

4.1. *Towards mapping models grounded in neurocognitive theory*

Retrieval-Integration theory makes explicit predictions about the spatiotemporal dynamics of LS and CS representation (Brouwer et al., 2012; Brouwer and Hoeks, 2013; Venhuizen and Brouwer, 2025a). Spatially, the model assumes the lpMTG to serve as a hub for the retrieval of LS representations that are stored across the association cortices. This model would thus predict that the lpMTG is sensitive to LS in general, but that different association cortices may show modality- or sensory-specific sensitivity to LS (e.g., visual *vs* motor concepts; see Pulvermüller, 1999). The lIFG, in turn, is predicted to be predominantly sensitive to CS. This prediction about spatial organization can be tested by complementing extant LS mapping results with novel mapping models targeting CS representation to decode or encode high spatial resolution neuroimaging data such as fMRI responses.

Beyond this spatial differentiation, the Retrieval-Integration model also predicts that the decodability of LS and CS representations may fluctuate over time post word onset, such that during retrieval, LS decodability is dominant, while during integrative compositional processing, CS decodability is dominant. Predictions such as these can be tested by using mapping models to investigate the difference in generalization across time of LS and CS representation decoding (or classification) post word onset (see Heikel et al., 2018; King & Dehaene, 2014). The idea behind this temporal generalization method is simple, yet powerful: assume a mapping model that is trained to succesfully decode a semantic representation at a timestep *t* from a neuroimaging signal with high temporal resolution, such as electroencephalography (EEG) or electrocorticography (ECoG) data. The decoding accuracy of this classifier can then be tested on all other time-steps (both backward: $t-2, t-1, \ldots$, and forward: $t+1, t+2, \ldots$), to see how well its decoding performance generalizes. If it does indeed generalize to time-steps other than *t*, this is taken to indicate that the mental representations at *t* also occur at those other time-steps (King & Dehaene, 2014). Testing each mapping model at each other time-step yields a Generalization Across Time (GAT) matrix, which reveals how different mental representations unfold in time, and to what degree they temporally overlap with each other. Indeed, if the prediction from the Retrieval-Integration model is right, for instance, relative decoding/classification accuracy should peak for LS earlier (e.g., in the 300–500 msec N400 time window) and for CS later (e.g., in the 600–1000 msec P600 time window), while there may also be a temporal window of overlap, where both LS and CS representations can be decoded with a certain degree of accuracy.

4.2. *Synthesis through neurocomputational modeling*

As commonly used neuroimaging methods have either high spatial or high temporal resolution, but typically not both, a further challenge is to integrate insights on the spatial organization with those on the temporal dynamics. One fruitful approach to synthesizing findings on the cortical organisation with those on the temporal dynamics is by complementing and connecting the mapping results with neurocomputational modeling of language comprehension that explicates the spatiotemporal dynamics of LS and CS. Beyond providing a formally precise instantiation of where, when, and how LS representations combine into an unfolding CS representation, the spatiotemporal dynamics of these models can be directly compared to those of the human comprehension system through temporal generalization of representation decoding or classification (see Rogers et al., 2021, for such an approach to LS).

To illustrate this approach, consider the explicit neurocomputational instantiation of RI theory, which has been shown to account for key semantic processing phenomena, and which is explicit about the cortical organisation of LS and CS (Brouwer et al., 2017, 2021). This neurocomputational model is a recurrent neural network model (see Fig. 3A) that processes sentences on a word-by-word basis. The model is built around the lpMTG and lIFG (see Fig. 3B) and produces N400 and P600 estimates at each word (Fig. 3C). Each acoustically (ac; auditory cortex) or orthographically (vc; visual cortex) perceived word enters the model through the ac/vc layer representing the perceived word form. This word form representation is then projected through the lpMTG, which retrieves an LS representation, while taking into account the model-internal CS representation residing in the lIFG. In the model, N400 amplitude (reflecting retrieval difficulty) is estimated as the degree to which the activity pattern in the lpMTG changes as a function of the incoming word, such that N400 amplitude is attenuated when the word-associated LS is more anticipated, and the other way around. In the model, an LS representation is a ‘vector space’-based conceptual vector. This retrieved LS representation is then projected to the lIFG, where it is integrated into the unfolding CS representation, and projected onto an interpretable CS representation, for instance, in form of a DFS vector (Brouwer et al., 2021). P600 amplitude (reflecting integration difficulty) is estimated as the degree to which the activation layer changes from one word to the next, such that larger P600 amplitude ensues for more substantial updates to the unfolding CS representation.

While the neurocomputational instantiation of the Retrieval-Integration model is explicit about the spatial organization of LS and CS (Brouwer et al., 2017, 2021), time in its

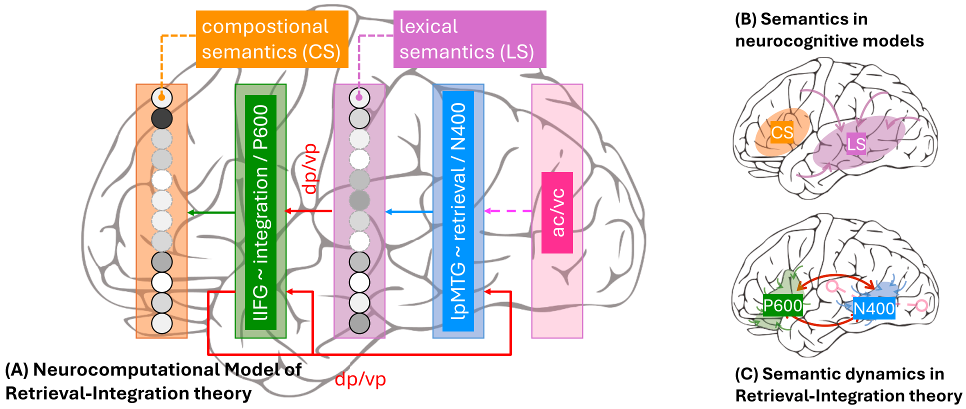
compostional semantics (CS)
lexical semantics (LS)
lIFG ~ integration / P600
lpMTG ~ retrieval / N400
ac/vc
dp/vp
dp/vp
(A) Neurocomputational Model of Retrieval-Integration theory
(B) Semantics in neurocognitive models
CS
LS
P600
N400
(C) Semantic dynamics in Retrieval-Integration theory

Fig. 3 – Synthesis through neurocomputational modeling. The neurocomputational model of Retrieval-Integration theory (A) provides a starting point for integrating the spatial organisation of lexical and compositional neural semantic representations (B) with their temporal dynamics (C).

current instantiation is discrete, and hence there are no temporal dynamics; that is, when processing an incoming word, activation flows through the model for a single discrete time tick. A critical adjustment to the model, therefore, would be to distribute word processing over multiple time ticks in order to investigate spatiotemporal dynamics of LS and CS representation (see Brouwer et al., 2017, for a technical discussion). Indeed, this would allow for modeling the spatiotemporally overlapping dynamics of the N400 and P600 underlying the observed EEG signal (Brouwer & Crocker, 2017), the temporally reverberating dynamics of frontal and temporal regions during sentence processing (e.g., Tse et al., 2007), as well as the relationship between these temporal and spatial LS and CS dynamics. Informed by mapping model results, neurocomputational modeling thus offers a means to mechanistically explicate the spatiotemporal representational dynamics of LS and CS in language comprehension, paving way towards an integrated theory of language in the brain.

5. Conclusion

The question of how the brain represents the meaning of concepts, phrases, and sentences, is a central question in cognitive science. Recent advances in neuroscience and artificial intelligence have led to the development of innovative mapping models that enable the decoding of neural semantic representations from neural activity, or conversely, the prediction of neural activity from neural semantic representations. These mapping models hold the promise to be instrumental in unraveling the organization, representation, and computation of meaning in the brain. However, I argue that in order to fulfil this promise, we need to first systematically investigate how assumptions about representational structure and the nature of neuroimaging data affect mapping results. Furthermore, mapping models should be developed that do not only capture the lexical semantics of individual words, but also the compositional semantic representations spanning multi-word utterances. While lexical semantic representations are relatively well-understood at the neural level, this is currently not the case for compositional semantics. This critically requires the development of a framework for compositional semantic representations at the neural level, and I have discussed how linguistic theory may be harnessed to arrive at such a framework. Finally, neurocognitive theories of language comprehension predict differential spatiotemporal representational dynamics for lexical and compositional semantics. Hence, in order to better understand mapping results, they should be grounded in neurocognitive theory, and complemented by, and connected to neurocomputational modeling of language comprehension, to explicate the spatiotemporal dynamics of lexical and compositional semantics. In the future, such theoretically-informed and grounded mapping models will significantly advance contemporary neurocognitive theory on meaning in the brain's language.

REFERENCES

Anderson, A. J., Binder, J. R., Fernandino, L., Humphries, C. J., Conant, L. L., Aguilar, M., Wang, X., Doko, D., & Raizada, R. D. (2017). Predicting neural activity patterns associated with sentences using a neurobiologically motivated model of semantic representation. *Cerebral Cortex, 27*(9), 4379–4395.

Antonello, R., & Huth, A. (2024). Predictive coding or just feature discovery? An alternative account of why language models fit brain data. *Neurobiology of Language, 5*(1), 64–79.

Asher, N., & Lascarides, A. (2003). *Logics of conversation*. Cambridge University Press.

Asher, N., Van de Cruys, T., Bride, A., & Abrusán, M. (2016). Integrating type theory and distributional semantics: A case study on adjective–noun compositions. *Computational Linguistics, 42*(4), 703–725.

Baggio, G. (2018). *Meaning in the brain*. MIT Press.

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

Belinkov, Y. (2022). Probing classifiers: Promises, shortcomings, and advances. *Computational Linguistics, 48*(1), 207–219.

Belinkov, Y., & Glass, J. (2019). Analysis methods in neural language processing: A survey. *Transactions of the Association for Computational Linguistics, 7*, 49–72.

Beltagy, I., Roller, S., Cheng, P., Erk, K., & Mooney, R. J. (2016). Representing meaning with a combination of logical and distributional models. *Computational Linguistics, 42*(4), 763–808.

Ben Shalom, D., & Poeppel, D. (2008). Functional anatomic models of language: Assembling the pieces. *The Neuroscientist, 14*(1), 119–127.

Bertolero, M. A., Moraczewski, D., Thomas, A., & Bassett, D. S. (2020). *Deep neural networks carve the brain at its joints*. arXiv preprint. arXiv:2002.08891.

Bos, J. (2003). Implementing the binding and accommodation theory for anaphora resolution and presupposition projection. *Computational Linguistics, 29*(2), 179–210.

Brouwer, H., & Crocker, M. W. (2017). On the proper treatment of the N400 and P600 in language comprehension. *Frontiers in Psychology, 8*, 1327.

Brouwer, H., Crocker, M. W., Venhuizen, N. J., & Hoeks, J. C. J. (2017). A neurocomputational model of the N400 and the P600 in language processing. *Cognitive Science, 41*, 1318–1352.

Brouwer, H., Delogu, F., Venhuizen, N. J., & Crocker, M. W. (2021). Neurobehavioral correlates of surprisal in language comprehension: A neurocomputational model. *Frontiers in Psychology, 12*, Article 615538.

Brouwer, H., Fitz, H., & Hoeks, J. C. J. (2012). Getting real about semantic illusions: Rethinking the functional role of the P600 in language comprehension. *Brain Research, 1446*, 127–143.

Brouwer, H., & Hoeks, J. C. J. (2013). A time and place for language comprehension: Mapping the N400 and the P600 to a minimal cortical network. *Frontiers in Human Neuroscience, 7*, 758.

Conneau, A., Kruszewski, G., Lample, G., Barrault, L., & Baroni, M. (2018). What you can cram into a single vector: Probing sentence embeddings for linguistic properties. In I. Gurevych, & Y. Miyao (Eds.), *Proceedings of the 56th annual meeting of the association for computational linguistics (Volume 1: Long papers)* (pp. 2126–2136). Melbourne, Australia: Association for Computational Linguistics.

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*. arXiv preprint. arXiv:1810.04805.

Dhar, R., & Søgaard, A. (2024). *From words to worlds: compositionality for cognitive architectures*. arXiv preprint. arXiv:2407.13419.

Frank, S. L., Haselager, W. F., & van Rooij, I. (2009). Connectionist semantic systematicity. *Cognition, 110*(3), 358–379.

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

Frisby, S. L., Halai, A. D., Cox, C. R., Ralph, M. A. L., & Rogers, T. T. (2023). Decoding semantic representations in mind and brain. *Trends in Cognitive Sciences, 27*(3), 258–281.

Golden, R. M., & Rumelhart, D. E. (1993). A parallel distributed processing model of story comprehension and recall. *Discourse Processes, 16*(3), 203–237.

Goldstein, A., Ham, E., Schain, M., Nastase, S., Zada, Z., Dabush, A., Aubrey, B., Gazula, H., Feder, A., Aubrey, B., Gazula, H., Feder, A., Doyle, W. K., Devore, S., Dugan, P., Friedman, D., Reichart, R., Brenner, M., Hassidim, A., Devinsky, O., Flinker, A., Levy, O., et al. (2023). *The temporal structure of language processing in the human brain corresponds to the layered hierarchy of deep language models*. arXiv preprint. arXiv:2310.07106.

Goldstein, A., Zada, Z., Buchnik, E., Schain, M., Price, A., Aubrey, B., Nastase, S. A., Feder, A., Emanuel, D., Nastase, S. A., Feder, A., Emanuel, D., Cohen, A., Jansen, A., Gazula, H., Choe, G., Rao, A., Kim, C., Casto, C., Fanda, L., Doyle, W., Friedman, D., et al. (2022). Shared computational principles for language processing in humans and deep language models. *Nature Neuroscience, 25*(3), 369–380.

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

Hagoort, P., Baggio, G., & Willems, R. M. (2009). Semantic unification. In *The cognitive neurosciences* (4th ed., pp. 819–836). MIT press.

He, T., Kong, R., Holmes, A. J., Nguyen, M., Sabuncu, M. R., Eickhoff, S. B., Bzdok, D., Feng, J., & Yeo, B. T. (2020). Deep neural networks and kernel regression achieve comparable accuracies for functional connectivity prediction of behavior and demographics. *Neuroimage, 206*, Article 116276.

Heikel, E., Sassenhagen, J., & Fiebach, C. J. (2018). Time-generalized multivariate analysis of EEG responses reveals a cascading architecture of semantic mismatch processing. *Brain and Language, 184*, 43–53.

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

Huth, A. G., Nishimoto, S., Vu, A. T., & Gallant, J. L. (2012). A continuous semantic space describes the representation of thousands of object and action categories across the human brain. *Neuron, 76*(6), 1210–1224.

Ivanova, A. A., Schrimpf, M., Anzellotti, S., Zaslavsky, N., Fedorenko, E., & Isik, L. (2023). *Beyond linear regression: Mapping models in cognitive neuroscience should align with research goals*. arXiv preprint. arXiv:2208.10668.

Johnson-Laird, P. N. (1980). Mental models in cognitive science. *Cognitive Science, 4*(1), 71–115.

Kaiser, D., Jacobs, A. M., & Cichy, R. M. (2022). Modelling brain representations of abstract concepts. *Plos Computational Biology, 18*(2), Article e1009837.

Kamp, H., & Reyle, U. (1993). *From discourse to logic: Introduction to modeltheoretic semantics of natural language, formal logic and discourse representation theory*. Dordrecht: Kluwer.

King, J.-R., & Dehaene, S. (2014). Characterizing the dynamics of mental representations: The temporal generalization method. *Trends in Cognitive Sciences, 18*(4), 203–210.

Kriegeskorte, N., Goebel, R., & Bandettini, P. (2006). Information-based functional brain mapping. *Proceedings of the National Academy of Sciences, 103*(10), 3863–3868.

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

Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (de)constructing the N400. *Nature Reviews Neuroscience, 9*(12), 920–933.

Manning, C. D., Clark, K., Hewitt, J., Khandelwal, U., & Levy, O. (2020). Emergent linguistic structure in artificial neural networks trained by self-supervision. *Proceedings of the National Academy of Sciences, 117*(48), 30046–30054.

Martin, A. E., & Baggio, G. (2020). Modelling meaning composition from formalism to mechanism. *Philosophical Transactions of the Royal Society B, 375*(1791), Article 20190298.

McClelland, J. L., & Rogers, T. T. (2003). The parallel distributed processing approach to semantic cognition. *Nature Reviews Neuroscience, 4*(4), 310–322.

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

Mitchell, T. M., Shinkareva, S. V., Carlson, A., Chang, K.-M., Malave, V. L., Mason, R. A., & Just, M. A. (2008). Predicting human brain activity associated with the meanings of nouns. *Science, 320*(5880), 1191–1195.

Muskens, R. (1996). Combining Montague semantics and discourse representation. *Linguistics and Philosophy, 19*(2), 143–186.

Nieuwland, M. S., Otten, M., & van Berkum, J. J. (2007). Who are you talking about? Tracking discourse-level referential processing with event-related brain potentials. *Journal of Cognitive Neuroscience, 19*(2), 228–236.

Partee, B. H. (1995). Lexical semantics and compositionality. *An Invitation to Cognitive Science: Language, 1*, 311–360.

Pavlick, E. (2022). Semantic structure in deep learning. *Annual Review of Linguistics, 8*, 447–471.

Pereira, F., Lou, B., Pritchett, B., Ritter, S., Gershman, S. J., Kanwisher, N., Botvinick, M., & Fedorenko, E. (2018). Toward a universal decoder of linguistic meaning from brain activation. *Nature Communications, 9*(1), 963.

Piantadosi, S. T. (2023). Modern language models refute Chomsky's approach to language. *From fieldwork to linguistic theory: A tribute to Dan Everett*, 353–414.

Poldrack, R. A. (2011). Inferring mental states from neuroimaging data: From reverse inference to large-scale decoding. *Neuron, 72*(5), 692–697.

Pulvermüller, F. (1999). Words in the brain's language. *Behavioral and Brain Sciences, 22*(2), 253–279.

Pylkkänen, L. (2019). The neural basis of combinatory syntax and semantics. *Science, 366*(6461), 62–66.

Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2018). *Improving language understanding by generative pre-training.* OpenAI.

Rogers, T. T., Cox, C. R., Lu, Q., Shimotake, A., Kikuchi, T., Kunieda, T., Miyamoto, S., Takahashi, R., Ikeda, A., Miyamoto, S., Takahashi, R., Ikeda, A., Matsumoto, R., & Lambon Ralph, M. A. (2021). Evidence for a deep, distributed and dynamic code for animacy in human ventral anterior temporal cortex. *eLife, 10*, Article e66276.

Schrimpf, M., Blank, I. A., Tuckute, G., Kauf, C., Hosseini, E. A., Kanwisher, N., Tenenbaum, J. B., & Fedorenko, E. (2021). The neural architecture of language: Integrative modeling converges on predictive processing. *Proceedings of the National Academy of Sciences, 118*(45), Article e2105646118.

Schulz, M.-A., Yeo, B. T., Vogelstein, J. T., Mourao-Miranada, J., Kather, J. N., Kording, K., Richards, B., & Bzdok, D. (2020). Different scaling of linear models and deep learning in UKBiobank brain images versus machine-learning datasets. *Nature Communications, 11*(1), 4238.

Tang, J., LeBel, A., Jain, S., & Huth, A. G. (2023). Semantic reconstruction of continuous language from non-invasive brain recordings. *Nature Neuroscience, 26*(5), 858–866.

Tse, C.-Y., Lee, C.-L., Sullivan, J., Garnsey, S. M., Dell, G. S., Fabiani, M., & Gratton, G. (2007). Imaging cortical dynamics of language processing with the event-related optical signal. *Proceedings of the National Academy of Sciences, 104*(43), 17157–17162.

Van Berkum, J. J. (2009). The neuropragmatics of 'simple' utterance comprehension: An ERP review. In *Semantics and pragmatics: From experiment to theory* (pp. 276–316). Palgrave Macmillan.

Van der Sandt, R. A. (1992). Presupposition projection as anaphora resolution. *Journal of Semantics, 9*(4), 333–377.

Van Noord, R., Abzianidze, L., Toral, A., & Bos, J. (2018). Exploring neural methods for parsing discourse representation structures. *Transactions of the Association for Computational Linguistics, 6*, 619–633.

Venhuizen, N. J., Bos, J., Hendriks, P., & Brouwer, H. (2018). Discourse semantics with information structure. *Journal of Semantics, 35*(1), 127–169.

Venhuizen, N. J., & Brouwer, H. (2025a). Referential retrieval and integration in language comprehension: An electrophysiological perspective. *Psychological Review*.

Venhuizen, N. J., & Brouwer, H. (2025b). Two models of meaning: Revisiting the principle of compositionality from the neurocognition of language. In *Psychology of learning and motivation*. Elsevier.

Venhuizen, N. J., Crocker, M. W., & Brouwer, H. (2019a). Expectation-based comprehension: Modeling the interaction of world knowledge and linguistic experience. *Discourse Processes, 56*(3), 229–255.

Venhuizen, N. J., Crocker, M. W., & Brouwer, H. (2019b). Semantic entropy in language comprehension. *Entropy, 21*(12), 1159.

Venhuizen, N. J., Hendriks, P., Crocker, M. W., & Brouwer, H. (2022). Distributional formal semantics. *Information and Computation, 287*, Article 104763.

Wang, S., Zhang, Y., Zhang, X., Sun, J., Lin, N., Zhang, J., & Zong, C. (2022). An fMRI dataset for concept representation with semantic feature annotations. *Scientific Data, 9*(1), 721.

Xu, Z., Shi, Z., & Liang, Y. (2024). *Do large language models have compositional ability? An investigation into limitations and scalability*. arXiv preprint. arXiv:2407.15720.

Zhao, H., Chen, H., Yang, F., Liu, N., Deng, H., Cai, H., Wang, S., Yin, D., & Du, M. (2024). Explainability for large language models: A survey. *ACM Transactions on Intelligent Systems and Technology, 15*(2), 1–38.

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin, 123*(2), 162.