Neurobiology of Language

Controlled Semantic Cognition: Precision Recordings converge with in silico Experiments to Reveal the Inner Workings of the Anterior Temporal Lobe Hub

--Manuscript Draft--

Manuscript Number:	NOL-25-00023R2
Article Type:	Review
Full Title:	Controlled Semantic Cognition: Precision Recordings converge with in silico Experiments to Reveal the Inner Workings of the Anterior Temporal Lobe Hub
Short Title:	In vivo and in silico approaches to semantics converge
Abstract:	The anterior temporal lobe (ATL) is crucial for learning and storing concepts, yet the inner workings of its multiple cytoarchitectonically distinct subregions remain a ‘black box’. Moreover, it is not yet clear how this region interacts with a distributed network of brain regions to access context-specific information. However, two recent papers have made crucial steps forward, while adopting radically different approaches. One (Jackson et al., 2021) used in silico experimentation to ask how a system can meet the core requirements of controlled semantic cognition, comparing multiple theoretical architectures on their ability to meet these demands. The other (Tiesinga et al., 2023) performed in vivo intracortical recordings with unprecedented coverage and spatial precision, comparing cytoarchitectonic ATL subregions. In this perspective, we bridge these recent in silico and in vivo explorations of controlled semantic cognition, demonstrating their remarkable convergence. Thereby, we expose the inner workings of the ATL semantic hub and the regulating effect of frontal regions, while highlighting the crucial next challenges. The result is a neuroanatomically-precise mechanistic model of controlled semantic cognition.
Corresponding Author:	Rebecca L Jackson University of York York, England UNITED KINGDOM
Other Authors:	Guy Orban
	Paul Tiesinga
Keywords:	Semantic cognition; computational modelling; intracortical EEG; semantic control; Granger causality; functional connectivity.
Manuscript Classifications:	Lexical and Combinatorial Semantics

Controlled Semantic Cognition: Precision Recordings converge with *in silico* Experiments to Reveal the Inner Workings of the Anterior Temporal Lobe Hub

Rebecca L. Jackson[1]*, Guy Orban[2]*, Paul Tiesinga[3]*

1 Department of Psychology & York Biomedical Research Institute, University of York, York, United Kingdom

2 Department of Medicine and Surgery, University of Parma, Parma, Italy

3 Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, The Netherlands

*These authors contributed equally.

Corresponding author: rebecca.jackson@york.ac.uk

Abbreviated title: In vivo and in silico approaches to semantics converge

Conflict of Interest

Authors report no conflict of interest.

Funding

This work was supported in part by the National Institute of Mental Health of the National Institutes of Health under Award Number R01MH123687 (PT).

Keywords

Semantic cognition; computational modelling; intracortical EEG; semantic control; Granger causality; functional connectivity.

Controlled Semantic Cognition: Precision Recordings converge with *in silico* Experiments to Reveal the Inner Workings of the Anterior Temporal Lobe Hub

Rebecca L. Jackson[1]*, Guy Orban[2]*, Paul Tiesinga[3]*

1 Department of Psychology & York Biomedical Research Institute, University of York, York, United Kingdom

2 Department of Medicine and Surgery, University of Parma, Parma, Italy

3 Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, The Netherlands

*These authors contributed equally.

Corresponding author: rebecca.jackson@york.ac.uk

Orcid: 0000-0001-5339-0266 (Jackson), 0000-0002-8179-9584 (Orban), 0000-0003-4509-6870 (Tiesinga)

Abbreviated title: In vivo and in silico approaches to semantics converge

Conflict of Interest

Authors report no conflict of interest.

Funding

This work was supported in part by the National Institute of Mental Health of the National Institutes of Health under Award Number R01MH123687 (PT).

Keywords

Semantic cognition; computational modelling; intracortical EEG; semantic control; Granger causality; functional connectivity.

Abstract

The anterior temporal lobe (ATL) is crucial for learning and storing concepts, yet the inner workings of its multiple cytoarchitectonically distinct subregions remain a ‘black box’. Moreover, it is not yet clear how this region interacts with a distributed network of brain regions to access context-specific information. However, two recent papers have made crucial steps forward, while adopting radically different approaches. One (Jackson et al., 2021) used in silico experimentation to ask how a system can meet the core requirements of controlled semantic cognition, comparing multiple theoretical architectures on their ability to meet these demands. The other (Tiesinga et al., 2023) performed in vivo intracortical recordings with unprecedented coverage and spatial precision, comparing cytoarchitectonic ATL subregions. In this perspective, we bridge these recent in silico and in vivo explorations of controlled semantic cognition, demonstrating their remarkable convergence. Thereby, we expose the inner workings of the ATL semantic hub and the regulating effect of frontal regions, while highlighting the crucial next challenges. The result is a neuroanatomically-precise mechanistic model of controlled semantic cognition.

Keywords: semantic cognition; computational modelling; intracortical EEG; semantic control; Granger causality; functional connectivity.

1 **Introduction**

Semantic cognition, the acquisition and controlled use of multimodal conceptual knowledge, underpins our everyday interactions with people and our environment (Patterson *et al.*, 2007; Lambon Ralph *et al.*, 2017). The brain combines information across multiple senses, moving beyond superficial perceptual similarities to uncover the conceptual structure necessary to drive meaningful behaviour (Rogers &

McClelland, 2004). For instance, using shape information in isolation, a pear would appear more similar to a lightbulb than an orange, resulting in an inappropriate behavioural response. To respond similarly to different fruits, we must also factor in somatosensory, gustatory and olfactory features. In semantic dementia (SD; also referred to as semantic variant primary progressive aphasia (Gorno-Tempini *et al.*, 2011)), these multimodal semantic representations are gradually lost as the bilateral anterior temporal lobes (ATL) progressively deteriorate (Warrington, 1975; Snowden *et al.*, 1989; Patterson *et al.*, 2007). This results in a profound impairment of verbal and nonverbal comprehension as patients lose their ability to communicate (e.g., not being able to name an orange) and interact in an informed fashion with their environment (e.g., knowing to peel the orange before eating it). Thus, the bilateral ATL was identified as the area responsible for bringing together information acquired in different senses. As a direct consequence of learning to map between sensory-specific ‘spoke’ areas, multimodal concept representations emerge in this ‘hub’ region (Rogers *et al.*, 2004; Patterson *et al.*, 2007; Lambon Ralph *et al.*, 2017). This is known as the hub and spoke theory of semantics. Subsequently, a broad range of neuroimaging and neurostimulation approaches have confirmed the importance of the ATL in multimodal semantic representation (Devlin *et al.*, 2000; Marinkovic *et al.*, 2003; Pobric *et al.*, 2007; Visser *et al.*, 2010b; Visser & Lambon Ralph, 2011; Binney *et al.*, 2012 ; Abel *et al.*, 2015; Shimotake *et al.*, 2015; Chen *et al.*, 2016; Jackson *et al.*, 2016). The relative engagement of the left and right ATL varies somewhat based on the modality of input (e.g., written words elicit more left ATL activity (Rice *et al.*, 2015)) and features of the concept (e.g., social concepts often elicit more right ATL activity (Olson *et al.*, 2007; Zahn *et al.*, 2007; Rice *et al.*, 2015)). However, both are engaged in semantic judgements across all concepts (Abel *et al.*, 2014; Rice *et al.*, 2015; Rice *et al.*, 2018; Platonov *et al.*, 2019), their connectivity is crucial (Jung & Ralph, 2016; Jung & Lambon Ralph, 2023) and only bilateral damage has severe consequences (Lambon Ralph *et al.*, 2010; Lambon Ralph *et al.*, 2012). They may therefore be considered two parts of one system which work in concert to represent multimodal semantic concepts (Rice *et al.*, 2018). Although several models have been proposed for the semantic system and much debate remains, the evidence increasingly supports both the importance of spokes (Meteyard *et al.*, 2012; Pulvermüller, 2013; Fernandino *et al.*, 2016) and the critical role of the ATL hub, which is now well accepted (Binder *et al.*, 2011; Lambon Ralph *et al.*, 2017; Kuhnke *et al.*, 2023; Fernandino & Binder, 2024). Additional brain areas respond to multimodal semantic concepts, yet damage to these regions does not result in the loss of semantic representations suggesting they fulfill distinct functional roles (Jefferies & Lambon Ralph, 2006; Humphreys & Ralph, 2015; Humphreys *et al.*, 2022). While the hub and spoke regions extract and represent stable concepts, additional regions are required to select the conceptual information relevant to a particular context or task (Jefferies, 2013; Noonan *et al.*, 2013; Davey *et al.*, 2015; Jackson, 2021; Reilly *et al.*, 2024). These ‘semantic control’ processes are underpinned by a left-lateralised network focused on inferior frontal and posterior temporal cortices (Jefferies & Lambon Ralph, 2006; Jefferies, 2013; Noonan *et al.*, 2013; Davey *et al.*, 2015; Thompson *et al.*, 2015; Jackson, 2021).

The brain learns multimodal semantic concepts as a consequence of mapping between features in different sensory modalities. Artificial neural networks (ANNs) also uncover relationships in their (training) environment by learning to map between features. Consequently these models can simulate how semantic concepts are learnt. Just like the cortical semantic system, ANNs can extract the higher-order structure of their environment as they learn to map from provided (‘input’) to required (‘output’) features. In a model of the semantic system, these input and output features may be presented to discrete regions (layers) reflecting different (real or artificial) senses. This model could then be assessed on its ability to accurately learn the multimodal conceptual structure necessary for semantically-driven behaviour (e.g., representing pears similarly to oranges) based on the relationships between features experienced across multiple sensory modalities and events (Saxe *et al.*, 2019; Cao *et al.*, 2020; Jackson *et al.*, 2021). This ANN could have a wide range of architectures (configurations of units and their connections) which may be i) better or worse at uncovering this underlying conceptual structure and ii) further from, or closer to, the way the brain solves this difficult problem. As such, constructing a single model may not accurately demonstrate how the cortical semantic system works. A powerful alternative is to systematically manipulate architectural features across multiple ANNs and measure the impact on performance. One can then examine whether this ‘reverse-engineered’ solution utilises a similar architecture to the cortical system.

2 Reverse-engineering the semantic network

In recent work, Jackson and colleagues ((Jackson *et al.*, 2021); hereafter ‘J21’), adopted this reverse-engineering approach to determine how a semantic system should be organised. Recurrent ANNs (ANNs with both feedforward and feedback connections) were constructed with seven different architectures, reflecting different theoretical perspectives on the cortical semantic system (Eggert & Wernicke, 1977; Damasio *et al.*, 1996; Martin & Chao, 2001; Patterson *et al.*, 2007; Binder & Desai, 2011; Lambon Ralph *et al.*, 2017). This included simulations with unimodal processing regions only, with a multimodal hub bridging all sensory modalities, or with multiple bimodal ‘hubs’ bridging between pairs of modalities (or a combination of these). In addition, the architectures varied in depth (whether the processing units formed one or two layers) and the presence of ‘shortcut’ connections; a small number of direct connections (in either direction) between the sensory input/output areas and the multimodal hub, which therefore skip the intervening layer. While conceptually similar to ‘skip connections’ in deep ANNs (He *et al.*, 2016), these shortcut connections do not project all information contained in one layer to a later layer, but simply connect pairs of units via learned weights (i.e., they are the same as any other connection in the model). They may therefore have greater biological plausibility as well as different functional consequences. All models had an equal number of resources (units and weights) and were trained until they could successfully associate the features of each concept across all possible combinations of three artificial sensory modalities, resulting in 9 different ‘tasks’ (see Figure 1A, top). After learning, the models were compared on how well they identified the full multimodal conceptual structure present in their environment, utilising all sensory features presented across multiple modalities and task contexts and not merely those active in the current task context. This is comparable to asking whether presenting a picture of a pear results in a similar activity pattern to an orange, and not just a lightbulb, using representational similarity analysis (Kriegeskorte *et al.*, 2008; Kriegeskorte & Kievit, 2013).

These *in silico* experiments highlighted three architectural features as beneficial: the presence of a single multimodal hub, a deep network and shortcut connections. The ‘winning’ architecture is shown in Figure 1A. Both J21 and Rogers et al., (2004) demonstrate that a single area responsible for mapping between the information in each sensory area forms multimodal concept representations, whilst separate connections between the various sensory regions cannot. Moreover, J21 demonstrated that this requirement for one multimodal representation area cannot be solved by different areas combining information from each pair of senses. Indeed, these bimodal regions cannot even be present in combination with the multimodal hub, as this obviates the need to utilise it for each mapping. This mechanistic account demonstrates why both unimodal sensory areas and a single multimodal hub, located in the ATL, are required for healthy semantic cognition. Additionally, J21 extended the hub and spoke theory by demonstrating the importance of depth (also see (Chen *et al.*, 2017)), which allows the system to perform multiple, sequential transformations on the input, allowing different layers to represent information at different levels of abstraction. This means shallow layers can extract the structure within a sensory modality and deeper layers can identify relationships across modalities. Furthermore, J21 highlighted the dramatic effects of a small number of direct, shortcut connections between sensory areas and the multimodal hub, a novel addition to the hub and spoke theory.

J21 also addressed the critical question of how the two crucial semantic processes, control and representation, coexist given their antagonistic demands. Semantic representation, the acquisition and storage of conceptual knowledge, is not sufficient for successful semantic cognition. The need for flexible context-appropriate behaviour, combined with the vast amount of information stored about any particular concept, necessitates additional control processes (Jefferies, 2013; Lambon Ralph *et al.*, 2017). This 'semantic control' inhibits irrelevant, yet often prepotent associations, allowing only selected task-relevant features to drive behaviour. For instance, while not wise to treat them similarly when seeking a food source, if given instructions to sort by shape, you could ignore the strong association with the orange to correctly match the pear and lightbulb. This requirement poses a challenge for the semantic system. On the one hand, it must combine features from different sensory modalities and different encounters to extract the underlying multimodal conceptual structure. On the other hand, it must act only upon the features relevant to a particular situation by shaping activity in the semantic network based on a goal or task context. For the first time, J21 simulated the inhibition and selection of semantic features in a recurrent framework, by limiting the subset of features allowed in the response. Instead of activating all features of a concept, different types of features were retrieved in different tasks. They demonstrated that a semantic system can support both the interacting, yet antagonistic, task context-independent multimodal representation and context-dependent control processes, if distinct areas develop a relative functional specialisation for each process. Architectures that represented task-context and task-invariant conceptual information in distinct areas were able to selectively respond with features in the required domain, whilst successfully extracting multimodal conceptual representations (see Figure 1A).

J21 chose their ‘optimal’ model architecture solely on the quality of the conceptual information extracted during learning, without considering similarity to the neural system. Despite this, it displays remarkable agreement with the coarse structure of the cortical semantic network. There is strong evidence for the presence of a single multimodal hub in the ATL with damage to this region alone resulting in a loss of multimodal semantic representations (Warrington, 1975; Snowden *et al.*, 1989; Patterson *et al.*, 2007). Similarly, there is a clear need for multiple layers to support the transformation from sensory to semantic representations, as demonstrated by the correspondence between differing layers of deep ANNs and regions along the ventral visual stream (Kriegeskorte, 2015; Yamins & DiCarlo, 2016); also see (Hauk *et al.*, 2023). Additionally, the cortex displays a relative functional specialisation for semantic control and representation (Jefferies & Lambon Ralph, 2006; Jefferies, 2013; Lambon Ralph *et al.*, 2017; Jackson *et al.*, 2021). While the ATL supports semantic representation, lateral frontal cortex and posterior temporal cortex (PTC) underpin semantic control (Jefferies, 2013; Noonan *et al.*, 2013; Jackson *et al.*, 2021), with damage to either control area resulting in inconsistent semantic access across task contexts, relying heavily on prepotent associations (Jefferies & Lambon Ralph, 2006; Thompson *et al.*, 2015). Critically, this correspondence suggests many of the architectural features identified in the reverse-engineering process are also utilised by the brain to meet the difficult requirements of a semantic system, and therefore the reverse-engineered model can serve as a mechanistic account of the neuronal semantic network. However, for some of the predicted features, there is currently only limited evidence that they are implemented within the cortex. Here, we highlight the results of a recent study by Tiesinga et al., ((2023) hereafter 'T23', Figure 1B) which begins to address the remaining hypotheses, supporting the reverse-engineered model by further demonstrating its close alignment with the cortical semantic system (Figure 1) and highlighting crucial areas for further exploration.

<Figure 1 about here>

3 Spatiotemporally precise intracortical measurement within the ATL hub

T23 recorded from stereo EEG (sEEG) leads implanted for diagnostic purposes in the most anterior aspect of the ATL, the temporal pole (TP), and the neighbouring cortex of epileptic patients (Figure 1B). Patients watched short (1167 ms) videos showing a male or female actor performing one of two object manipulations (dragging or grasping), preceded by a static presentation of the first frame (for 275 or 875 ms). T23 employed a factorial design with two tasks: a semantic gender discrimination and a non-semantic action discrimination, which were performed equally well (97% correct, (Platonov *et al.*, 2019)). ‘Male humans’ and ‘female humans’ are multimodal semantic concepts, and discriminating them required, according to T23, the combination of visual information about the face and hand, which are known to remain segregated even at the highest levels of the visual system (Premereur *et al.*, 2016). On the other hand, discriminating basic actions is independent of the semantic system, instead depending upon praxis systems within parietal cortex (Orban *et al.*, 2019; Aflalo *et al.*, 2020), as demonstrated by neuropsychological double dissociations (Corbett *et al.*, 2009; Kalénine *et al.*, 2010; Adlam *et al.*, 2013) and neuroimaging (e.g., (Culham & Valyear, 2006; Urgen & Orban, 2021)). Note, that here the task is a simple discrimination between two actions which does not utilise action events that may also recruit the ATL (Binder & Desai, 2011; Fernandino *et al.*, 2013; Desai *et al.*, 2023). Analyses were limited to the broad gamma frequency band as this reflects neuronal activity (Ray *et al.*, 2008; Manning *et al.*, 2009; Ray & Maunsell, 2011; Burke *et al.*, 2015; Xie *et al.*, 2024). An earlier sEEG study utilising these videos and tasks revealed responses restricted to TP and orbitofrontal cortex (OFC) bilaterally but with a clear right hemisphere bias (Platonov *et al.*, 2019). Therefore, T23’s analyses were restricted to the right hemisphere. In both studies, TP responses were brief (<150ms), independent of stimulus duration, and task-dependent. T23 first investigated how the semantic task responses differed across subregions of the right TP.

T23 distinguished 4 TP subregions based on the cytoarchitectonic regions defined in Ding et al., (Ding *et al.*, 2009) and their differences in functional connectivity

(Pascual *et al.*, 2015) (see Figure 1B). The most anterior region (cytoarchitectonic TG) was labeled Tip. Posterior to this, the TP was separated into Dorsal (anterior TA), Ventrolateral (VL; anterior TE), and Medial (TI, 35, 36; entorhinal cortex) subregions (see Figure 2G). These four subregions have differing functional connectivity, focused on auditory and somatosensory regions, default mode network and local paralimbic regions, respectively. Tip and VL were separated despite having similar default mode network functional connectivity (Pascual *et al.*, 2015), as they critically differ in cytoarchitectonics (dysgranular vs. granular), chemoarchitecture (density of SMI 32 (nonphosphorylated neurofilament protein) and WFA (wisteria floribunda agglutinin) labelling; (Ding *et al.*, 2009)), structural connectivity (Sasaki *et al.*, 2023) and intra-TP and subcortical functional connectivity (Pascual *et al.*, 2015). In contrast, areas grouped within the Medial subregion share chemoarchitectural and cytoarchitectonic features (Ding *et al.*, 2009). Semantic task responses were focused in Tip and, to a lesser extent, VL.

<Box 1 about here>

T23 then used Granger causality (GC) to trace the route and direction of information flow between pairs of leads within TP and between TP and other right hemisphere areas (Box 1). Granger causality allows the measurement of directional information flow (Chapeton *et al.*, 2022). GC was assessed at a short (4 ms) and a long (16 ms) delay, covering the range of neuronal interareal connections (from 0 to 20ms (van Kerkoerle *et al.*, 2014; Banaie Boroujeni & Womelsdorf, 2023)). As expected, connections with a short delay were more frequent between TP subregions than with non-TP areas. Remarkably, the task-dependent functional connections were as short-lived as the TP responses (Figure 2A-C). Within the TP, information travelled first from VL to Tip and then back again (Figure 2B-D). Intriguingly, these backward connections were stronger than their forward counterparts (Figure 2E), and more frequently direct (assessed using conditional GC, Box 1; also see Figure 3B). Additionally, T23 identified regions providing input to, or receiving output from, the TP, highlighting the dynamic connectivity with prefrontal cortex and PTC. The time courses of the TP responses were highly related to the timing of their incoming connections (more than the timing of activity in their input regions). This intracortical assessment provided a crucial window into the time-resolved activity and connectivity of ATL subregions during a semantic task, with high spatio-temporal precision and coverage across the TP.

T23 noted considerable agreement between their findings and the reverse-engineered model of J21. Note that while the formal model only contains three layers, this is a simplification; increasing the depth was beneficial and the cortical system requires many steps to go from sensory input to conceptual knowledge (Kriegeskorte, 2015; Yamins & DiCarlo, 2016). Thus, the three model layers may be mapped to the many cortical layers at various different spatial scales, representing more or less of the perceptual-conceptual system. Crucially, the deepest layer must correspond to the location where different sensory streams converge. The neuroimaging simulations in J21 represent a relatively large scale, with the input layer reflecting higher-order sensory information in occipitotemporal cortices and the deepest multimodal layer reflecting the entire ATL. Here the focus is on a more granular scale; an in-depth look at the computations within the temporal lobe, starting with high-level sensory features. Specifically, T23 mapped PTC to the sensory input layer, VL to the shallower processing layer (with unimodal visual feedforward input) and Tip to the deepest processing layer (where the feedforward input from multiple modalities converges; see Figure 1B, bottom). This is an appropriate spatial resolution for comparison with intracortical recordings which allows us to address novel questions regarding fine scale TP neuroanatomy. Here, we highlight the strong convergence between the reverse-engineered model and the intracortical data, providing further evidence for the similarity of the model architecture to the brain and adding neuroanatomical detail (Figure 1). We consider the wider implications of the three main points of convergence in turn: the functional organisation of the multimodal ATL hub, the importance of feedback and shortcut connections, and the interaction between semantic control and representation areas.

<Figure 2 about here>

4 The Functional Organisation of the ATL Multimodal Hub

The findings reported by T23 are in strong accord with earlier intracerebral electrode studies demonstrating evidence for a multimodal hub in the ATL (Shimotake *et al.*, 2015; Chen *et al.*, 2016; Rogers *et al.*, 2021; Cox *et al.*, 2024). However, the ATL is large and inhomogeneous, consisting of multiple cytoarchitectural regions (Ding *et al.*, 2009) (see Figure 1B). Where precisely do the multiple sensory processing streams converge in the ATL? This ‘multimodal convergence point’ corresponds to the deepest layer of the reverse-engineered model and accurate identification of this locale is crucial to comprehend the wider organisation of the ATL. Indeed, J21 demonstrated the emergence of a graded hub centred around this region; conceptual representations are maximally multimodal here and, while all prior layers have bottom-up input in a single sensory modality, their representations are increasingly multimodal as they get closer to this area, due to its top-down feedback. This aligns with the cortical evidence demonstrating a gradual transition from largely unimodal to highly multimodal involvement across temporal cortex (Bajada *et al.*, 2017; Lambon Ralph *et al.*, 2017; Jackson *et al.*, 2018; Rice *et al.*, 2018; Cox *et al.*, 2024). However, questions remain regarding the precise location of the multimodal convergence point. T23 may better separate the role of ATL subregions than earlier ECoG studies as their direct sEEG recordings of the grey matter have three critical advantages: 1) covering the entire cortical surface rather than just the relatively flat cortical gyri, 2) better spatial resolution, with leads separated by 3.5 mm rather than 2 cm, and 3) more anterior coverage of the temporal lobe, extending into the tip of the temporal pole. Note, that while Abel et al., (Abel *et al.*, 2014; Abel *et al.*, 2015) used a specialised ECoG array to extend their coverage anteriorly, their spatial resolution remained limited with only a few electrodes in Tip.

T23 demonstrated the importance of distinguishing TP subregions, as both Tip and VL had more semantic task-responsive electrodes than the Medial and Dorsal subregions. Critically, this aligns with prior ECoG and fMRI studies highlighting the basal anterior temporal surface, focused around VL, as the key site for multimodal conceptual representation (Shimotake *et al.*, 2015; Chen *et al.*, 2016; Rogers *et al.*, 2021). This basal area (described as between 3.8 and 5.2 cm from the pole within anterior fusiform gyrus) is active and necessary across modalities and tasks (Shimotake *et al.*, 2015) and was demonstrated to encode semantic content using a searchlight representational similarity analysis (Chen *et al.*, 2016; Cox *et al.*, 2024). T23 provided additional supporting evidence for the importance of this basal region. During the semantic task, VL received information from both fusiform and lingual gyri and had a high number of task-responsive leads. Moreover, the latency of VL responses correlated with reaction times in this task. However, the prior ECoG assessments specifically identify this ventral area as the critical ATL region for multimodal conceptual processing, while T23 also highlighted the key role of Tip. Tip had the most task-responsive leads and provided the majority of the TP outputs. Furthermore, the GC analyses indicated that information first flows from VL to Tip, and only later travels in the opposite direction (Figure 2B-D). This pattern of receiving early input from VL is more consistent with Tip being the deepest processing layer where the different modalities converge, rather than VL, as the input must first travel from sensory areas to the multimodal convergence point before this region feeds back to influence processing within the different sensory streams (Figure 2F). We could instead assume that VL is the multimodal convergence point and Tip sits atop a sensory pathway, for instance, by providing emotion information from OFC when viewing faces (i.e., the valence pathway postulated by Rice et al., 2015). However, if this was the main sensory pathway driving VL activity one would expect the initial influence to be from Tip to VL, and not vice versa. Additionally, if Tip was responsible for valence, we would expect a greater impact of OFC on Tip than VL activity, yet the direct effects on the two regions were equivalent. Hence, parsimoniously, T23 inferred an alignment between the cortical data such that Tip is the likely locus of the multimodal convergence point, corresponding to the deepest layer of the J21 model, while VL corresponds to the intermediate layer between the multimodal Tip and the visual lingual and fusiform gyri (see Figure 1B, bottom). As such, all of VL's bottom-up input would be expected to be unimodally visual, yet it would be hypothesised to show multimodal responses due to the strong feedback from the proximal Tip region.

T23 proposed that the information represented in Tip consists solely of a ‘label’ which binds together all the features of a specific multimodal concept (such as female), which are represented within shallower regions, including VL. This identifier should not be confused with a verbal label, such as the word ‘male’. Note that T23 used a visual task only and inferred that Tip likely has multimodal representations because its responses and connectivity identified it as the deepest part of the processing stream for the multimodal concepts, male and female. Direct tests of the convergence of information from multiple different senses have indicated the multimodal nature of VL (Shimotake *et al.*, 2015) and, to some extent, Tip (Abel *et al.*, 2015; Shimotake *et al.*, 2015). Systematically assessing the convergence across the ATL of information presented in different sensory modalities remains a critical goal for future research, due to the unimodal nature of T23’s tasks and the fallibility of Granger causality estimates. However, bridging the computational and intracortical investigations, has provided testable hypotheses for these crucial future investigations; that the Tip would indeed respond to stimuli from various presentation modalities and that the relative difference in activation across modalities would be somewhat larger in VL than Tip.

By comparing their sEEG results to the computational model, T23 identified Tip (cytoarchitectonic region TG) as the most likely hypothetical location of the multimodal convergence point, while prior ECoG studies have focused on VL (cytoarchitectonic anterior TE). In either case, VL would demonstrate multimodal responses when tested directly, due to the impact of conceptual information fed backwards from Tip. Thus, the only discrepancy of T23’s proposal with prior intracortical studies is that these did not explicitly identify Tip. Since only the most posterior edge of Tip was assessed in prior ECoG studies and these most anterior electrodes were typically heavily implicated (Shimotake *et al.*, 2015; Chen *et al.*, 2016; Rogers *et al.*, 2021), the focus on VL may simply result from the extremely limited Tip coverage in these earlier studies (see Figure 2G). Consistent with a multimodal role, Tip demonstrates resting-state functional connectivity with all other TP subregions (Pascual *et al.*, 2015), while VL is functionally connected to the pulvinar (Pascual *et al.*, 2015; Zhou *et al.*, 2016), a property typical of higher order visual areas. Furthermore, the strong dominance of temporal cortex structural connections in the anterior-posterior over the dorsal-ventral direction (Moran *et al.*, 1987; Ding *et al.*, 2009; Binney *et al.*, 2010; Herbet *et al.*, 2018; Jung & Lambon Ralph, 2023) appears more compatible with processing streams converging at the anterior tip of the temporal pole and not on the basal surface. However, fMRI assessments demonstrate peak semantic activation in a ventral ATL region overlapping VL and not Tip (see Figure 2G), supporting the prior intracortical studies' focus on VL. These fMRI peaks are consistently posterior to the caudal border of Tip, which we estimate to be located at a Y coordinate of approximately +4 in MNI space (by averaging the 3 points on the boundary shown in Figure 1B, top panel). Eight fMRI studies (Devlin *et al.*, 2000; Sharp *et al.*, 2004; Binney *et al.*, 2010; Visser *et al.*, 2010a; Visser & Lambon Ralph, 2011; Visser *et al.*, 2012; Hoffman *et al.*, 2015; Jackson *et al.*, 2015), performed between 2000 and 2015 yielded local maxima in VL with a Y coordinate ranging from -24 to -5, all at least 9 mm behind the estimated caudal border of Tip (Figure 2G). Indeed, the peak coordinate from (Binney *et al.*, 2010) which has frequently been used to define a ventral ATL ROI (Jackson *et al.*, 2016; Rice *et al.*, 2018; Gore *et al.*, 2021; Jung & Lambon Ralph, 2023), is located at (- 36, -15, -30), approximately 20 mm posterior to Tip. Several explanations may account for the discrepancy between the sEEG analyses and these fMRI peak coordinates in locating the likely multimodal convergence point within the ATL. Firstly, T23 could have misidentified Tip as the most critical subregion for conceptual processing if either, Tip and VL are functionally indistinct despite their cytoarchitectural differences, or if Tip and VL do indeed subserve distinct purposes, but the boundary between them was mislocated. While both VL and Tip had many task-responsive leads with similar response strengths and timing, T23 provided evidence of a functional distinction between these areas. Specifically, the temporal leads providing input to the Tip and VL were clearly segregated along a dorso-ventral axis with VL receiving the majority of input from ventral PTC (vPTC) visual regions (FFA and lingual gyrus) and Tip receiving more dorsal PTC (dPTC) input, which occurred with a longer delay (see Figure 3A). In addition, the feedback connections within TP were much stronger in Tip than VL (Figure 2E). However, since subregion definition involved warping a group-level template onto individual brains, the border between VL and Tip may have been mislocated. Thus, the critical region could be the most anterior portion of VL, yet its effects may have been mislabelled as Tip. To assess the likelihood of this possibility, we performed an additional analysis based on T23's data. We first identified single electrodes which crossed from VL into Tip in two patients. If T23's results were caused by mislocalising the boundary, only measurements of Tip activity near the border would show these effects, while more anterior locations would not. This was not the case: both patients demonstrated strong responses and backward intra-TP connectivity throughout Tip, which reduced around the border with VL, resulting in significant differences between the Tip and VL locations on both metrics (Figure 2H). Thus, these data show a functional distinction along the longitudinal axis of the TP. Whether these differences correspond to smooth graded changes in function or a sharp boundary, they support the functional distinction between Tip and VL, implicating Tip as a more likely site for multimodal convergence than VL.

A second possibility is the social nature of the categories discriminated in T23. Some have proposed that TP is predominantly involved in processing social categories, particularly on the right (e.g., (Olson *et al.*, 2007; Zahn *et al.*, 2007; Ross & Olson, 2010; Olson *et al.*, 2013; Wang *et al.*, 2017)), perhaps due to the input from a valence spoke in OFC (Rice *et al.*, 2015; Lambon Ralph *et al.*, 2017; Rice *et al.*, 2018). However, while the precise area identified across different fMRI comparisons of social and non-social semantics varies dramatically, it appears more frequently located in dorsal TP than in Tip (Zahn *et al.*, 2007; Binney *et al.*, 2016; Rice *et al.*, 2018; Binney & Ramsey, 2020) and the proposed anterior temporal face patch is focused on the basal surface and not the pole (Rajimehra *et al.*, 2009). Indeed, using their specialised ECoG array, Abel et al., (Abel *et al.*, 2014) only identified stronger responses to people than landmarks or tools in leads focused in left and posterior ATL regions, with no such leads in right Tip. Thus, both Tip and VL regions of the left and right ATL appear to represent both social and non-social knowledge, and the social nature of the task is unlikely to explain the discrepancy with fMRI studies (Binney & Ramsey, 2020).

A more likely explanation relates to the limitations of fMRI and the nature of semantic representations. As the temporal resolution of fMRI is poor (~3 seconds), it may simply miss the short-lived TP responses T23 have described. Semantic representations are highly distributed and the contribution of a particular location varies over time. As a result, activity at the same ATL electrode can change strength and even sign as a representation unfolds across a trial (Rogers *et al.*, 2021). Poor temporal (or spatial) resolution means summarising over these changing positive and negative values, and therefore potentially missing task critical regions. This is why we require precision measurements with high spatial and temporal resolution, such as T23's sEEG recordings. These issues are likely true of both Tip and VL to some extent, yet, crucially these factors are likely to affect Tip more if it is indeed the multimodal convergence point. The variability in a region's contribution over time increases from sensory to multimodal conceptual areas and was identified in ATL but not PTC (Rogers *et al.*, 2021). Increased temporal variability results in a lower likelihood of identifying a strong differential response, resulting in the observation that stronger responses tend to be identified earlier in processing streams. If this temporal variability continues to increase as we move further from sensory input, it would be highest at the multimodal convergence point. Indeed, plotting the amount of variance (Rogers *et al.*, 2021) found in each electrode by its Y coordinate (Figure 2G), we can see that the variance does appear to increase in a posterior to anterior direction, even within TP, reaching the highest value in the single Tip electrode. Thus, the effects in Tip may be missed in fMRI, or these effects may be identified alongside stronger activation (and therefore a peak) in VL, detracting attention from them. This possibility is supported by the apparent presence of Tip activity within many of the fMRI studies which simply focus their attention on the vATL peaks (Spitsyna *et al.*, 2006; Binney *et al.*, 2010; Visser & Lambon Ralph, 2011; Jackson *et al.*, 2015; Rice *et al.*, 2018)). The greater strength of the peaks in VL tends to be a measure of the larger or more consistent engagement of this region compared to baseline not how multimodal the response is. Combined with the observation that more peripheral areas have more consistent responses, fMRI peaks would be most likely to occur in the most peripheral area that overlaps between sensory modalities, not the most central. Overall, these dynamic changes being maximal in Tip may be the most likely explanation for the discrepancy between the fMRI peaks and the sEEG results, as crucially many fMRI studies find activity spanning both VL and Tip.

Additionally, their impact may be exacerbated by the smoothing effects of fMRI, the focus on the precise location of a numerical peak, and an overreliance on visual stimuli (which may lead to more frequent identification of VL than the ATL areas that form part of other modalities' processing streams). Further intracortical experiments comparing the engagement and connectivity of Tip and VL in a range of semantic tasks presented across different modalities will be critical to adjudicate this discrepancy of fMRI peaks with spatiotemporally precise measurements which so far were collected only with a simple visual semantic task.

5 The Importance of Feedback and Shortcut Connections

To better emulate brain connectivity, J21 extended prior feedforward-only assessments (Rogers & McClelland, 2004) to assess the impact of semantic network architecture within a recurrent framework. Feedback connections were expected to be crucial to capture the complex dynamics of the semantic system, whereby activity patterns unfold over the full time course of conceptual access as they allow regions to continually influence each other (Clarke & Tyler, 2015; Kietzmann *et al.*, 2019; Rogers *et al.*, 2021; von Seth *et al.*, 2023). Indeed, deep visual ANNs are better able to fit time-resolved neural data when they include both feedforward and feedback connections (Güçlü & van Gerven, 2017; Kietzmann *et al.*, 2019; Nayebi *et al.*, 2022; Thorat *et al.*, 2022; von Seth *et al.*, 2023). Both the reverse-engineered model and time-resolved investigations of the cortical semantic network demonstrate a high degree of interaction between regions at different levels in the processing stream. For instance, MEG studies demonstrate recurrent connectivity between the ATL and vPTC during semantic task performance (Clarke *et al.*, 2011; Rahimi *et al.*, 2022), a pattern also identified by T23 (involving Tip more frequently than VL). However, with the greater spatial resolution afforded by sEEG, T23 were also able to demonstrate the importance and timings of both feedforward and feedback connections between VL and Tip.

T23 found the latency of TP responses and connections varied considerably across patients, ranging from 125 to 225 ms (consistent with earlier studies; (Clarke *et al.*, 2011; Chen *et al.*, 2016; Rogers *et al.*, 2021; Rahimi *et al.*, 2023; von Seth *et al.*, 2023) and lasting approximately 100 ms. However, within patients these effects followed a stereotyped sequence, starting with feedforward connections from VL to Tip, on average 140 ms after stimulus onset and 40 ms (20-70 ms across patients) before the TP responses (Figure 2D). This forward connection was quickly followed by a recurrent interaction, starting on average 12 ms (range 10-13 ms) before the TP responses. Remarkably, the feedback connections were stronger than the feedforward connections (both between Tip and VL and within Tip, Figure 2E), demonstrating the importance of feedback within the ATL. While the specific timings may be affected by the simple nature of the task employed, finding evidence of feedback connections at both inter- and intra-regional levels supports the importance of the recurrent semantic processing simulated by J21. The spatial proximity of input and output regions in the reverse-engineered model precludes a detailed comparison of the timing of feedforward versus feedback connections. However, the simulation does result in a very short initial feedforward-only period while the unimodal sensory information propagates through the network to activate the multimodal hub, followed by persistent recurrent connectivity, consistent with the findings of T23.

How does this recurrent connectivity support semantic cognition? Visual ANNs demonstrate that feedforward processing alone is sufficient for accurate object identification, particularly in relatively easy tasks, such as the gender discrimination employed by T23. However, this may not provide the efficiency needed to complete hard semantic tasks, quickly, with limited resources (Nayebi *et al.*, 2022). T23 proposed that the feedback from the label in Tip allowed pattern completion, activating features of the concept across the shallower TP regions. Hence, the complete representation of a concept would involve both Tip and VL (Box 2, Figure A). Indeed, in the reverse-engineered model, the early activation of the multimodal hub allows the system to start to represent related sensory features across other senses and for the multimodal conceptual structure in the deep hub to influence processing in, and information transmission from, shallower sensory layers. The initial feedforward sweep may be sufficient to provide a coarse estimate of what an object might be, even when recognition is difficult. Employing multimodal information in a top-down manner at this stage could allow further processing to target specific diagnostic features promoting efficient recognition of objects, words and people (Kveraga *et al.*, 2007; Clarke & Tyler, 2015).

The reverse-engineered model demonstrated the importance of shortcut connections in semantic cognition for the first time. The addition of direct connections between shallow sensory and deep multimodal regions had a dramatic positive effect on both representation quality and speed of learning, despite the potential competition between fulfilling these two metrics. Moreover, these improvements were seen despite the shortcut connections representing a very small proportion of the total connections between sensory areas and the hub (there were around nine times as many connections at each step in the indirect route as there were shortcut connections). While different connectivity proportions were examined, a level of sparsity is necessary to maintain biological plausibility; long-range white matter connections are present in the brain, yet sparse due to metabolic and packing constraints (Nelson & Bower, 1990; Plaut, 2002). While evidence for shortcut connections within the cortical semantic network is limited, the idea of multiple processing routes is highly compatible with the existence of two broadly distinguishable sets of white matter supporting interareal connectivity; the occipito-temporal projection system of U-fibres connecting neighbouring areas and the inferior longitudinal fasciculus bridging across longer distances (Catani *et al.*, 2003; Bajada *et al.*, 2017; Herbet *et al.*, 2018). Moreover, tractography assessments demonstrate greater ATL connectivity with more distal posterior fusiform than mid-fusiform cortex (Bouhali *et al.*, 2014). Similarly, connections are identified in non-human primates between higher order visual areas and cortical areas two levels further down in the hierarchy, for example, TEad receiving mainly from TEpd supplemented by a weak projection from TEOd (Kravitz *et al.*, 2013). However, these structural connectivity assessments provide no information about whether and how these routes are utilised for semantic cognition. Nor can they tell us the direction of these possible shortcut connections.

By assessing the directed, effective connectivity of the ATL with high spatial and temporal precision, T23 demonstrated support for the involvement of both a multi-step route and shortcut connections in semantic cognition. GC analyses identified 21 direct connections in the forward direction from vPTC to TP leads: 16 into VL and 5 directly projecting into Tip (Supplementary Tables 10 & 11 in T23). At least 2 of these 5 connections operated at 4ms (Figure 3A). In the feedback direction, there were 20 direct connections from TP to vPTC leads, with 15 of these arising in the Tip. At least 5 operated at 4ms. Shortcut connections in the feedback direction therefore appear numerous, although it should be noted that there were more Tip than VL leads (24 versus 18) and all VL leads received feedback from Tip, suggesting that many more connections supported feedback via a multi-step route as well. Despite these factors, the number of feedback shortcut connections may be surprising given the need to have sparse structural connections over long distances due to packing and metabolic constraints. It may be that the long-distance structural connections are sparse, yet interactions in the feedback direction heavily depend upon this route. Alternatively, it may be that the number of shortcut connections identified by T23 was artificially inflated somewhat due to the inclusion of indirect connections which were estimated to be direct over a relatively long lag time. Connections at the 16ms delay may reflect weak direct connections involving thin fibers or may be an incorrect characterisation of multiple rapid indirect connections. Overall, T23 provide initial novel evidence of sparse shortcut connections supporting bidirectional interactions between vPTC and Tip (bypassing VL, Figure 1B, bottom). It is important to note that evidence for these shortcut connections were derived from Granger causality computations. Granger Causality is an inferential technique, with the direct connections identified open to alternative interpretations, such as very rapid multi-synaptic pathways, particularly with the longer 16ms delay. However, even when restricting the analyses to the 4ms lag only, there was evidence of sparse shortcut connections. Future investigations should utilise additional methods to determine whether there is converging evidence of shortcut connections into the ATL and provide better estimations of their sparsity.

Why should there be shortcut connections between vPTC and Tip? J21 found shortcut connections were effective at helping ameliorate the slowing caused by increasing depth. Error-driven learning across many layers is initially slow as it requires the manipulation of multiple connection weights (He *et al.*, 2016; Saxe *et al.*, 2019), which shortcut connections avoid by creating an alternative path through fewer layers. However, this does not explain why these connections also improved the quality of representations. Perhaps, quickly projecting coarse sensory information into the multimodal hub, activates conceptual representations able to support top-down semantic processing. This conceptual information could then impact sensory processing quickly by utilising the feedback shortcut connections from the Tip to vPTC. Thus, shortcut connections may improve representations by encouraging dynamic interactions between perceptual and conceptual regions. These possibilities are highly compatible with the sEEG data, which showed relatively early responses in the TP and even earlier intra-TP connectivity, speculated to be dependent on responses below the detection threshold. Further investigation of shortcut connections will likely prove crucial to a mechanistic understanding of the dynamics of the cortical semantic system.

<Figure 3 about here>

6 The Interaction between Control and Representation Regions

By varying connectivity *in silico*, J21 demonstrated that information signalling task-context should be received in more peripheral regions of the semantic network and not the core multimodal region. Shielding the multimodal hub from the direct effects of task context is imperative to allow it to extract the full conceptual structure based on the relationship between features presented across different contexts and episodes. To date, the functional connectivity of semantic control and representation regions has been assessed with fMRI (Jackson *et al.*, 2016; Jung & Lambon Ralph, 2023), which has a notoriously poor temporal resolution, and MEG (Farahibozorg *et al.*, 2022; Rahimi *et al.*, 2022; Rahimi *et al.*, 2023), which has a limited spatial resolution. Thus, J21 provided a clear prediction as to how semantic representation and control regions interact (Figure 1A, bottom), yet existing studies provide little evidence to adjudicate this hypothesis. With the excellent spatiotemporal resolution of sEEG, T23 were able to provide unique evidence illuminating how and where frontal regions interact with subregions of the ATL hub during basic semantic processing. Specifically, the strength and timing of the feedforward connection from vPTC to VL was mediated by the OFC. By controlling the timing of these inputs, OFC indirectly determined the duration of the TP responses. Furthermore, within the TP, the feedforward connection from VL to Tip was mediated by both OFC and inferior frontal gyrus (IFG; Figure 3B). Both findings are consistent with J21's prediction that control regions should not act directly at the multimodal convergence point but instead influence more peripheral areas (Figure 1). Moreover, T23 specifically showed the influence of these potential control areas on the early feedforward input from unimodal sensory streams and not the strong multimodal feedback signal. Further investigations will be crucial to assess the interaction between control and representation regions in more controlled, demanding semantic tasks. In particular, recording of the left hemisphere will allow a better assessment of the role of semantic control regions, including the IFG which demonstrates relatively left-lateralised activation (Jackson, 2021). The interaction between the frontal cortex and the ATL has previously been hypothesised to occur via posterior temporal control regions (Davey *et al.*, 2015; Jackson, 2021; Hodgson *et al.*, 2023), based on the known connectivity (Matsumoto *et al.*, 2004; Catani *et al.*, 2005; Jackson *et al.*, 2016) and the apparent functional similarity of IFG and PTC (Jefferies, 2013;

Thompson *et al.*, 2015). T23 found little evidence for this hypothesis. However, critically, these assessments were only performed in the right hemisphere, while only the left PTC appears important for semantic control (Jefferies & Lambon Ralph, 2006; Jefferies, 2013; Noonan *et al.*, 2013; Davey *et al.*, 2015; Thompson *et al.*, 2015; Jackson, 2021). Future intracortical studies should determine whether this is an additional important route in the left hemisphere.

How do these fronto-temporal connections support controlled semantic cognition?

Semantic control assessments typically focus on the ability to selectively retrieve task-relevant information, as modelled in J21. Frontal areas may shape the information retrieved within the hub by controlling its inputs, manipulating the interaction between the hub and its spokes in a task-appropriate manner. Activity in the hub would then affect responses in earlier areas via its strong feedback connections, explaining the effects of task context in sensory areas (e.g., (Martin *et al.*, 1995) as simulated in J21). The dynamic interaction between control regions, the multimodal hub and sensory regions may collectively shape ongoing retrieval, informed by both the information accessed so far and the task context. While the specific interactions would differ between different tasks, a high degree of interaction would support efficient, task-relevant processing. For instance, when given the task to decide if an object is dangerous, determining that it is a plant would allow further processing to focus on different features (e.g., toxicity, colour) than if it were an animal (e.g., size, sharp teeth). Similarly, when determining whether two words are related, we must access features of the first word and use these to guide our search for relevant features of the second word. Intriguingly, most research on semantic control has focused on the IFG as the critical frontal area, although meta-analyses (Figure 3C) show consistent activation across a broad region, including lateral OFC

(Binder & Desai, 2011; Noonan *et al.*, 2013; Jackson *et al.*, 2021) , which has strong structural connections with the ATL (Von Der Heide *et al.*, 2013). In contrast, T23 highlighted a greater role for OFC, which includes most of the electrodes mediating intra-TP connectivity (Figure 3B), and all those responsible for controlling the timing of the inputs to TP. The differing spatial focus of the two methods could be explained, at least in part, by poor fMRI signal within OFC, or the focus of the intracortical investigation on the right hemisphere, as right IFG appears less important for the control of meaningful stimuli (Jefferies, 2013; Thompson *et al.*, 2016; Jackson *et al.*, 2021; Hodgson *et al.*, 2023). Indeed, comparison of these regions across the two hemispheres will be crucial, and T23's analyses focusing on the right hemisphere should not be considered evidence against the crucial role of the left IFG in semantic control. Moreover, the importance of the OFC in T23's analyses, may not be a general semantic control effect. For instance, the social nature of the gender task may drive the OFC response. Within the graded hub-and-spoke view, OFC has predominantly been considered a 'spoke' for emotional content or valence, providing input to the ATL via the rostral TP (Patterson *et al.*, 2007; Rice *et al.*, 2018). This is consistent with the direct input from OFC to VL identified by T23, but less so with the initial direction of effective connectivity being from VL to Tip. Alternatively, the IFG and OFC may underpin distinct control-related processes, consistent with their differing functional and structural connectivity patterns; the OFC is structurally connected to TP via the uncinate fasciculus whereas IFG connects to PTC control areas via the arcuate fasciculus (Catani *et al.*, 2005). IFG may be central for semantic control as typically assessed with tasks requiring the manipulation of semantic knowledge to resolve ambiguity, inhibit dominant features and select less dominant features as required, yet show limited involvement in T23 simply as the semantic control demands were low in the gender identification task, which was relatively easy (reaction times were around 820 ms) and did not emphasise these key semantic control processes. In contrast, the OFC involvement may not reflect semantic control as typically conceived but a related, yet separable process. For instance, OFC has previously been associated with the top-down support of visual object recognition, providing contextual support to help identify an object efficiently based on the statistical probabilities of its environment (Kveraga *et al.*, 2007). Alternatively, T23's analyses may lead us to consider an alternative hypothesis that OFC may assess whether an appropriate amount of evidence has accumulated to make the specific judgement required by the task. This would allow response times to differ between tasks varying in the required depth of semantic processing (e.g., 'Is it a bird?' versus 'Is it a robin?') by curtailing unnecessary processing within TP once the relevant behavioural output is possible. Manipulating this threshold allows an accurate, yet fast response to any given task. T23 demonstrated that OFC modulated the duration of the input to VL from vPTC, which determined the timing of the TP responses, perhaps allowing them to be very brief for the easy gender task. Within the conceptualisation by T23, this would mean OFC modulates the input to the Tip neurons representing the gender concept labels, to determine how stringently it should test the evidence provided by the semantic features collected by VL and identify whether sufficient evidence has accumulated. Alternatively, the role of OFC in reward processing (Balleine *et al.*, 2011; Rudebeck & Murray, 2011; Walton *et al.*, 2011; Rolls, 2023) could point to a broader role in driving semantic processing to favour rewarding, task-relevant or socially significant, features. Currently, the functional significance of the OFC in semantic control is unknown and there are multiple ways to interpret T23's findings. However, the apparent control OFC exerts over TP responses is an intriguing avenue for future exploration. Further intracortical investigations of executively demanding semantic tasks are necessary to distinguish the possible roles of left and right OFC and IFG in controlled semantic cognition.

7 Implications

In this article, we highlight the remarkable convergence between recent computational modelling experiments and precision physiological recordings of the semantic system (Figure 1). A key reason for this success is the high spatial and, in particular, temporal resolution of the stereo EEG technique used by T23. Indeed, the recorded ATL responses would likely go undetected in fMRI due to both their extremely short duration and the dynamic nature of the code utilised by the distributed concept representations (Rogers *et al.*, 2021). Of course, the spatial specificity could be improved even further with single cell recordings (Leonard *et al.*, 2023; Xie *et al.*, 2024). However, these have yet to be performed in the ATL. A second crucial factor is the reverse-engineering approach J21 adopted. While a single model could be more or less suited to a function, systematic comparison across models allows the identification of beneficial architectural features. However, the reverse-engineered model and the brain could improve performance in distinct ways. Indeed, recent advances in the performance of deep visual ANNs have actually reduced their similarity to the cortical visual system (Schrimpf *et al.*, 2020). This discrepancy is likely a function of the models' increasing complexity, with higher performing models including far more units and connections, a solution that is not available to the biologically-constrained brain. J21 circumvented this issue by matching model complexity. The convergence between the cortical data and the optimal model provides a crucial validation of the reverse-engineering approach's ability to identify brain-like solutions. This occurred despite T23's use of a simple semantic task and focus on the lesser observed right hemisphere. Consequently, additional investigations employing precision recording will be crucial to validate and extend our knowledge of these points of convergence. Most intriguingly, T23's results find initial evidence supporting a novel prediction generated by the reverse-engineered model of controlled semantic cognition; *the importance of utilising shortcut connections* between distal layers. Shortcut connections may be critical to accurately simulate the complex dynamics of the semantic system in future neuroanatomically-inspired computational models. Indeed, more neurobiologically plausible patterns of recurrent connectivity and sparse long-range connections may improve the ability of deep ANNs to provide brain-like simulations of a multitude of cognitive processes. Furthermore, if shortcut connections allow multimodal regions to impact distal sensory areas with top-down conceptual knowledge, all state-of-the-art deep ANNs performing visual or auditory identification based on a single sensory modality in isolation, are inherently limited.

T23's sEEG recordings represent the end of a long wait for data pertaining to the operations within the semantic hub, allowing comparison of the J21 modelling to the brain at a more granular level. Together the neural and computational assessments highlighted the *functional heterogeneity* of cytoarchitectonically distinct ATL subregions, including the importance of Tip and VL. Simulating multiple subregions at different processing depths within the ATL provides a mechanistic explanation of the graded changes in the modality of information represented across this hub. T23 aligned the deepest model layer with Tip, while previous work has typically considered VL to be the multimodal convergence point. Determining which subregion sits atop this organisational hierarchy and is therefore the centre of the graded hub, will be critical to understand the functional organisation of the ATL and inform neuroanatomically-precise models of semantic cognition in health and disorder. Tip and VL appear to have complementary roles, with T23 speculating that VL represents the higher-order visual features present in the input and Tip represents the multimodal combination of features defining a concept (the concept label; Box 2). They hypothesised that VL must provide adequate evidence of a particular concept to activate Tip, which then engages additional sensory features in the other shallower unimodal regions. J21 demonstrated how this process of pattern completion across different spokes may be performed in a context-dependent manner. Indeed, aligning the neural and computational evidence illuminated how prefrontal control regions impact ATL responses. Surprisingly, this revealed a *potential additional role of the OFC in semantic control* which may have been obscured by the prior focus on nearby IFG. This possible involvement in warping the semantic space to fit the task context is consistent with recent suggestions that OFC represents the task state space (Wilson *et al.*, 2014; Schuck *et al.*, 2016; Liao *et al.*, 2023). Alternatively, OFC may have a related but distinct role, for instance, in determining when sufficient semantic processing has occurred. The relative contribution of the OFC and IFG in a more controlled, demanding semantic task remains to be seen. OFC controlled both the timing of the input to VL and the connection from VL to Tip (alongside IFG). As predicted by the reverse-engineered model, this PFC mediation did not directly impact the area proposed to be the multimodal convergence point.

By bridging the computational and neural literature on the ATL hub, this perspective has generated greater understanding of how the brain can meet the core yet conflicting, requirements of controlled semantic cognition. Now is a crucial time to gain greater understanding of how this is made possible by the architecture and mechanics of the human cortical system as generative AI models, such as ChatGPT, based on large language models (LLMs) are having a disruptive impact on society (Hutson, 2021a; b; Conroy, 2023). These models have the semblance of semantic cognition skills, appearing to store conceptual knowledge and use it flexibly as the context requires, using a vector based representation similar to those proposed earlier for the semantic system (Piantadosi *et al.*, 2024). Yet they do not understand meanings *per se*, instead simply repeating previously used phrases that seem relevant for the context, known as “parroting” (Bender *et al.*, 2021; Fayyad, 2023; Mitchell & Krakauer, 2023). Moreover they require a vast amount of training and resources to learn (Bartoldson *et al.*, 2023; Luccioni *et al.*, 2023). In contrast, the human brain is capable of a much more remarkable feat of extracting concepts from multiple events, each containing only a subset of the important features, while also producing only the subset of features relevant to the current task demands. Moreover, it meets these core, yet conflicting, requirements of controlled semantic cognition with far fewer resources and training. Understanding how this is possible is crucial to advance both our understanding of the brain and the development of new generations of AI technology.

8 Conclusions and Future Directions

Within this perspective, we demonstrate how we can advance our understanding of the neural architecture for controlled semantic cognition by combining evidence from *in silico* experiments and *in vivo* precision measurements. These approaches are symbiotic; computational models must be encouraged to be brain-like and neural measures require formal, mechanistic explanation (Doerig *et al.*, 2023). Thus, we demonstrate the power of computational cognitive neuroscience to bridge the explanatory gap between brain and behaviour. Only through continual translation between computational modelling and neuronal observations are we able to refine and test precise, mechanistic theories of how complex functions arise from the brain. However, the process of applying this to understand controlled semantic cognition has only just begun. Indeed, the convergence identified here has only been demonstrated within the right hemisphere, for unimodal visual stimuli of a single type with a very simple semantic task. Confronting the model with more sEEG (Rogers *et al.*, 2021; Mercier *et al.*, 2022; Zhu *et al.*, 2022) and single cell recordings (Bausch *et al.*, 2021; Jamali *et al.*, 2021; Paulk *et al.*, 2022; Zheng *et al.*, 2022; Jamali *et al.*, 2024; Khanna *et al.*, 2024) (Xie *et al.*, 2024) across hemispheres and varied semantic tasks remains an important task for the future. Future experiments should: 1) directly compare the responses and connections of each cytoarchitectural region in the left and right ATL using sEEG; 2) present semantic stimuli across different modalities combined with sEEG to directly test which cytoarchitectural regions show multimodal responses across time, and 3) distinguish the connectivity patterns associated with semantic control versus representation processes and the roles of IFG and OFC by manipulating semantic control demands. In addition, neuroanatomically-constrained models of controlled semantic cognition should be constructed to allow closer convergence with these cortical recordings, and their ability to simulate the dynamics of semantic processing across tasks tested.

Box 1: Granger causality analyses characterise the directed connectivity within the semantic network

sEEG recordings summarise the synaptic input and spiking activity over approximately a cubic mm around each contact point (channel) (Mercier *et al.*, 2022). The dominance of broad gamma band power in the response periods, indicated that T23's measurements did reflect the spiking activity of neurons, at least in part (Leonard *et al.*, 2023; Xie *et al.*, 2024). Unlike coherence measures (Anand *et al.*, 2023), Granger causality (GC) analyses allow inference of the direction of interaction between two electrode channels. To that end, GC quantifies the extent to which predictions of the activity in one channel are improved by taking into account the previous activity on the other channel, over and above its own history. T23 achieved this by fitting an autoregressive (AR) model to the channel activity. While GC is designed for stationary signals, the nonstationary dynamics of the interactions could be assessed by applying GC to short (100 sample) windows that could be considered approximately stationary, which are then shifted across the trial interval in order to obtain a time-varying GC. The number of model parameters, which is the product of the square of the number of channels considered and the number of previous samples included in the AR model, needs to be balanced with the amount of data available in the window. For the analysis presented in T23, this amounted to 4 prior samples for a pair of channels. Taking samples every 1 or 4 ms resulted in maximum delays of 4 and 16 ms, covering the appropriate delays for probing direct communication between neighbouring and distal regions (van Kerkoerle *et al.*, 2014; Banaie Boroujeni & Womelsdorf, 2023), respectively. As some interactions may be task-independent, the GC measure was z-scored relative to the baseline prior to the onset of the static image. While standard GC analyses are directional, they fall short of identifying direct causal influences as a common input affecting the two channels could lead to the detection of spurious connections. This can be addressed by conditioned GC, where the predictive model includes additional channels to factor out such common input. Using this procedure it was possible to distinguish direct connections from indirect interactions mediated by a third region. Taken together, z-scored conditioned and unconditioned GC are a powerful tool to dissect the flow of information in a spatiotemporally precise way revealing the dynamic interactions across a hitherto unappreciated regional heterogeneity in the temporal lobe.

Box 2: Benefits of a multi-level representation of concepts

From their pioneering recordings, T23 hypothesised a two-level representation of concrete concepts in TP with only the multimodal concept label stored in Tip, and the defining unimodal, sensory features stored in other TP subregions, including VL for vision. This organisation allows many concepts to be represented in a small space, as many labels can be stored in Tip and the same unimodal features can subserve multiple concepts. Such efficient storage is crucial given the many concepts humans can acquire. Moreover, the activation of shared features allows generalisation between similar concepts. Importantly, these advantages depend on the conceptual representations being distributed across many neurons, as assumed by both T23, and the J21 model. Indeed, T23 demonstrated that about two thirds of all Tip neurons are activated by a single exemplar. This distributed processing stands in contrast to the sparse coding in hippocampus (Quiroga, 2012), consistent with the general view that episodic and semantic memory should use sparse and distributed coding, respectively (McClelland *et al.*, 1995; Frankland & Bontempi, 2005). T23 derived the two-level representation by distinguishing 3 steps of TP activity when matching a visual item to a concept (Box Figure A): 1) VL collects the evidence in favour of a concept (here either male or female); 2) when this evidence is sufficient for one particular exemplar, as determined by PFC, Tip is activated; 3) this Tip activation automatically feeds back to VL and the other unimodal parts of TP, such as the dorsal area for audition, without outside control. Thus, the feedback from Tip results in a pattern completion process, whereby all the concepts' features are activated, even those that were not initially activated by the presentation of the item. This is only possible due to the convergence of the different sensory streams at one location (i.e., a multimodal convergence point or hub), in a network with multiple layers (i.e., depth) and as J21 demonstrates, this results in gradual changes in the extent to which areas are multimodal as the feedback provides multimodal influences on earlier regions. Thus, performing pattern completion across modalities requires a graded deep hub-and-spoke architecture.

The two-level representation, proposed by T23, could also allow the implementation of a hierarchical structure of categories (Box Figure B&C) (Quillian, 1968; Warrington, 1975) whereby Tip includes both class and exemplar labels. Neuronal Tip populations acting as labels of two exemplars sit at the same level (here male and female) and may be linked by reciprocal inhibitory connections, along with populations representing labels of the same class level. On the other hand, representations of the higher class level (humans) may be based on the overlap of features activated in the lower exemplar level (in this case, the features associated with both male and female humans) resulting in both direct (excitatory) connections between the class label and these features, and indirect conncetions via the exemplar labels. While this extended T23 two-level scheme is reminiscent of Quillian's hierarchical propositional model of living things (Quillian, 1968), semantic knowledge is accessible at multiple levels of the hierarchy and not just the superordinate level, as labels at different hierarchical levels in Tip are linked to features in VL. Furthermore, given its similarity to recurrent neural networks of the hub-and-spoke theory (Rogers & McClelland, 2004; Jackson *et al.*, 2021), the extended T23 scheme does not suffer from the drawbacks of hierarchical propositional models (McClelland & Rogers, 2003). First, both class and exemplar labels are linked directly to their semantic features, leading to equally fast verification (Mccloskey & Glucksberg, 1979; Murphy & Brownell, 1985), while the stronger link between class labels and their properties may reflect the combination of their direct links and indirect links (via the exemplar level; Box Figure C). Second, step 2 in the T23 scheme is reminiscent of category verification in ANNs, which is based on similarity between representations of items and categories, a more efficient strategy than searching a hierarchical tree (McClelland & Rogers, 2003). Finally, the extended T23 scheme predicts an increasing duration of ATL activity as tasks address higher levels of the hierarchy. Indeed, the indirect links between label and features for superordinate classes (such as living things versus artefacts) would involve a greater number of synapses than the connections between features and subordinate exemplar labels (such as female, male), allowing multiple recurrent loops using the feedback connections within Tip and between Tip and VL. Consistent with this prediction, the TP activation for superordinate level animacy judgements reported by Rogers et al., (2021) was much longer than those for the exemplar-level gender judgements in T23. However, because higher level categories also have direct label-features connections (Box Figure C), tasks involving low and high hierarchical levels elicited equally fast EEG responses (Thorpe *et al.*, 1996). Thus, it seems that the small modification to the T23 scheme for representing concepts in TP, introduced here, allows us to reconcile parallel distributed and hierarchical semantic processing. The modified scheme also reconciles category and feature based views of semantic representation (Frisby *et al.*, 2023), and could even accomodate vector space views, as the activity in VL could represent non-interpretable dimensions as much as sensorimotor features, and activity in Tip the point in vector space defined by these dimensions. Furthermore, semantic representation in the scheme is grounded (Frisby *et al.*, 2023), as activity in VL feeds back to the visual and other sensorimotor areas that support its feature/dimension representation. However, because it exhibits pattern completion it also has a self-contained element.

In addition to providing insights into the semantic representations, the two level concept representation in TP may have wider implications. Indeed, the functional connections of Tip with superior posterior temporal regions of the language network (Price *et al.*, 2005; Hickok & Poeppel, 2007; Price, 2012; Fedorenko *et al.*, 2024) may allow one-to-one links between the concept labels and the corresponding verbal labels. Combined with those between the hub and the spokes, these links may explain the wide distribution of linguistic meaning across cortex, both within and outside the language network (Pereira *et al.*, 2018; Wang *et al.*, 2018).

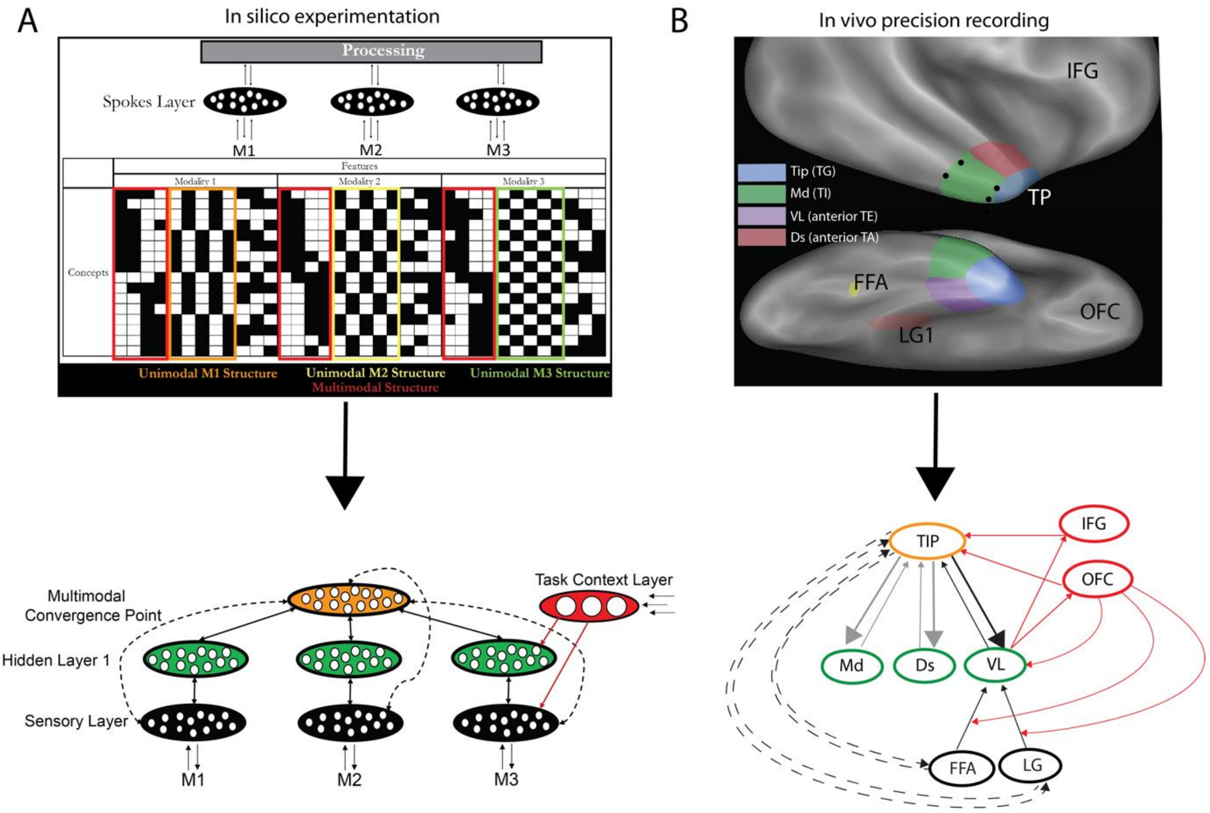


Figure 1: The convergence between precision sEEG recordings of the cortex (from T23) and a model reverse-engineered to promote controlled semantic cognition (from J21). **A:** *In silico* experimentation compared a series of architectures with this basic setup on their ability to learn the underlying conceptual structure while responding with the task-appropriate subset of features. **Top:** The grid shows the features of each concept in the model environment (black features are present for that concept). The environment included features in 3 different artificial sensory modalities, each of which had its own structure, as well as features related to a multimodal structure (shown by the coloured boxes). The architectures varied on their connections and the presence of one or two hidden layers but all had sensory regions to allow the input and output of features in each sensory modality and the same number of units and connections overall. The models received features of a concept in one of the three modalities, and a task-context signal designating the modality in which output was required, resulting in 9 different possible ‘tasks’ (e.g., experiencing a concept’s features in modality 2 and

responding with its features in modality 3). **Bottom**: The optimal model architecture identified by J21. Black: sensory input and output regions in each of the three modalities; green: hidden units for processing in a shallow hidden layer connected to the sensory areas, orange: hidden units for processing that connect the different processing streams resulting in multimodal convergence; red: control regions providing the task-context. **B:** The TP subregions explored with sEEG and their proposed correspondence with the model layers. **Top**: The four TP subregions are shown on lateral and ventral views of the inflated MNI brain; blue: Tip, green: VL, red: Dorsal, and purple: Medial subregions. The black dots indicate the positions used to calculate the approximate Y coordinate of the rostral and caudal borders of VL used in Figure 2. The yellow and red hatching indicate the FFA and lingual gyrus regions providing visual input to TP. Red labels refer to regions included in the network shown in D. STS: superior temporal sulcus, ITS: inferior temporal sulcus, OTS: occipito-temporal sulcus, Coll S: collateral sulcus, IPS: intraparietal sulcus, Cen S: central sulcus, IFS: inferior frontal sulcus, FG: fusiform gyrus, LG: lingual gyrus, IFG: inferior frontal gyrus, OFC: orbitofrontal cortex. **Bottom**: Functional connections between Tip, VL and vPTC regions (FFA and lingual gyrus), OFC and rostral IFG found using Granger causality in T23. This utilises the same colour code as the model in A. Connections identified with putative control regions are coloured red. The shortcut connections are displayed with dashed lines. Black (and red) lines represent connections identified in T23, while grey lines show connections that are assumed to be important for non-visual processing based on prior work and showed weaker involvement in T23. These connections link Tip with the Dorsal (D) region, (corresponding to anterior TA), processing auditory and also somatosensory information, and within the Medial subregion (specifically with TI), processing olfactory information (Ding *et al.*, 2009).

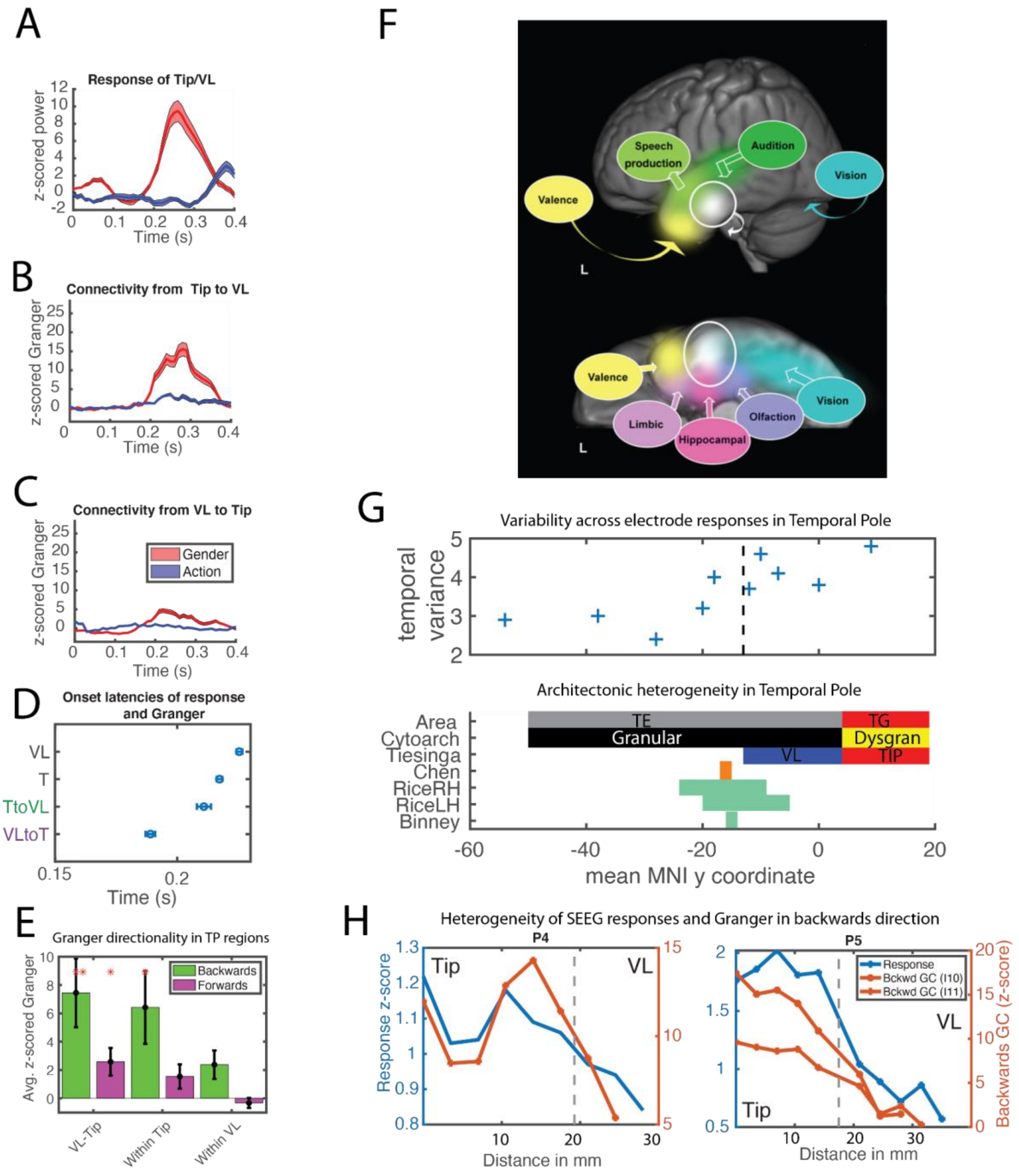


Figure 2: Key findings from T23's sEEG recordings and their relation to prior sEEG and fMRI studies. **A-D**: The sEEG data of patient 5 (see T23 for all 4 patients). The time courses of the responses in Tip and VL (A) and the connectivity (Granger causality at the short delay) backward from Tip to VL (B) and forward from VL to Tip (C) in the semantic and non-semantic tasks. D shows the onset of these responses and connections (at the short delay). In each patient, the connectivity from VL to Tip preceded the connectivity from Tip to VL and the Tip and VL responses. **E**: The group average strength of the forward (purple) and backward (green) connections within and between Tip and VL for the semantic task. **F:** The graded hub-and-spokes scheme from (Rice *et al.*, 2015) shown on the lateral and ventral views of the left hemisphere, including a basal multimodal convergence point shown in white corresponding approximately to VL. **G**: Locating ATL structure and activity on the rostrocaudal dimension. The top diagram shows each electrode in Rogers et al., (Rogers *et al.*, 2021) plotted at its location on the Y axis in MNI space by the amount of variance in the coefficient change for the classifier fitted to the raw sEEG signal (adapted from Rogers et al., 2021), reflecting the changing nature of the representation at that location over time. This is stronger for more anterior electrodes, and strongest for the one electrode judged to be within Tip (see scale in the lower part of the diagram). The lower diagram utilises the same scale and displays the location of areas TG and TE, and their cytoarchitectural differences, as well as Tip and VL as defined by T23 (corresponding to TG and anterior TE, respectively). TP extends from the rostral end of the temporal pole to the limen insulae, as in Ding et al., (Ding *et al.*, 2009) and Pascual et al., (Pascual *et al.*, 2015). Additionally, the locations of local maxima in prior studies are shown, including the earlier ECoG study (Chen *et al.*, 2016) (brown), the range at which right

and left hemipshere peaks (green) were identified in the multiple fMRI studies, presented in Table 1 of Rice et al., (Rice *et al.*, 2018) and the peak fMRI location from Binney et al., (Binney *et al.*, 2010) (green), which has been repeatedly used as a vATL ROI (+/- 36, -15, -30). **H:** For the two electrodes in two patients which included both Tip and VL regions in T23, response strength and feedback connectivity (towards posterior TP) are plotted as a function of distance from the most rostral lead of the electrode, inserted almost perfectly parallel to the Y axis. The grey dashed line shows the division between VL and Tip. More anterior locations showed stronger feedback connections and stronger responses, resulting in the overall differences between Tip and VL.

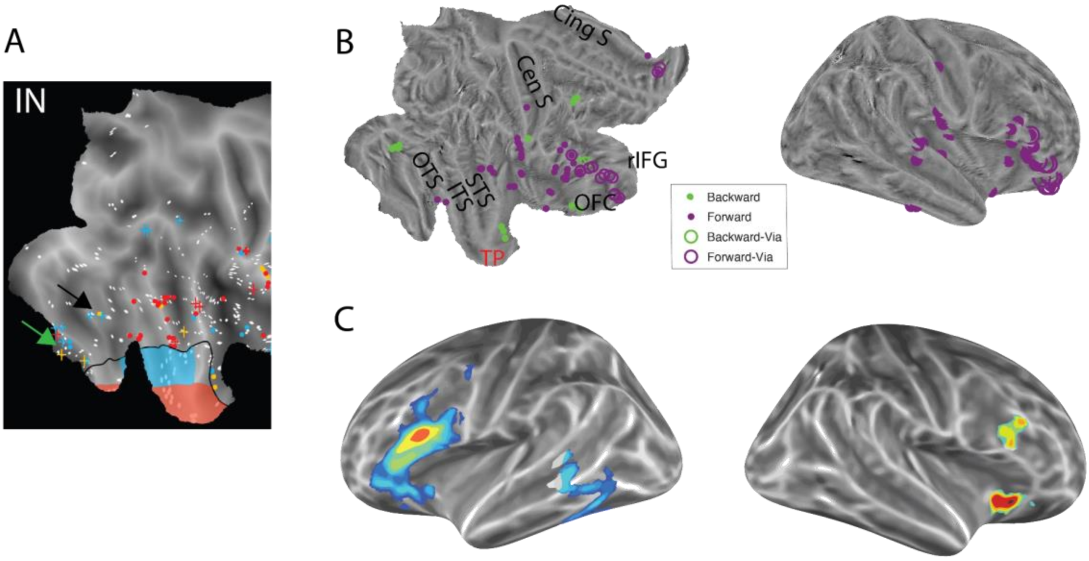
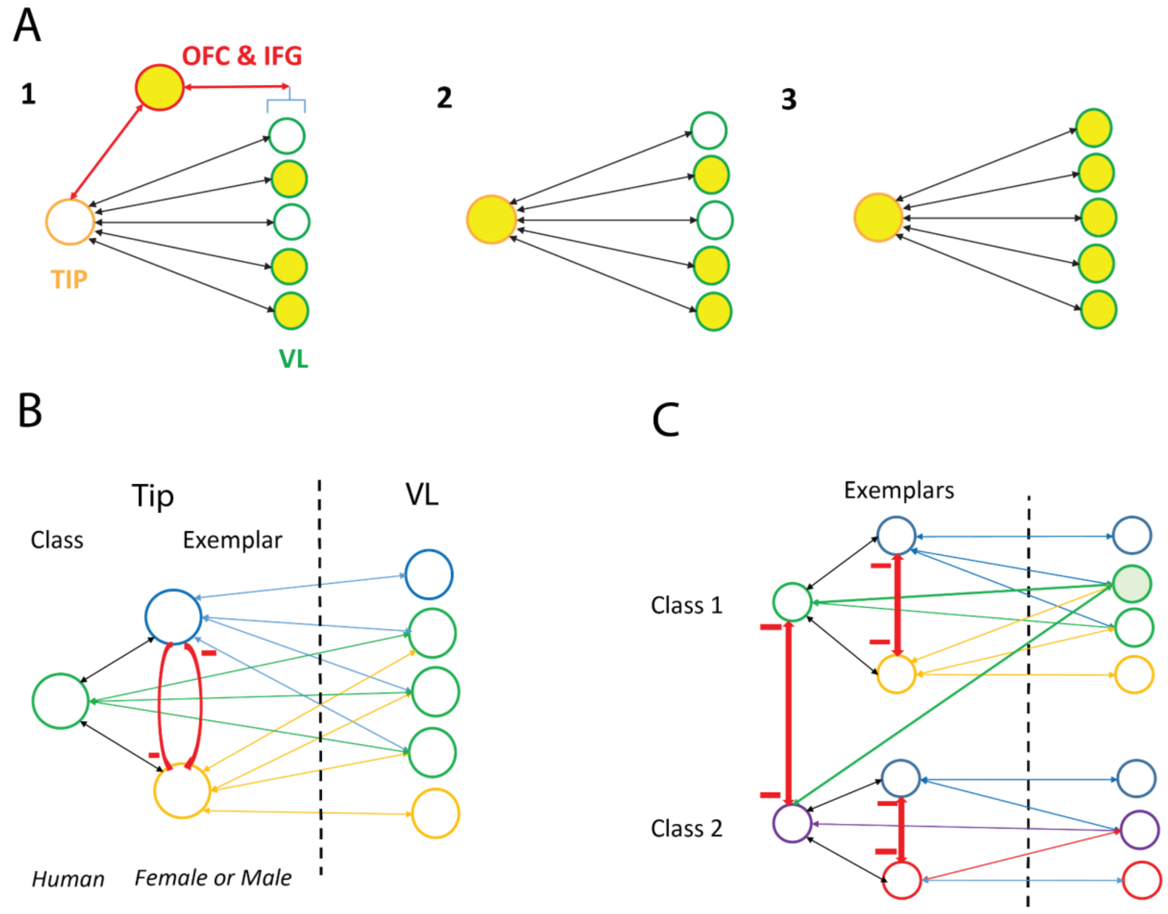


Figure 3: A: The location of leads providing input to VL (blue), Tip (red) or both (yellow) over a short (dots) or long delay (crosses), as determined by T23 using GC. Green and black arrows point to lingual gyrus and FFA respectively. **B**: The location of leads conditioning the short delay GC from VL to Tip (purple), shown on both a flatmap and an inflated right hemisphere (from T23). Via leads (open circles) indicate conditioning leads which also have established connections with each region. TP: temporal pole; OTS: occipito-temporal sulcus; ITS: inferior temporal sulcus; STS: superior temporal sulcus; OFC: orbito-frontal cortex; rIFG: rostral inferior frontal gyrus; Cen S: Central sulcus; Cing S: cingulate sulcus. **C**: The areas consistently implicated in semantic control across both hemispheres from an ALE meta-analysis of neuroimaging studies (Jackson, 2021) projected onto the cortical surface of left and right hemispheres for comparison with B.



Box Figure: A: A putative scheme of how concepts are represented within TP, based on T23. Concepts are represented by the combination of an abstract label in Tip and unimodal features in other TP subregions, including visual features in VL. The 3 panels are a schematic representation of the 3 steps-sequence of activation of the gender concept representation proposed by T23. Orange and green circles:

neural populations in Tip and VL respectively; red circle: control regions; yellow filled circles: active populations. Note that each little circle represents many neurons and that while Tip is proposed to link features, this is much more than a simple linear combination as for example, nonlinear effects can be created by presynaptic inhibition. **B** and **C:** extension of the label-features scheme of T23 to two levels of categories: ordinate (classes) and exemplar levels. The semantic hierarchy arises from connections between label populations in Tip. B: interaction between two exemplars of the same class (e.g., male and female humans); C: interactions between classes (e.g., humans and apes). All connections are proposed to be excitatory except those in red, which are inhibitory links between label populations at the same hierarchical level.

Data and code availability

No new data or code were produced for this manuscript.

References

Abel, T.J., Rhone, A.E., Nourski, K.V., Granner, M.A., Oya, H., Griffiths, T.D., Tranel, D.T., Kawasaki, H. & Howard, M.A. (2014) Mapping the temporal pole with a specialized electrode array: technique and preliminary results. *Physiol Meas*, **35**, 323-337.

Abel, T.J., Rhone, A.E., Nourski, K.V., Kawasaki, H., Oya, H., Griffiths, T.D., Howard, M.A. & Tranel, D. (2015) Direct Physiologic Evidence of a Heteromodal Convergence Region for Proper Naming in Human Left Anterior Temporal Lobe. *J Neurosci*, **35**, 1513-1520.

Adlam, A.L.R., de Haan, M., Hodges, J.R. & Patterson, K. (2013) Memory for action sequences in semantic dementia. *Neuropsychologia*, **51**, 1481-1487.

Aflalo, T., Zhang, C.Y., Rosario, E.R., Pouratian, N., Orbans, G.A. & Andersen, R.A. (2020) A shared neural substrate for action verbs and observed actions in human posterior parietal cortex. *Sci Adv*, **6**.

Anand, S., Cho, H., Adamek, M., Burton, H., Moran, D., Leuthardt, E. & Brunner, P. (2023) High gamma coherence between task-responsive sensory-motor cortical regions in a motor reaction-time task. *J Neurophysiol*, **130**, 628-639.

Bajada, C.J., Haroon, H.A., Azadbakht, H., Parker, G.J.M., Lambon Ralph, M.A. & Cloutman, L.L. (2017) The tract terminations in the temporal lobe: Their location and associated functions. *Cortex*, **97**, 277-290.

Balleine, B.W., Leung, B.K. & Ostlund, S.B. (2011) The orbitofrontal cortex, predicted value, and choice. *Ann Ny Acad Sci*, **1239**, 43-50.

Banaie Boroujeni, K. & Womelsdorf, T. (2023) Routing states transition during oscillatory bursts and attentional selection. *Neuron*, **111**, 2929-2944 e2911.

Bartoldson, B.R., Kailkhura, B. & Blalock, D. (2023) Compute-Efficient Deep Learning: Algorithmic Trends and Opportunities. *J Mach Learn Res*, **24**.

Bausch, M., Niediek, J., Reber, T.P., Mackay, S., Boström, J., Elger, C.E. & Mormann, F. (2021) Concept neurons in the human medial temporal lobe flexibly represent abstract relations between concepts. *Nat Commun*, **12**.

Bender, E.M., Gebru, T., McMillan-Major, A. & Shmitchell, S. (2021) On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? *Proceedings of the 2021 Acm Conference on Fairness, Accountability, and Transparency, Facct 2021*, 610-623.

Binder, J.R. & Desai, R.H. (2011) The neurobiology of semantic memory. *Trends Cogn Sci*, **15**, 527-536.

Binder, J.R., Gross, W.L., Allendorfer, J.B., Bonilha, L., Chapin, J., Edwards, J.C., Grabowski, T.J., Langfitt, J.T., Loring, D.W., Lowe, M.J., Koenig, K., Morgan, P.S., Ojemann, J.G., Rorden, C., Szaflarski, J.P., Tivarus, M.E. & Weaver, K.E. (2011) Mapping anterior temporal lobe language areas with fMRI: A multicenter normative study. *Neuroimage*, **54**, 1465-1475.

Binney, R.J., Embleton, K.V., Jefferies, E., Parker, G.J.M. & Lambon Ralph, M.A. (2010) The Ventral and Inferolateral Aspects of the Anterior Temporal Lobe Are Crucial in Semantic Memory: Evidence from a Novel Direct Comparison of Distortion-Corrected fMRI, rTMS, and Semantic Dementia. *Cereb Cortex*, **20**, 2728-2738.

Binney, R.J., Hoffman, P. & Lambon Ralph, M.A. (2016) Mapping the Multiple Graded Contributions of the Anterior Temporal Lobe Representational Hub to Abstract and Social Concepts: Evidence from Distortion-corrected fMRI. *Cereb Cortex*, **26**, 4227-4241.

Binney, R.J., Parker, G.J.M. & Lambon Ralph, M.A. (2012) Convergent Connectivity and Graded Specialization in the Rostral Human Temporal Lobe as Revealed by Diffusion-Weighted Imaging Probabilistic Tractography. *J Cognitive Neurosci*, **24**, 1998-2014.

Binney, R.J. & Ramsey, R. (2020) Social Semantics: The role of conceptual knowledge and cognitive control in a neurobiological model of the social brain. *Neurosci Biobehav Rev*, **112**, 28-38.

Bouhali, F., Thiebaut de Schotten, M., Pinel, P., Poupon, C., Mangin, J.F., Dehaene, S. & Cohen, L. (2014) Anatomical connections of the visual word form area. *J Neurosci*, **34**, 15402-15414.

Burke, J.F., Ramayya, A.G. & Kahana, M.J. (2015) Human intracranial high-frequency activity during memory processing: neural oscillations or stochastic volatility? *Curr Opin Neurobiol*, **31**, 104-110.

Cao, Y., Summerfield, C. & Saxe, A.M. (2020) Characterizing emergent representations in a space of candidate learning rules for deep networks. *Advances in Neural Information Processing Systems*, **33**.

Catani, M., Jones, D.K., Donato, R. & ffytche, D.H. (2003) Occipito-temporal connections in the human brain. *Brain*, **126**, 2093-2107.

Catani, M., Jones, D.K. & Ffytche, D.H. (2005) Perisylvian language networks of the human brain. *Ann Neurol*, **57**, 8-16.

Chapeton, J.I., Wittig, J.H., Inati, S.K. & Zaghloul, K.A. (2022) Micro-scale functional modules in the human temporal lobe. *Nat Commun*, **13**.

Chen, L., Lambon Ralph, M.A. & Rogers, T.T. (2017) A unified model of human semantic knowledge and its disorders. *Nat Hum Behav*, **1**.

Chen, Y., Shimotake, A., Matsumoto, R., Kunieda, T., Kikuchi, T., Miyamoto, S., Fukuyama, H., Takahashi, R., Ikeda, A. & Lambon Ralph, M.A. (2016) The 'when' and 'where' of semantic coding in the anterior temporal lobe: Temporal representational similarity analysis of electrocorticogram data. *Cortex*, **79**, 1-13.

Clarke, A., Taylor, K.I. & Tyler, L.K. (2011) The Evolution of Meaning: Spatio-temporal Dynamics of Visual Object Recognition. *J Cognitive Neurosci*, **23**, 1887-1899.

Clarke, A. & Tyler, L.K. (2015) Understanding What We See: How We Derive Meaning From Vision. *Trends Cogn Sci*, **19**, 677-687.

Conroy, G. (2023) How ChatGPT and other AI tools could disrupt scientific publishing. *Nature*, **622**, 234-236.

Corbett, F., Jefferies, E., Ehsan, S. & Ralph, M.A.L. (2009) Different impairments of semantic cognition in semantic dementia and semantic aphasia: evidence from the non-verbal domain. *Brain*, **132**, 2593-2608.

Cox, C.R., Rogers, T.T., Shimotake, A., Kikuchi, T., Kunieda, T., Miyamoto, S., Takahashi, R., Matsumoto, R., Ikeda, A. & Lambon Ralph, M.A. (2024) Representational similarity learning reveals a graded multidimensional semantic space in the human anterior temporal cortex. *Imaging Neuroscience*, **2**, 1-22.

Culham, J.C. & Valyear, K.F. (2006) Human parietal cortex in action. *Curr Opin Neurobiol*, **16**, 205-212.

Damasio, H., Grabowski, T.J., Tranel, D., Hichwa, R.D. & Damasio, A.R. (1996) A neural basis for lexical retrieval. *Nature*, **380**, 499-505.

Davey, J., Rueschemeyer, S.A., Costigan, A., Murphy, N., Krieger-Redwood, K., Hallam, G. & Jefferies, E. (2015) Shared neural processes support semantic control and action understanding. *Brain Lang*, **142**, 24-35.

Desai, R.H., Hackett, C.T., Johari, K., Lai, V.T. & Riccardi, N. (2023) Spatiotemporal characteristics of the neural representation of event concepts. *Brain Lang*, **246**.

Devlin, J.T., Russell, R.P., Davis, M.H., Price, C.J., Wilson, J., Moss, H.E., Matthews, P.M. & Tyler, L.K. (2000) Susceptibility-induced loss of signal: Comparing PET and fMRI on a semantic task. *Neuroimage*, **11**, 589-600.

Ding, S.L., Van Hoesen, G.W., Cassell, M.D. & Poremba, A. (2009) Parcellation of Human Temporal Polar Cortex: A Combined Analysis of Multiple Cytoarchitectonic, Chemoarchitectonic, and Pathological Markers. *J Comp Neurol*, **514**, 595-623.

Doerig, A., Sommers, R.P., Seeliger, K., Richards, B., Ismael, J., Lindsay, G.W., Kording, K.P., Konkle, T., van Gerven, M.A.J., Kriegeskorte, N. & Kietzmann, T.C. (2023) The neuroconnectionist research programme. *Nat Rev Neurosci*, **24**, 431-450.

Eggert, G.H. & Wernicke, C. (1977) *Wernicke's works on aphasia : a sourcebook and review*. Mouton, The Hague.

Farahibozorg, S.R., Henson, R.N., Woollams, A.M. & Hauk, O. (2022) Distinct roles for the anterior temporal lobe and angular gyrus in the spatiotemporal cortical semantic network. *Cereb Cortex*, **32**, 4549-4564.

Fayyad, U.M. (2023) From Stochastic Parrots to Intelligent Assistants-The Secrets of Data and Human Interventions. *Ieee Intell Syst*, **38**, 63-67.

Fedorenko, E., Ivanova, A.A. & Regev, T.I. (2024) The language network as a natural kind within the broader landscape of the human brain. *Nat Rev Neurosci*, **25**, 289-312.

Fernandino, L. & Binder, J.R. (2024) How does the "default mode" network contribute to semantic cognition? *Brain Lang*, **252**.

Fernandino, L., Binder, J.R., Desai, R.H., Pendl, S.L., Humphries, C.J., Gross, W.L., Conant, L.L. & Seidenberg, M.S. (2016) Concept Representation Reflects Multimodal Abstraction: A Framework for Embodied Semantics. *Cereb Cortex*, **26**, 2018-2034.

Fernandino, L., Conant, L.L., Binder, J.R., Blindauer, K., Hiner, B., Spangler, K. & Desai, R.H. (2013) Parkinson's disease disrupts both automatic and controlled processing of action verbs. *Brain Lang*, **127**, 65-74.

Frankland, P.W. & Bontempi, B. (2005) The organization of recent and remote memories. *Nat Rev Neurosci*, **6**, 119-130.

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 Cogn Sci*, **27**, 258-281.

Gore, K.R., Woollams, A.M., Bruehl, S., Halai, A.D. & Lambon Ralph, M.A. (2021) Direct Neural Evidence for the Contrastive Roles of the Complementary Learning Systems in Adult Acquisition of Native Vocabulary. *Cereb Cortex*.

Gorno-Tempini, M.L., Hillis, A.E., Weintraub, S., Kertesz, A., Mendez, M., Cappa, S.F., Ogar, J.M., Rohrer, J.D., Black, S., Boeve, B.F., Manes, F., Dronkers, N.F., Vandenberghe, R., Rascovsky, K., Patterson, K., Miller, B.L., Knopman, D.S., Hodges, J.R., Mesulam, M.M. & Grossman, M. (2011) Classification of primary progressive aphasia and its variants. *Neurology*, **76**, 1006-1014.

Güçlü, U. & van Gerven, M.A.J. (2017) Modeling the Dynamics of Human Brain Activity with Recurrent Neural Networks. *Front Comput Neurosc*, **11**.

Hauk, O., Jackson, R.L. & Rahimi, S. (2023) Transforming the neuroscience of language: estimating pattern-to-pattern transformations of brain activity. *Lang Cogn Neurosci*.

He, K.M., Zhang, X.Y., Ren, S.Q. & Sun, J. (2016) Deep Residual Learning for Image Recognition. *Proc Cvpr Ieee*, 770-778.

Herbet, G., Zemmoura, I. & Duffau, H. (2018) Functional Anatomy of the Inferior Longitudinal Fasciculus: From Historical Reports to Current Hypotheses. *Frontiers in Neuroanatomy*, **12**.

Hickok, G. & Poeppel, D. (2007) Opinion - The cortical organization of speech processing. *Nat Rev Neurosci*, **8**, 393-402.

Hodgson, V.J., Lambon Ralph, M.A. & Jackson, R.L. (2023) The cross-domain functional organization of posterior lateral temporal cortex: insights from ALE meta-analyses of 7 cognitive domains spanning 12,000 participants. *Cereb Cortex*, **33**, 4990-5006.

Hoffman, P., Binney, R.J. & Ralph, M.A.L. (2015) Differing contributions of inferior prefrontal and anterior temporal cortex to concrete and abstract conceptual knowledge. *Cortex*, **63**, 250-266.

Humphreys, G.F., Jung, J. & Ralph, M.A.L. (2022) The convergence and divergence of episodic and semantic functions across lateral parietal cortex. *Cereb Cortex*, **32**, 5664-5681.

Humphreys, G.F. & Ralph, M.A. (2015) Fusion and Fission of Cognitive Functions in the Human Parietal Cortex. *Cereb Cortex*, **25**, 3547-3560.

Hutson, M. (2021a) The Language Machines. *Nature*, **591**, 22-25.

Hutson, M. (2021b) Lyin' AIs. *Ieee Spectrum*, **58**, 40-45.

Jackson, R.L. (2021) The neural correlates of semantic control revisited. *Neuroimage*, **224**.

Jackson, R.L., Bajada, C.J., Rice, G.E., Cloutman, L.L. & Lambon Ralph, M.A. (2018) An emergent functional parcellation of the temporal cortex. *Neuroimage*, **170**, 385-399.

Jackson, R.L., Hoffman, P., Pobric, G. & Lambon Ralph, M.A. (2015) The Nature and Neural Correlates of Semantic Association versus Conceptual Similarity. *Cereb Cortex*, **25**, 4319-4333.

Jackson, R.L., Hoffman, P., Pobric, G. & Lambon Ralph, M.A. (2016) The Semantic Network at Work and Rest: Differential Connectivity of Anterior Temporal Lobe Subregions. *J Neurosci*, **36**, 1490-1501.

Jackson, R.L., Rogers, T.T. & Lambon Ralph, M.A. (2021) Reverse-engineering the cortical architecture for controlled semantic cognition. *Nat Hum Behav*, **5**, 774-+.

Jamali, M., Grannan, B., Cai, J., Khanna, A.R., Munoz, W., Caprara, I., Paulk, A.C., Cash, S.S., Fedorenko, E. & Williams, Z.M. (2024) Semantic encoding during language comprehension at single-cell resolution. *Nature*, **631**, 610-616.

Jamali, M., Grannan, B.L., Fedorenko, E., Saxe, R., Báez-Mendoza, R. & Williams, Z.M. (2021) Single-neuronal predictions of others' beliefs in humans. *Nature*, **591**, 610-+.

Jefferies, E. (2013) The neural basis of semantic cognition: Converging evidence from neuropsychology, neuroimaging and TMS. *Cortex*, **49**, 611-625.

Jefferies, E. & Lambon Ralph, M.A. (2006) Semantic impairment in stroke aphasia versus semantic dementia: a case-series comparison. *Brain*, **129**, 2132-2147.

Jung, J. & Lambon Ralph, M.A. (2023) Distinct but cooperating brain networks supporting semantic cognition. *Cereb Cortex*, **33**, 2021-2036.

Jung, J. & Ralph, M.A.L. (2016) Mapping the Dynamic Network Interactions Underpinning Cognition: A cTBS-fMRI Study of the Flexible Adaptive Neural System for Semantics. *Cereb Cortex*, **26**, 3580-3590.

Kalénine, S., Buxbaum, L.J. & Coslett, H.B. (2010) Critical brain regions for action recognition: lesion symptom mapping in left hemisphere stroke. *Brain*, **133**, 3269-3280.

Khanna, A.R., Muñoz, W., Kim, Y.J., Kfir, Y., Paulk, A.C., Jamali, M., Cai, J., Mustroph, M.L., Caprara, I., Hardstone, R., Mejdell, M., Meszéna, D., Zuckerman, A., Schweitzer, J., Cash, S. & Williams, Z.M. (2024) Single-neuronal elements of speech production in humans. *Nature*, **626**, 603-+.

Kietzmann, T.C., Spoerer, C.J., Sorensen, L.K.A., Cichy, R.M., Hauk, O. & Kriegeskorte, N. (2019) Recurrence is required to capture the representational dynamics of the human visual system. *P Natl Acad Sci USA*, **116**, 21854-21863.

Kravitz, D.J., Saleem, K.S., Baker, C.I., Ungerleider, L.G. & Mishkin, M. (2013) The ventral visual pathway: an expanded neural framework for the processing of object quality. *Trends Cogn Sci*, **17**, 26-49.

Kriegeskorte, N. (2015) Deep Neural Networks: A New Framework for Modeling Biological Vision and Brain Information Processing. *Annual Review of Vision Science, Vol 1*, **1**, 417-446.

Kriegeskorte, N. & Kievit, R.A. (2013) Representational geometry: integrating cognition, computation, and the brain. *Trends Cogn Sci*, **17**, 401-412.

Kriegeskorte, N., Mur, M. & Bandettini, P. (2008) Representational similarity analysis - connecting the branches of systems neuroscience. *Front Syst Neurosci*, **2**.

Kuhnke, P., Beaupain, M.C., Arola, J., Kiefer, M. & Hartwigsen, G. (2023) Meta-analytic evidence for a novel hierarchical model of conceptual processing. *Neurosci Biobehav R*, **144**.

Kveraga, K., Ghuman, A.S. & Bar, M. (2007) Top-down predictions in the cognitive brain. *Brain Cognition*, **65**, 145-168.

Lambon Ralph, M.A., Cipolotti, L., Manes, F. & Patterson, K. (2010) Taking both sides: do unilateral anterior temporal lobe lesions disrupt semantic memory? *Brain*, **133**, 3243-3255.

Lambon Ralph, M.A., Ehsan, S., Baker, G.A. & Rogers, T.T. (2012) Semantic memory is impaired in patients with unilateral anterior temporal lobe resection for temporal lobe epilepsy. *Brain*, **135**, 242-258.

Lambon Ralph, M.A., Jefferies, E., Patterson, K. & Rogers, T.T. (2017) The neural and computational bases of semantic cognition. *Nat Rev Neurosci*, **18**, 42-55.

Leonard, M.K., Gwilliams, L., Sellers, K.K., Chung, J.E., Xu, D., Mischler, G., Mesgarani, N., Welkenhuysen, M., Dutta, B. & Chang, E.F. (2023) Large-scale single-neuron speech sound encoding across the depth of human cortex. *Nature*.

Liao, J.J., Li, J.H., Qiu, Y.D., Wu, X.Y., Liu, B.Y., Zhang, L., Zhang, Y.T., Peng, X.Q. & Huang, R.W. (2023) Dissociable contributions of the hippocampus and orbitofrontal cortex to representing task space in a social context. *Cereb Cortex*.

Luccioni, A.S., Viguier, S. & Ligozat, A.L. (2023) Estimating the Carbon Footprint of BLOOM, a 176B Parameter Language Model. *J Mach Learn Res*, **24**.

Manning, J.R., Jacobs, J., Fried, I. & Kahana, M.J. (2009) Broadband Shifts in Local Field Potential Power Spectra Are Correlated with Single-Neuron Spiking in Humans. *J Neurosci*, **29**, 13613-13620.

Marinkovic, K., Dhond, R.P., Dale, A.M., Glessner, M., Carr, V. & Halgren, E. (2003) Spatiotemporal dynamics of modality-specific and supramodal word processing. *Neuron*, **38**, 487-497.

Martin, A. & Chao, L.L. (2001) Semantic memory and the brain: structure and processes. *Curr Opin Neurobiol*, **11**, 194-201.

Martin, A., Haxby, J.V., Lalonde, F.M., Wiggs, C.L. & Ungerleider, L.G. (1995) Discrete Cortical Regions Associated with Knowledge of Color and Knowledge of Action. *Science*, **270**, 102-105.

Matsumoto, R., Nair, D.R., LaPresto, E., Najm, I., Bingaman, W., Shibasaki, H. & Lüders, H.O. (2004) Functional connectivity in the human language system:: a cortico-cortical evoked potential study. *Brain*, **127**, 2316-2330.

McClelland, J.L., Mcnaughton, B.L. & Oreilly, R.C. (1995) Why There Are Complementary Learning-Systems in the Hippocampus and Neocortex - Insights from the Successes and Failures of Connectionist Models of Learning and Memory. *Psychol Rev*, **102**, 419-457.

McClelland, J.L. & Rogers, T.T. (2003) The parallel distributed processing approach to semantic cognition. *Nat Rev Neurosci*, **4**, 310-322.

Mccloskey, M. & Glucksberg, S. (1979) Decision-Processes in Verifying Category Membership Statements - Implications for Models of Semantic Memory. *Cognitive Psychol*, **11**, 1-37.

Mercier, M.R., Dubarry, A.S., Tadel, F., Avanzini, P., Axmacher, N., Cellier, D., Del Vecchio, M., Hamilton, L.S., Hermes, D., Kahana, M.J., Knight, R.T., Llorens, A., Megevand, P., Melloni, L., Miller, K.J., Piai, V., Puce, A., Ramsey, N.F., Schwiedrzik, C.M., Smith, S.E., Stolk, A., Swann, N.C., Vansteensel, M.J., Voytek, B., Wang, L., Lachaux, J.P. & Oostenveld, R. (2022) Advances in human intracranial electroencephalography research, guidelines and good practices. *Neuroimage*, **260**.

Meteyard, L., Cuadrado, S.R., Bahrami, B. & Vigliocco, G. (2012) Coming of age: A review of embodiment and the neuroscience of semantics. *Cortex*, **48**, 788-804.

Mitchell, M. & Krakauer, D.C. (2023) The debate over understanding in AI?s large language models. *P Natl Acad Sci USA*, **120**.

Moran, M.A., Mufson, E.J. & Mesulam, M.M. (1987) Neural Inputs into the Temporopolar Cortex of the Rhesus-Monkey. *J Comp Neurol*, **256**, 88-103.

Murphy, G.L. & Brownell, H.H. (1985) Category differentiation in object recognition: typicality constraints on the basic category advantage. *J Exp Psychol Learn Mem Cogn*, **11**, 70-84.

Nayebi, A., Sagastuy-Brena, J., Bear, D.M., Kar, K., Kubilius, J., Ganguli, S., Sussillo, D., DiCarlo, J.J. & Yamins, D.L.K. (2022) Recurrent Connections in the Primate Ventral Visual Stream Mediate a Tradeoff Between Task Performance and Network Size During Core Object Recognition. *bioRxiv*.

Nelson, M.E. & Bower, J.M. (1990) Brain Maps and Parallel Computers. *Trends Neurosci*, **13**, 403-408.

Noonan, K.A., Jefferies, E., Visser, M. & Lambon Ralph, M.A. (2013) Going beyond Inferior Prefrontal Involvement in Semantic Control: Evidence for the Additional Contribution of Dorsal Angular Gyrus and Posterior Middle Temporal Cortex. *J Cognitive Neurosci*, **25**, 1824-1850.

Olson, I.R., McCoy, D., Klobusicky, E. & Ross, L.A. (2013) Social cognition and the anterior temporal lobes: a review and theoretical framework. *Soc Cogn Affect Neur*, **8**, 123-133.

Olson, I.R., Ploaker, A. & Ezzyat, Y. (2007) The Enigmatic temporal pole: a review of findings on social and emotional processing. *Brain*, **130**, 1718-1731.

Orban, G.A., Ferri, S. & Platonov, A. (2019) The role of putative human anterior intraparietal sulcus area in observed manipulative action discrimination. *Brain Behav*, **9**.

Pascual, B., Masdeu, J.C., Hollenbeck, M., Makris, N., Insausti, R., Ding, S.L. & Dickerson, B.C. (2015) Large-Scale Brain Networks of the Human Left Temporal Pole: A Functional Connectivity MRI Study. *Cereb Cortex*, **25**, 680-702.

Patterson, K., Nestor, P.J. & Rogers, T.T. (2007) Where do you know what you know? The representation of semantic knowledge in the human brain. *Nat Rev Neurosci*, **8**, 976-987.

Paulk, A.C., Kfir, Y., Khanna, A.R., Mustroph, M.L., Trautmann, E.M., Soper, D.J., Stavisky, S.D., Welkenhuysen, M., Dutta, B., Shenoy, K.V., Hochberg, L.R., Richardson, R.M., Williams, Z.M. & Cash, S.S. (2022) Large-scale neural recordings with single neuron resolution using Neuropixels probes in human cortex. *Nat Neurosci*, **25**, 252-+.

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. *Nat Commun*, **9**.

Piantadosi, S.T., Muller, D.C.Y., Rule, J.S., Kaushik, K., Gorenstein, M., Leib, E.R. & Sanford, E. (2024) Why concepts are (probably) vectors. *Trends Cogn Sci*.

Platonov, A., Avanzini, P., Pelliccia, V., LoRusso, G., Sartori, I. & Orban, G.A. (2019) Rapid and specific processing of person-related information in human anterior temporal lobe. *Commun Biol*, **2**.

Plaut, D.C. (2002) Graded modality-specific specialisation in semantics: A computational account of optic aphasia. *Cogn Neuropsychol*, **19**, 603-639.

Pobric, G., Jefferies, E. & Lambon Ralph, M.A. (2007) Anterior temporal lobes mediate semantic representation: Mimicking semantic dementia by using rTMS in normal participants. *P Natl Acad Sci USA*, **104**, 20137-20141.

Premereur, E., Taubert, J., Janssen, P., Vogels, R. & Vanduffel, W. (2016) Effective Connectivity Reveals Largely Independent Parallel Networks of Face and Body Patches. *Curr Biol*, **26**, 3269-3279.

Price, C., Thierry, G. & Griffiths, T. (2005) Speech-specific auditory processing: where is it? *Trends Cogn Sci*, **9**, 271-276.

Price, C.J. (2012) A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. *Neuroimage*, **62**, 816-847.

Pulvermüller, F. (2013) How neurons make meaning: brain mechanisms for embodied and abstract-symbolic semantics. *Trends Cogn Sci*, **17**, 458-470.

Quillian, M.R. (1968) Semantic Networks. In Minsky, M. (ed) *Semantic Information Processing*. MIT Press, Cambridge, Massachusetts.

Quiroga, R.Q. (2012) Concept cells: the building blocks of declarative memory functions. *Nat Rev Neurosci*, **13**, 587-597.

Rahimi, S., Farahibozorg, S.R., Jackson, R. & Hauk, O. (2022) Task modulation of spatiotemporal dynamics in semantic brain networks: An EEG/MEG study. *Neuroimage*, **246**.

Rahimi, S., Jackson, R., Farahibozorg, S.R. & Hauk, O. (2023) Time-Lagged Multidimensional Pattern Connectivity (TL-MDPC): An EEG/MEG pattern transformation based functional connectivity metric. *Neuroimage*, **270**, 119958.

Rajimehra, R., Young, J.C. & Tootell, R.B.H. (2009) An anterior temporal face patch in human cortex, predicted by macaque maps. *P Natl Acad Sci USA*, **106**, 1995-2000.

Ray, S., Crone, N.E., Niebur, E., Franaszczuk, P.J. & Hsiao, S.S. (2008) Neural Correlates of High-Gamma Oscillations (60-200 Hz) in Macaque Local Field Potentials and Their Potential Implications in Electrocorticography. *J Neurosci*, **28**, 11526-11536.

Ray, S. & Maunsell, J.H.R. (2011) Different Origins of Gamma Rhythm and High-Gamma Activity in Macaque Visual Cortex. *Plos Biol*, **9**.

Reilly, J., Shain, C., Borghesani, V., P., K., Vigliocco, G., Peele, J., Mahon, B., Buxbaum, L.J., Majid, M., Brysbaert, M., Borghi, A., DeDeyne, S., Dove, G., Papeo, L., Pexman, P., Poeppel, D., Lupyan, G., Boggio, P., Hickok, G., Gwilliams, L., Fernandino, L., Mirman, D., Chrysikou, E., Sandberg, C., Crutch, S., Pylkanen, L., Yee, E., Jackson, R., Rodd, J., Bedny, M., Connell, L., Kiefer, M., Kemmerer, D., de Zubicaray, G., Jefferies, E., Lynott, D., Siew, C., Desau, R., McRae, K., Diaz, M., Bolognesi, M., Fedorenko, E., Kiran, S., Montefinese, M., Binder, J.R., Yap, M., Hartwigsen, G., Canton, J., Bi, Y.C., Hoffman, P., Garcea, F. & Vinson, D. (2024) What we mean when we say semantic: Toward a multidisciplinary semantic glossary. *Psychonomic Bulletin & Review*.

Rice, G.E., Hoffman, P., Binney, R.J. & Lambon Ralph, M.A. (2018) Concrete versus abstract forms of social concept: an fMRI comparison of knowledge about people versus social terms. *Philos Trans R Soc Lond B Biol Sci*, **373**.

Rice, G.E., Lambon Ralph, M.A. & Hoffman, P. (2015) The Roles of Left Versus Right Anterior Temporal Lobes in Conceptual Knowledge: An ALE Meta-analysis of 97 Functional Neuroimaging Studies. *Cereb Cortex*, **25**, 4374-4391.

Rogers, T.T., Cox, C.R., Lu, Q.H., Shimotake, A., Kikuchi, T., Kunieda, T., 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**.

Rogers, T.T., Lambon Ralph, M.A., Garrard, P., Bozeat, S., McClelland, J.L., Hodges, J.R. & Patterson, K. (2004) Structure and deterioration of semantic memory: A neuropsychological and computational investigation. *Psychol Rev*, **111**, 205-235.

Rogers, T.T. & McClelland, J.L. (2004) *Semantic cognition : a parallel distributed processing approach*. MIT Press, Cambridge, Mass.

Rolls, E.T. (2023) Emotion, motivation, decision-making, the orbitofrontal cortex, anterior cingulate cortex, and the amygdala. *Brain Struct Funct*, **228**, 1201-1257.

Ross, L.A. & Olson, I.R. (2010) Social cognition and the anterior temporal lobes. *Neuroimage*, **49**, 3452-3462.

Rudebeck, P.H. & Murray, E.A. (2011) Balkanizing the primate orbitofrontal cortex: distinct subregions for comparing and contrasting values. *Ann Ny Acad Sci*, **1239**, 1-13.

Sasaki, T., Makris, N., Shenton, M.E., Savadjiev, P., Rathi, Y., Eckbo, R., Bouix, S., Yeterian, E., Dickerson, B.C. & Kubicki, M. (2023) Structural connectivity of cytoarchitectonically distinct human left temporal pole subregions: a diffusion MRI tractography study. *Front Neuroanat*, **17**, 1240545.

Saxe, A.M., McClelland, J.L. & Ganguli, S. (2019) A mathematical theory of semantic development in deep neural networks. *P Natl Acad Sci USA*, **116**, 11537-11546.

Schrimpf, M., Kubilius, J., Hong, H., Majaj, N.J., Rajalingham, R., Issa, E.B., Kar, K., Bashivan, P., Prescott-Roy, J., Geiger, F., Schmidt, K., Yamins, D.L.K. & DiCarlo, J.J. (2020) Brain-Score: Which Artificial Neural Network for Object Recognition is most Brain-Like? *bioRxiv*.

Schuck, N.W., Cai, M.B., Wilson, R.C. & Niv, Y. (2016) Human Orbitofrontal Cortex Represents a Cognitive Map of State Space. *Neuron*, **91**, 1402-1412.

Sharp, D.J., Scott, S.K. & Wise, R.J.S. (2004) Retrieving meaning after temporal lobe infarction: The role of the basal language area. *Ann Neurol*, **56**, 836-846.

Shimotake, A., Matsumoto, R., Ueno, T., Kunieda, T., Saito, S., Hoffman, P., Kikuchi, T., Fukuyama, H., Miyamoto, S., Takahashi, R., Ikeda, A. & Lambon Ralph, M.A. (2015) Direct Exploration of the Role of the Ventral Anterior Temporal Lobe in Semantic

Memory: Cortical Stimulation and Local Field Potential Evidence From Subdural Grid Electrodes. *Cereb Cortex*, **25**, 3802-3817.

Snowden, J.S., Goulding, P.J. & Neary, D. (1989) Semantic dementia: a form of circumscribed cerebral atrophy. *Behav. Neurol.* , **2**, 167-182.

Spitsyna, G., Warren, J.E., Scott, S.K., Turkheimer, F.E. & Wise, R.J.S. (2006) Converging language streams in the human temporal lobe. *J Neurosci*, **26**, 7328-7336.

Thompson, H.E., Henshall, L. & Jefferies, E. (2016) The role of the right hemisphere in semantic control: A case-series comparison of right and left hemisphere stroke. *Neuropsychologia*, **85**, 44-61.

Thompson, H.E., Robson, H., Lambon Ralph, M.A. & Jefferies, E. (2015) Varieties of semantic 'access' deficit in Wernicke's aphasia and semantic aphasia. *Brain*, **138**, 3776-3792.

Thorat, S., Aldegheri, G. & Kietzmann, T.C. (2022) Category-orthogonal object features guide information processing in recurrent neural networks trained for object categorization. *arXiv*.

Thorpe, S., Fize, D. & Marlot, C. (1996) Speed of processing in the human visual system. *Nature*, **381**, 520-522.

Tiesinga, P., Platonov, A., Pelliccia, V., LoRusso, G., Sartori, I. & Orban, G.A. (2023) Uncovering the fast, directional signal flow through the human temporal pole during semantic processing. *Sci Rep-Uk*, **13**.

Urgen, B.A. & Orban, G.A. (2021) The unique role of parietal cortex in action observation: Functional organization for communicative and manipulative actions. *Neuroimage*, **237**.

van Kerkoerle, T., Self, M.W., Dagnino, B., Gariel-Mathis, M.A., Poort, J., van der Togt, C. & Roelfsema, P.R. (2014) Alpha and gamma oscillations characterize feedback and feedforward processing in monkey visual cortex. *Proc Natl Acad Sci U S A*, **111**, 14332-14341.

Visser, M., Embleton, K.V., Jefferies, E., Parker, G.J. & Lambon Ralph, M.A. (2010a) The inferior, anterior temporal lobes and semantic memory clarified: Novel evidence from distortion-corrected fMRI. *Neuropsychologia*, **48**, 1689-1696.

Visser, M., Jefferies, E., Embleton, K.V. & Lambon Ralph, M.A. (2012) Both the Middle Temporal Gyrus and the Ventral Anterior Temporal Area Are Crucial for Multimodal Semantic Processing: Distortion-corrected fMRI Evidence for a Double Gradient of Information Convergence in the Temporal Lobes. *J Cognitive Neurosci*, **24**, 1766-1778.

Visser, M., Jefferies, E. & Lambon Ralph, M.A. (2010b) Semantic Processing in the Anterior Temporal Lobes: A Meta-analysis of the Functional Neuroimaging Literature. *J Cognitive Neurosci*, **22**, 1083-1094.

Visser, M. & Lambon Ralph, M.A. (2011) Differential Contributions of Bilateral Ventral Anterior Temporal Lobe and Left Anterior Superior Temporal Gyrus to Semantic Processes. *J Cognitive Neurosci*, **23**, 3121-3131.

Von Der Heide, R.J., Skipper, L.M., Klobusicky, E. & Olson, I.R. (2013) Dissecting the uncinate fasciculus: disorders, controversies and a hypothesis. *Brain*, **136**, 1692-1707.

von Seth, J., Nicholls, V.I., Tyler, L.K. & Clarke, A. (2023) Recurrent connectivity supports higher-level visual and semantic object representations in the brain. *Commun Biol*, **6**, 1207.

Walton, M.E., Behrens, T.E.J., Noonan, M.P. & Rushworth, M.F.S. (2011) Giving credit where credit is due: orbitofrontal cortex and valuation in an uncertain world. *Ann Ny Acad Sci*, **1239**, 14-24.

Wang, X.S., Wu, W., Ling, Z.H., Xu, Y.W., Fang, Y.X., Wang, X.Y., Binder, J.R., Men, W.W., Gao, J.H. & Bi, Y.C. (2018) Organizational Principles of Abstract Words in the Human Brain. *Cereb Cortex*, **28**, 4305-4318.

Wang, Y., Collins, J.A., Koski, J., Nugiel, T., Metoki, A. & Olson, I.R. (2017) Dynamic neural architecture for social knowledge retrieval. *Proc Natl Acad Sci U S A*, **114**, E3305-E3314.

Warrington, E.K. (1975) The selective impairment of semantic memory. *Q J Exp Psychol*, **27**, 635-657.

Wilson, R.C., Takahashi, Y.K., Schoenbaum, G. & Niv, Y. (2014) Orbitofrontal Cortex as a Cognitive Map of Task Space. *Neuron*, **81**, 267-279.

Xie, W.Z., Wittig, J.J.r., Chapeton, J.I., El-Kalliny, M., Jackson, S.N., Inati, S.K. & Zaghloul, K.A. (2024) Neuronal sequences in population bursts encode information in human cortex. *Nature*.

Yamins, D.L.K. & DiCarlo, J.J. (2016) Using goal-driven deep learning models to understand sensory cortex. *Nat Neurosci*, **19**, 356-365.

Zahn, R., Moll, J., Krueger, F., Huey, E.D., Garrido, G. & Grafman, J. (2007) Social concepts are represented in the superior anterior temporal cortex. *Proc Natl Acad Sci U S A*, **104**, 6430-6435.

Zheng, J., Schjetnan, A.G.P., Yebra, M., Gomes, B.A., Mosher, C.P., Kalia, S.K., Valiante, T.A., Mamelak, A.N., Kreiman, G. & Rutishauser, U. (2022) Neurons detect cognitive boundaries to structure episodic memories in humans. *Nat Neurosci*, **25**, 358-+.

Zhou, H.H., Schafer, R.J. & Desimone, R. (2016) Pulvinar-Cortex Interactions in Vision and Attention. *Neuron*, **89**, 209-220.

Zhu, Y.M., Xu, M., Lu, J.F., Hu, J.H., Kwoks, V.P.Y., Zhou, Y.L., Yuan, D., Wu, B., Zhang, J., Wu, J.S. & Tan, L. (2022) Distinct spatiotemporal patterns of syntactic and semantic processing in human inferior frontal gyrus. *Nat Hum Behav*, **6**, 1104-1111.

Figure 1

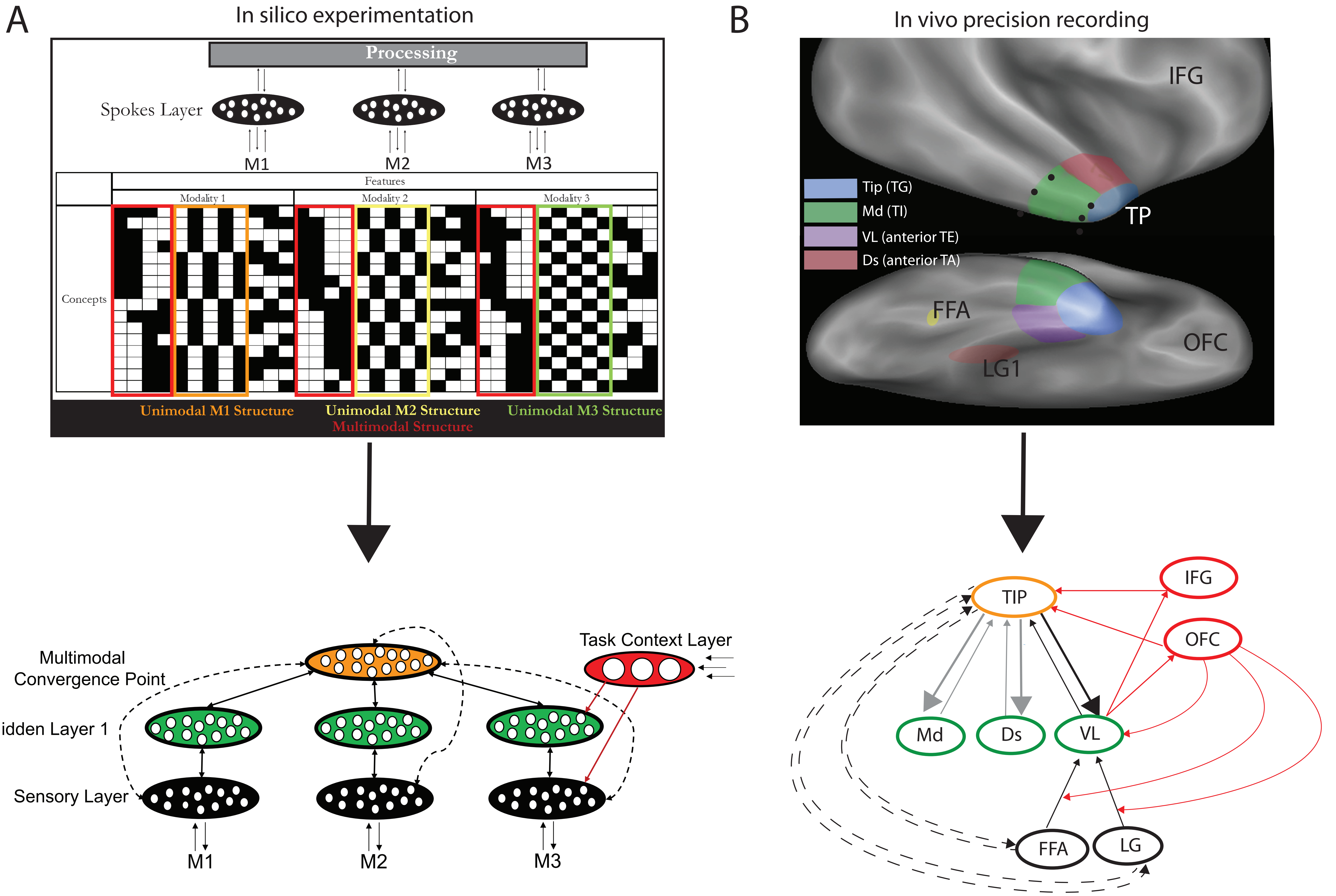

A
In silico experimentation
Processing
Spokes Layer
M1
M2
M3
Features
Modality 1
Modality 2
Modality 3
Concepts
Unimodal M1 Structure
Unimodal M2 Structure
Unimodal M3 Structure
Multimodal Structure
Multimodal Convergence Point
Task Context Layer
Hidden Layer 1
Sensory Layer
M1
M2
M3
B
In vivo precision recording
IFG
Tip (TG)
Md (TI)
VL (anterior TE)
Ds (anterior TA)
TP
FFA
OFC
LG1
TIP
IFG
OFC
Md
Ds
VL
FFA
LG

Figure 2

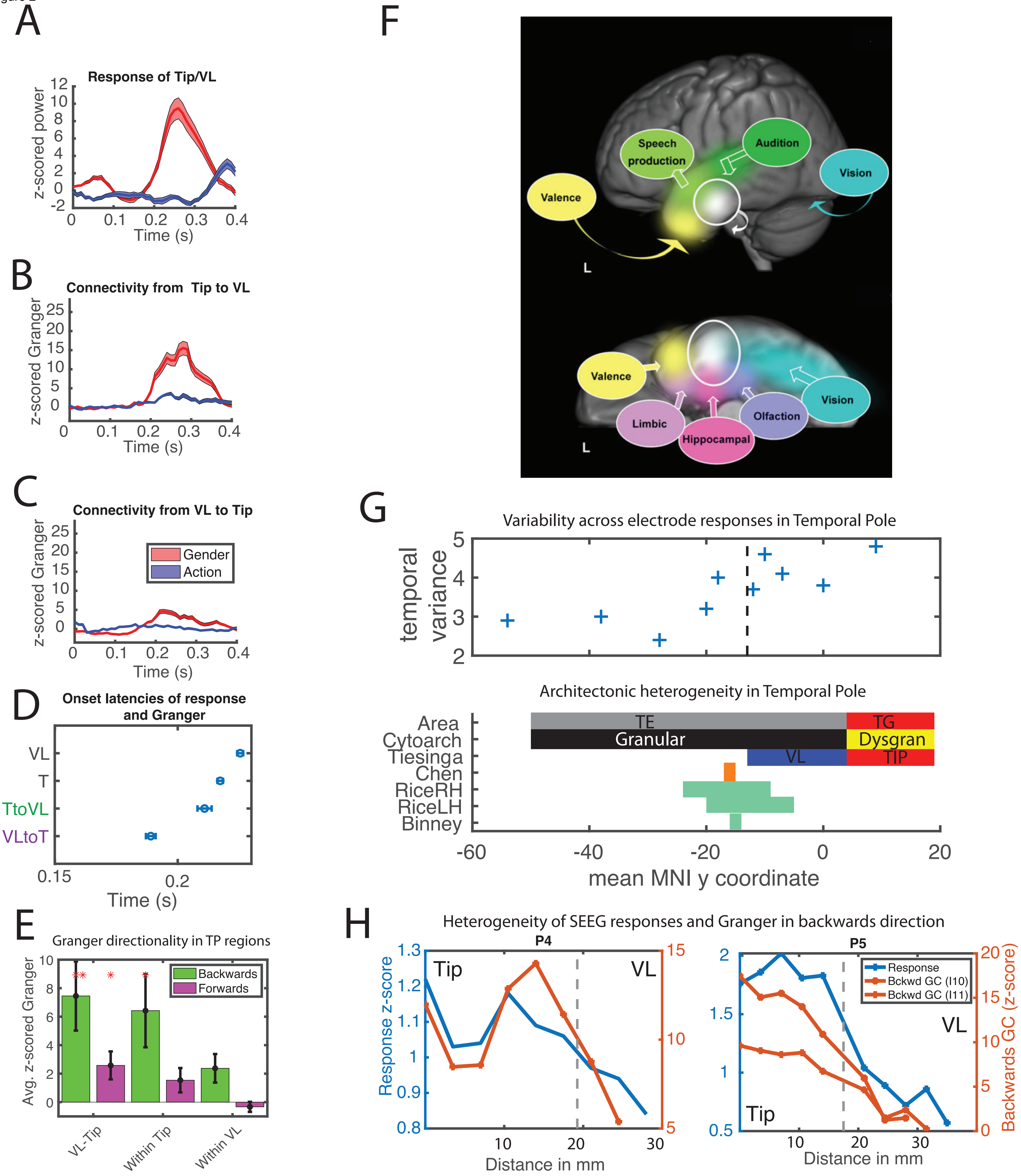
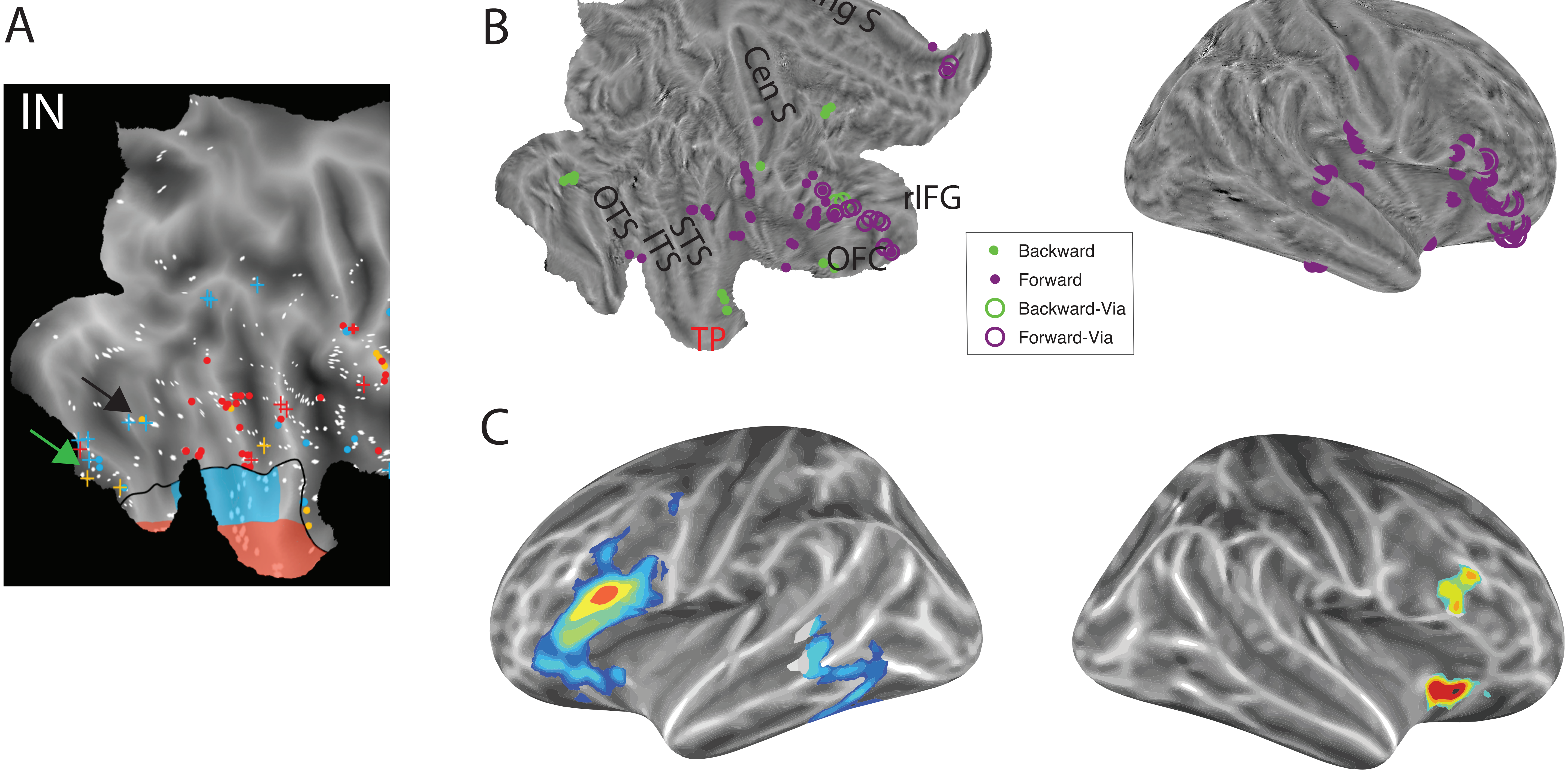
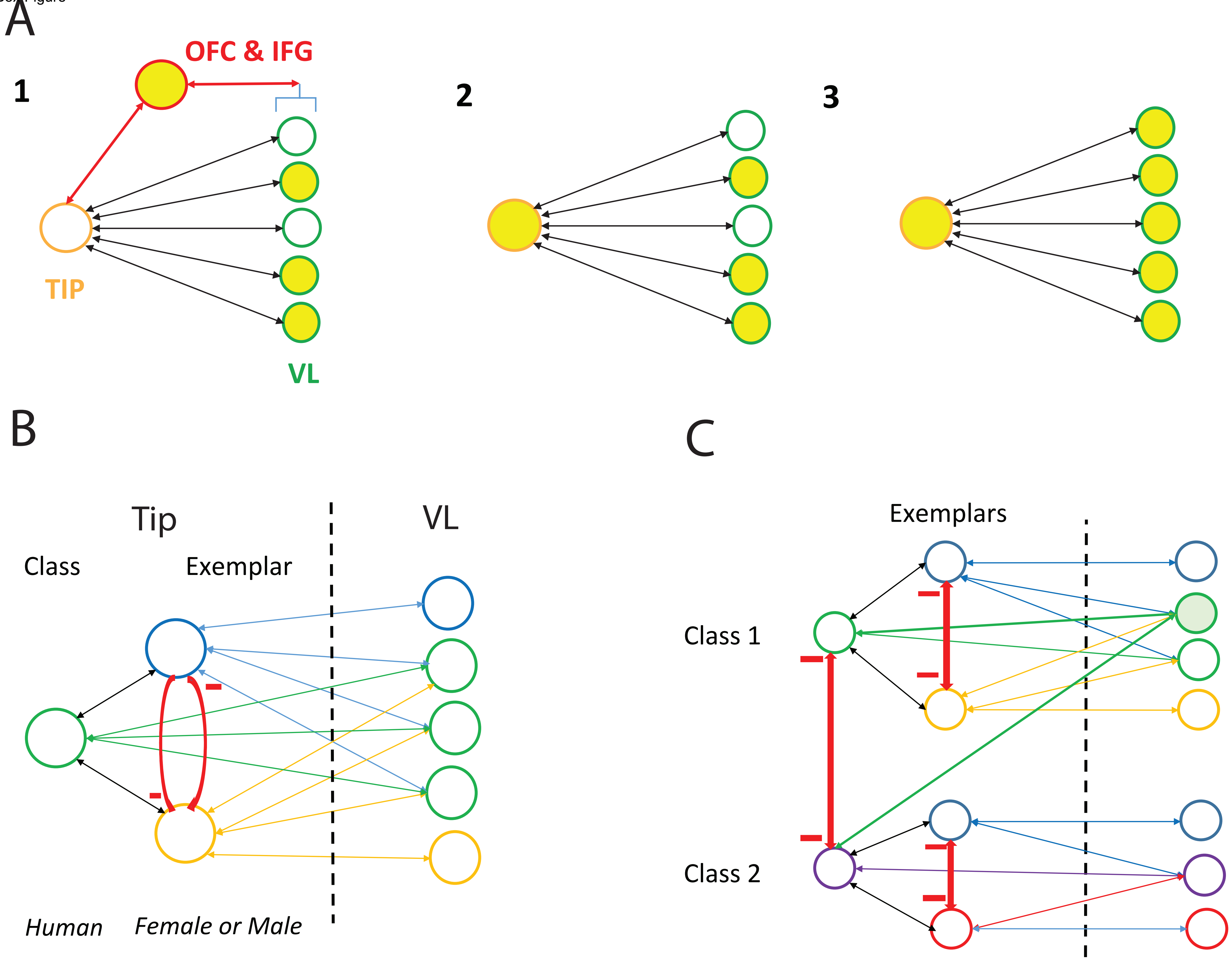

A
Response of Tip/VL
z-scored power
12
10
8
6
4
2
0
-2
0
0.1
0.2
0.3
0.4
Time (s)
B
Connectivity from Tip to VL
z-scored Granger
25
20
15
10
5
0
0
0.1
0.2
0.3
0.4
Time (s)
C
Connectivity from VL to Tip
z-scored Granger
25
20
15
10
5
0
Gender
Action
0
0.1
0.2
0.3
0.4
Time (s)
D
Onset latencies of response and Granger
VL
T
TtoVL
VLtoT
0.15
0.2
Time (s)
E
Granger directionality in TP regions
Avg. z-scored Granger
10
8
6
4
2
0
**
*
*
Backwards
Forwards
VL-Tip
Within Tip
Within VL
F
Speech production
Audition
Vision
Valence
L
Valence
Vision
Limbic
Hippocampal
Olfaction
L
G
Variability across electrode responses in Temporal Pole
temporal variance
5
4
3
2
Architectonic heterogeneity in Temporal Pole
Area
Cytoarch
Tiesinga
Chen
RiceRH
RiceLH
Binney
TE
TG
Granular
Dysgran
VL
TIP
-60
-40
-20
0
20
mean MNI y coordinate
H
Heterogeneity of SEEG responses and Granger in backwards direction
P4
Response z-score
1.3
1.2
1.1
1
0.9
0.8
Tip
VL
15
10
5
10
20
30
Distance in mm
P5
2
1.5
1
0.5
Response
Bckwd GC (I10)
Bckwd GC (I11)
VL
Tip
20
15
10
5
0
Backwards GC (z-score)
10
20
30
Distance in mm

Figure 3


A
IN
B
Cing S
Cen S
OTS
STS
ITS
TP
rIFG
OFC
● Backward
● Forward
○ Backward-Via
○ Forward-Via
C

Box Figure


A
1
OFC & IFG
TIP
VL
2
3
B
Tip
VL
Class
Exemplar
Human
Female or Male
C
Exemplars
Class 1
Class 2

Please sign two copies, retain one for your files and return the other to the editorial office.

Publication Agreement

Agreement: We are pleased to have the privilege of publishing your Article in a forthcoming issue of *Neurobiology of Language*, an open access journal. By your signature below, you hereby grant all your right, title, and interest in the Article to the Society for the Neurobiology of Language (SNL) and Massachusetts Institute of Technology (MIT). By assigning copyright for your work to MIT, The MIT Press will be able to fully disseminate your work while protecting it from inappropriate use.

Open Access Publication of your Article: Your Article will be published under the Creative Commons CC BY attribution license in *Neurobiology of Language*: https://direct.mit.edu/nol. As such, you are free to: 1.) copy and redistribute the material in any medium or format for any purpose, including commercial use, 2.) remix, transform, and build upon the material for any purpose, provided that full attribution of the source is included.

Rights of the Author: Furthermore, you may 1.) photocopy the Article for use in your own teaching activities, 2.) publish the Article, or permit it to be published, as a part of any book you may write, or in any anthology of which you are an editor, in which the Article is included or which expands or elaborates on the Article, 3.) self-archive the Article on your website and institutional repositories in accordance with the CC BY license described at the guidelines found at https://creativecommons.org/licenses/by/4.0/ . As a condition of reserving this right, you agree that The MIT Press and *Neurobiology of Language* will be given first publication credit, and proper copyright notice will be displayed on the work (both on the work as a whole and, where applicable, on the Article as well) whenever such publication occurs.

Rights of MIT Press: This agreement means that MIT Press will have, in all media, the following rights among others: 1.) to license abstracts, quotations, extracts, reprints, and/or translations of the work for publication, 2.) to license reprints of the Article to third persons for educational photocopying, 3.) to license others to create abstracts of the Article, 4.) to license secondary publishers to reproduce the Article in print, microform, or any computer readable form including electronic on-line databases. This also includes licensing the Article for document delivery.

Warranties: You warrant that: (a) the Article has not been published before in any form, (b) you have made no license or other transfer to anyone with respect to your copyright in it, (c) you are its sole author (or that you and the coauthors listed on the article are its sole authors), (d) the Article was not substantially generated by a large language model (LLM) or generative artificial intelligence (GAI) and, to the extent LLM or GAI was used in connection with the creation of your Article, all of the aforementioned warranties still apply, and (e) generally, you have all the necessary rights to make the grants you make to us. Any exceptions are to be noted below. You also warrant that the Article does not violate anyone's copyright, right of privacy, or other right whatsoever of any third party, and is not defamatory of any third party or otherwise unlawful in any respect. You agree to indemnify us against any claim or action alleging facts, which, if true, constitute a breach of any of the foregoing warranties.

3rd Party Copyrighted Content: If your article contains third-party copyrighted material (images, illustrations, etc.) that you do not own copyright to, please check this box ☑ to confirm that you have obtained and submitted a copy of the required releases. You must include proper copyright notice as required by the original copyright holder. Unless you have provided a copyright notice or credit line that attributes the content to another copyright holder, the terms of this agreement will govern the content as well.

Use of LLM and GAI: If, and to the extent, you use LLM or GAI in connection with the creation of your Article, you will notify the Press and describe with specificity the nature of such use. Further, you agree to cooperate with the Press to answer any questions concerning aligning such use with the generally accepted principles of scholarly research and publishing standards. You understand that the Press may, in its reasonable discretion, request that you remove or adjust portions of your Article that were created in connection with the use of LLM or GAI.

Supplementary Material Agreement: If you are submitting supplementary material, check here ☐ and return an initialed page 3.

Multiple Authors: If there is more than one author of the Article, the word “you” includes all authors jointly and severally. The corresponding author may sign on behalf of all authors if he or she has the authority to act as their agent. Please check off the “for all authors” box if it is applicable.

Publication Ethics: You acknowledge that you have read and understood the MIT Press Journal Publication Ethics (found at https://direct.mit.edu/journals/pages/publication-ethics), which is applicable to you as an Article contributor to an MIT Press publication.

Please sign two copies, retain one for your files and return the other to the editorial office.

Concerning Promotional Material: If you wish to give *Neurobiology of Language* permission to use your illustrations in promotional materials for the journal, please check this box ☑. In the event that one of your illustrations is used for this purpose, you and the artwork will be appropriately credited.

Concerning U.S. (Federal) Government Employees: Some of the foregoing grants and warranties will not apply if the Article was written by U.S. Government employees acting within the scope of their employment. If you are a U.S. Government employee who prepared this work as part of your official duties and there is no copyright to transfer, please check this box ☐.

Electronic signature: You agree that this agreement may be signed with an electronic signature, that an electronic signature shall be valid and binding for all purposes, and hereby waive any objection to use of an electronic version of this agreement as a substitute for the original for any legally recognized purpose.

In Conclusion: This is the entire agreement between you and us, and it may be modified only in writing. It will bind and benefit our respective successors in interest. It will terminate if we do not publish your Article within two years of the date of your signature(s).

I (we) concur in this letter of agreement: [☑ for all authors] The MIT Press hereby acknowledge their consent to the terms of the foregoing agreement.

Rebecca Jackson	UK
Print Name	Citizenship

	24/3/25	
Signature	Date	Nick Lindsay, Journals Director, The MIT Press

Controlled Semantic Cognition: Precision Recordings converge with in silico Experiments to Reveal the Inner Workings of the Anterior Temporal Lobe Hub	
Article Title	Manuscript #

Please initial two copies. Retain one for your records and return the other to the editorial office.

Publication Agreement
Supplementary Material for *Neurobiology of Language* Articles

Supplementary Material Agreement: If you are submitting additional text, video, or other material (Supplementary Material) now or in the future, that you consider relevant to your article, the following three clauses will govern the posting of that material.

You hereby give to MIT Press a nonexclusive license to publish the Supplementary Material by hosting it on the Press's servers, accessible through a link from the Press's web page where your article is listed, and from any other MIT Press website or related website. You understand that the Supplementary Material has not been reviewed or approved by the editor of *Neurobiology of Language* is not considered part of your Article or of *Neurobiology of Language,* and is not included in any licensing arrangements that the Press may have with respect to *Neurobiology of Language.*

You are free to include your personal contact information and any copyright or other notice that you wish to include in the Supplementary Material, provided it is not in the nature of an advertisement or solicitation of employment or business. MIT Press reserves the right to take down the Supplementary Material at any time if in its sole judgment it believes that the Supplementary Material is not suitable for posting on its website.

You warrant that you are the sole author of the Supplementary Material, or if not, that you have a complete and unencumbered right to grant MIT Press the license stated above. You also warrant that the Supplementary Material does not libel anyone, invade anyone's copyright or otherwise violate any statutory or common law right of anyone. You agree to indemnify us against any claim or action alleging facts which, if true, constitute a breach of any of the foregoing warranties.

I concur in this Supplementary Materials Agreement: Initials here: ____________

____
Print Name

____
Article Title Manuscript #