

Published in final edited form as:

*Nat Hum Behav.* 2017 March ; 1(3): . doi:10.1038/s41562-016-0039.

## A unified model of human semantic knowledge and its disorders

Lang Chen<sup>1,3,\*</sup>, Matthew A. Lambon Ralph<sup>2,\*</sup>, and Timothy T. Rogers<sup>1,\*</sup>

<sup>1</sup>Department of Psychology, University of Wisconsin-Madison, 1202 West Johnson Street, Madison, WI 53705, USA

<sup>2</sup>Neuroscience and Aphasia Research Unit (NARU), Division of Neuroscience & Experimental Psychology, School of Biological Sciences, University of Manchester, M13 9PL, UK

<sup>3</sup>Stanford Cognitive and Systems Neuroscience Laboratory, 1070 Arastradero Rd. Suite 220, Palo Alto, CA 94304, USA

### Summary

How is knowledge about the meanings of words and objects represented in the human brain? Current theories embrace two radically different proposals: either distinct cortical systems have evolved to represent different kinds of things, or knowledge for all kinds is encoded within a single domain-general network. Neither view explains the full scope of relevant evidence from neuroimaging and neuropsychology. Here we propose that graded category-specificity emerges in some components of the semantic network through joint effects of learning and network connectivity. We test the proposal by measuring connectivity amongst cortical regions implicated in semantic representation, then simulating healthy and disordered semantic processing in a deep neural network whose architecture mirrors this structure. The resulting neuro-computational model explains the full complement of neuroimaging and patient evidence adduced in support of both domain-specific and domain-general approaches, reconciling long-standing disputes about the nature and origins of this uniquely human cognitive faculty.

---

Semantic memory supports the human ability to infer important but unobserved states of affairs in the world, such as object names (“that’s a mushroom”), properties (“it is poisonous”), predictions (“it appears in autumn”), and the meaning of statements (“it is edible after cooking”). Such inferences are generated within a cross-modal cortical network that encodes relationships amongst perceptual, motor, and linguistic representations of objects, actions, and statements (henceforth surface representations<sup>1</sup>). The large-scale architecture and organizational principles of the semantic network remain poorly understood, however. Theories about the nature and structure of this network have long been caught between two proposals: (a) the system is modular and domain-specific, with

---

**Materials and Correspondence** Correspondence and requests for materials should be addressed to Lang Chen (lchen32@stanford.edu), Matthew A. Lambon Ralph (matt.lambon-ralph@manchester.ac.uk), or Timothy T. Rogers (trogers@wisc.edu).

#### Author contributions

All authors contributed to the entire process of this project, including project planning, experiment work, data analysis, and writing the paper.

#### Competing financial interests

The authors declare no competing financial interests.

components that have evolved to support different knowledge domains<sup>2,3</sup>, e.g. animals, tools, people, etc., or (b) it is interactive and domain-general, with all components contributing to all knowledge domains<sup>4–6</sup>. Despite profoundly different implications about the nature and roots of human cognition, these views have proven difficult to adjudicate<sup>3,7</sup>.

We consider a third proposal which arises from a general approach to functional specialization in the brain that we call *connectivity-constrained cognition* -  $C^3$  for convenience. This view proposes that functional specialization in the cortex is jointly caused by (1) learning/experience, (2) perceptual, linguistic, and motor structures in the environment and (3) anatomical connectivity in the brain. Connectivity is important because, within a given neuro-cognitive network, robustly connected components exert strong mutual influences and so, following learning, come to respond similarly to various inputs. In the case of semantic representation, these factors suggest a new approach that reconciles domain-specific and domain-general views. Specifically, learning, environmental structure, and connectivity together produce graded domain-specificity in some network components because conceptual domains differ in the surface representations they engage<sup>8–10</sup>. For instance, tools engage praxis more than animals<sup>11</sup> so regions that interact with action systems come to respond more to tool stimuli. Yet such effects emerge through domain-general learning of environmental structure, and centrally-connected network components contribute critically to all semantic domains<sup>12,13</sup>.

This  $C^3$  proposal coheres with those of several other groups<sup>8,14–17</sup>, but its potential to reconcile divergent views remains unclear because prior studies have focused on fairly specific questions about local network organization. The current paper tests the proposal by first measuring the anatomical connectivity of a broad cortical semantic network, and then assessing the consequences of that connectivity for healthy and disordered network behavior using simulations with a deep neural network model. Specifically, from a new literature review and meta-analysis of functional brain imaging studies we delineated cortical regions involved in semantic representation of words and visually-presented objects and identified those showing systematic semantic category effects. We then measured white-matter tracts connecting these regions using probabilistic diffusion-weighted tractography, resulting in a new characterization of cortical semantic network connectivity. From these results we constructed a deep neural network model and trained it to associate surface representations of objects: their visual structure, associated functions and praxis, and words used to name or describe them. The resulting model is able to explain evidence adduced in support of both domain-specific and domain-general theories, including (a) patterns of functional activation in brain imaging studies, (b) impairments observed in the primary disorders of semantic representation, and (c) the anatomical bases of these disorders.

## Activation likelihood estimate (ALE) analysis

Prior empirical, modeling, and neuroimaging work (SI-Discussion 1) has identified several cortical regions that contribute to semantic processing and their respective functional roles, including: (1) the posterior fusiform gyrus (pFG), which encodes visual representations of objects<sup>18,19</sup>; (2) the superior temporal gyrus (STG), which encodes auditory representations of speech<sup>20</sup>; (3) lateral parietal cortex, which encodes representations of object function and

praxis<sup>19,21,22</sup>; and (4) the ventral anterior temporal lobe (ATL), thought to serve as a cross-modal hub that encodes semantic similarity structure<sup>23,24</sup>. To assess which of these regions show reliable semantic category sensitivity, and to identify additional category-sensitive regions not included among these, we conducted an ALE meta-analysis of functional imaging studies seeking semantic category effects. ALE provides a way of statistically assessing which category effects are reliably observed in the same location across studies. Like a prior meta-analysis<sup>25</sup>, we included studies of activations generated by words or pictures denoting animals or artifacts (manmade objects). We identified 49 studies<sup>9,19,21,26–71</sup> with 73 independent experiments and 270 foci, making this the largest such analysis to date (details in Methods). Using recently updated ALE methods<sup>72</sup>, we tested for cortical regions showing systematically different patterns for animals versus artifacts, or systematically elevated responses for both domains relative to baseline (see Table S1). Results are shown in Figure 1 and Figure S1.

### Regions identified in prior work

Medial pFG is activated more for artifacts than animals bilaterally. Lateral pFG is activated above baseline for animals but not artifacts in both hemispheres, though the animal vs. artifact contrast was only significant in the right, possibly because the homologous left-hemisphere region is “sandwiched” between two areas showing the reverse pattern (pMTG and medial pFG; see Fig. S2). The differential engagement of lateral/medial pFG by animal/artifact is well documented and typically thought to be bilateral<sup>73</sup>.

STG did not show reliable category effects, consistent with the view from prior models<sup>4,12,74</sup> that it processes spoken word input and so should be equally engaged by animals and artifacts.

Ventral ATL did not exhibit activations above baseline for either domain, though this is not surprising for methodological reasons established in prior work<sup>4,75</sup>. Converging evidence from patient studies<sup>76</sup>, brain imaging with appropriate methodology<sup>4,75</sup>, transcranial magnetic stimulation<sup>23</sup>, electro-corticography<sup>77</sup>, and lesion-symptom mapping<sup>78</sup> have established the importance of ventral ATL for domain-general semantic processing. Prior models<sup>5,12,79</sup> included ventral ATL as a cross-modal semantic hub (see SI-Discussion 1.4 and 2.3).

### Regions not specified or included in prior work

In the left parietal lobe, artifacts produced more activations than animals, consistent with the proposal that this region encodes representations of object-directed action<sup>19,80</sup>. One prior model incorporated function representations in the lateral parietal cortex<sup>12</sup>. The cluster spanned inferior and superior parietal lobes (IPL and SPL), which patient and imaging literatures suggest encode different aspects of action knowledge<sup>80</sup>. Thus we included both as separate regions of interest in the connectivity analysis and the model.

Posterior middle temporal gyrus (pMTG) exhibited more activations for artifacts than animals consistent with the literature implicating this region in the semantic representation of tools<sup>25,73</sup>. Accordingly, we included pMTG as a region of interest in the connectivity analysis and the model.

Lateral occipital complex (LOC) activated more often for animals than artifacts, which probably reflects domain differences in visual structure including greater complexity and more overlap among animals relative to manmade objects<sup>81</sup>. We thus identify LOC as a source of visual input to inferotemporal cortex, and assume that animals generate more activation here because they have richer and more overlapping visual representations.

## Semantic network connectivity

We next measured white-matter connectivity amongst all temporal regions of interest, and between temporal and parietal regions, using probabilistic diffusion-weighted tractography. We did not investigate intra-lobe connectivity within the parietal cortex<sup>82,83</sup>, since these areas contribute to other non-semantic cognitive and perceptual abilities beyond the scope of this study. Diffusion-weighted images were collected from 24 participants using methods optimized to reduce the susceptibility artifact in entral ATL<sup>84</sup>. Seeds were placed in the white matter underlying the regions of interest from the ALE analysis or the literature (Fig. 1 and 2; for ROI definition, see Methods), mapped back to native space for each subject<sup>85</sup>. STG and LOC were excluded from the analysis since their connectivity is well-studied<sup>86,87</sup> and they are posited to provide spoken-word and visual input, respectively, to the semantic network. Results are shown in Figure 2.

### Intratemporal connections

Both lateral and medial pFG projected into ATL (> 5% in more than two thirds of the participants) and to one another (Fig.2A and Table S2; for thresholding, see Methods). ATL also projected to both pFG regions and to the pMTG (>2.5% in more than half participants). Streamlines from pMTG terminated in the ATL neighbourhood (yellow stream in Fig.2B) and projected to lateral pFG with high probability and to medial pFG with moderate probability (> 1% in more than half participants).

### Temporo-parietal connections

Streamlines from the ATL did not extend into parietal cortex as also found previously<sup>83,88</sup>. Streamlines from pMTG, however, projected both to ATL and to IPL, providing an indirect route from IPL to the ATL via pMTG (Fig.2B). Likewise, the IPL streamlines projected to pMTG but not to ATL. Medial pFG did not stream to IPL, but did project more superiorly within the parietal lobe. Recent neuroanatomical studies from MR tractography and tracing studies in non-human primates have suggested that the inferior longitudinal fasciculus (ILF), which connects ventral aspects of ATL, occipito-temporal, and occipital cortex, also branches dorsally in its posterior extent to terminate in dorsoparietal regions<sup>89,90</sup>—potentially connecting ATL to SPL indirectly via the medial pFG. To test this possibility, we assessed the posterior trajectory of a seed more anteriorly along the ILF. The streamline passed through the medial pFG neighborhood and branched superiorly into SPL (Fig.2C). Likewise, SPL streamlines descended to intersect the ILF streamline. A waypoint seed placed at this junction streamed to SPL, the anterior ILF seed and medial pFG. Thus the tractography reveals two pathways from temporal to parietal regions of the network: one that connects ATL to IPL via the pMTG (Fig.2D), and a second connecting ATL to the SPL via

the medial pFG (Fig.2E). This provides an in-vivo demonstration of the dorsal-projecting ILF branch in humans.

## An anatomically-constrained computational model

Figure 3A shows a schematic of the ALE and connectivity results. We next constructed a neurocomputational model whose architecture mirrors these results, shown in Figure 3B. The model is a deep recurrent neural network that computes mappings amongst visual representations of objects (coded in LOC), verbal descriptors (STG), and functional (IPL) and praxic (SPL) action representations<sup>85</sup>. The model was trained with predictive error-driven learning to generate an item's full complement of visual, verbal, function and praxic properties, given a subset of these as input. Surface representations were generated to capture three well-documented aspects of environmental structure: (a) hierarchical similarity with few properties shared across domains, more shared within domains, and many shared within basic categories<sup>11</sup>; (b) many more praxic and functional features for artifacts and somewhat more visual features for animals<sup>10,11</sup>; and (c) more feature overlap amongst animals than artifacts<sup>5</sup> (see SI-Methods 5 and Table S3)

We used the model to assess whether connectivity and learning jointly explain the category-specific patterns observed in the ALE meta-analysis. Fifteen models with different initial random weights were trained, providing analogs of fifteen subjects in a brain imaging study. Models were tested with simulations of both word and picture comprehension. The activation patterns generated by these inputs were treated as analogs of the BOLD response and analyzed to identify model regions showing systematic category effects<sup>12</sup> (see Methods).

Results are shown in Figure 3C. All category effects observed in the ALE analysis emerged in the corresponding model layers for both words and pictures. Medial pFG, pMTG, IPL and SPL responded more to artifacts because they strongly interact with function or praxis representations. Lateral pFG responded more to animals because the medial units had partially "specialized" to represent artifacts. Thus model connectivity, learning and environmental structure together produced the category-sensitive activations observed in the ALE analysis.

Category-specific activations have also been observed during word comprehension in congenitally blind participants, providing important support for domain-specific views since such results cannot arise from domain differences in visual structure<sup>9,63,91</sup>. To assess whether learning and connectivity also explain such patterns, we replicated the simulations in models trained without visual inputs or targets. The animal advantage in lateral pFG disappeared, presumably because these units no longer communicate activation from early vision<sup>12</sup>. Artifacts, however, continued to elicit greater activation in medial pFG, posterior pMTG, IPL, and SPL, because these units continue to participate in generating function and praxis representations for object-directed action. The absence of a category effect in lateral pFG and tool/praxis-specific activation patterns in pMTG, IPL, and SPL have all been reported in this population<sup>9,63,91</sup>.

## Disorders of semantic representation

We next considered whether learning and connectivity explain the primary disorders of semantic representation and their anatomical basis. By primary we mean acquired disorders that (a) reflect degraded semantic representation rather than access/retrieval deficits<sup>92,93</sup> and (b) have been shown in case-series studies to manifest predictable patterns of impairment. These include: (a) semantic dementia (SD), where progressive bilateral ATL atrophy produces a category-general semantic impairment<sup>79</sup>; (b) herpes simplex viral encephalitis (HSVE), where acute bilateral ATL pathology produces chronic impairments disproportionately affecting animals<sup>94</sup>; (c) temporo-parietal tumor resection (TPT), which produces greater impairment for artifacts<sup>95</sup>; and (d) forms of visual agnosia (VA) producing slower and less accurate recognition for animals<sup>81,96</sup>.

Both SD and HSVE were simulated by removing increasing proportions of ATL connections. To capture the progressive nature of SD, performance was assessed without relearning after connections were removed. For HSVE we considered two damage models. In the homogeneous variant (HSVE), damage was identical to SD but the network was then retrained to simulate acute injury with recovery. In the asymmetric variant (HSVE+), lateral connections between ATL and pFG were more likely to be removed than medial connections, consistent with a possible difference noted in a direct comparison of white-matter pathology in SD vs. HSVE<sup>94</sup> (see SI-Discussion 3.2). The damaged model was again retrained. TPT was simulated by removing a proportion of connections within/between PMTG and IPL layers, while VA was simulated by removing weights between LOC and pFG layers. At each level of damage for each disorder, we simulated picture naming for all animal and artifact items<sup>85</sup>.

Results are shown in Figure 4. The model captures the direction and magnitude of several key phenomena including: (a) no category effect in SD, (b) a substantial animal disadvantage in both HSVE variants (results of HSVE+ in Fig. S5), (c) a modest artifact disadvantage in TPT, (d) an animal disadvantage in response time in VA, (e) worse anomia in SD than HSVE, and (f) a smaller and opposite category effect in TPT compared to HSVE.

The pattern of network connectivity transparently explains the key results for two patient groups: TPT, where disrupted interactions with function representations in IPL disproportionately affect artifacts, and SD, where ATL damage produces a domain-general impairment. In VA the category effect arises from “visual crowding”<sup>81</sup>: because animals overlap more in their visual properties<sup>5,11</sup>, they are more difficult to discriminate (and hence to name) when inputs from vision are impoverished<sup>97</sup>. In HSVE, the model pathology is identical to SD—the category effect thus arises through re-learning, via two mechanisms. First, intact functional and praxic layers can support new learning for items with these properties, that is, for artifacts. Second, because animals share more properties than artifacts their ATL representations are also “conceptually crowded,” compromising relearning of inter-item differences when ATL representations are damaged<sup>79</sup>. In HSVE+, the effect is magnified when lateral ATL-pFG connections are disproportionately removed, since these connections provide more support for animal knowledge as shown in the simulation of imaging results (see Figs. S4 and S5).

The simulation suggests a novel resolution to the long-standing puzzle of why patients with HSVE and SD show qualitatively distinct impairments despite largely overlapping pathology<sup>79</sup>. Category-specific deficits may arise when white-matter pathology is distributed asymmetrically in the ATL, but even when pathology is identical they may emerge through relearning following the acute injury (see SI-Discussion 3.2 and 4). To assess this hypothesis, we first evaluated model predictions by regressing the magnitude of the category effect (artifact accuracy – animal accuracy) on the total amount of damage, the amount of relearning and their interaction, in the simulation of both HSVE and HSVE+ (see details in SI-Methods 6). In both cases the two factors interacted reliably: when damage was severe, relearning produced a larger category effect, but when damage was mild, relearning shrunk the category effect (Fig. S7A&B; interaction for HSVE  $t = 2.501$ ,  $p = .014$ ; interaction term for HSVE+  $t = 2.137$ ,  $p = .035$ ). We then assessed whether the same pattern is observed in the literature. Across 19 previously-reported HSVE cases of category-specific impairment (Table S6), we regressed the reported category effect on the overall severity of the impairment, the amount of relearning (assessed as the time elapsed between injury and test), and their interaction term. Consistent with model predictions, these factors interacted reliably [ $t = 3.298$ ,  $p < .01$ ]: relearning produced larger effects when deficits were severe but smaller effects when deficits were mild (Fig. S7C). The same pattern was also observed longitudinally in 4 patients with HSVE<sup>98,99</sup> (Fig. S7D). By contrast, this pattern was not found in the non-HSVE cases (for full results, see Table S7). Thus the model's account of category-specific impairment is consistent with the existing literature.

Finally, we considered classic lesion-symptom mapping results suggesting that animal-selective deficits occur with ventro-temporal damage while artifact-selective deficits occur with temporo-parietal pathology<sup>100</sup>. We conducted a model lesion-symptom analysis by grouping simulated patients across all four disorders into a single dataset. We quantified regional pathology in every model patient as the proportion of connections removed from each layer and measured category selectivity as the difference in accuracy naming artifacts vs animals. We then computed, across all patients at each layer, the correlation between pathology and category selectivity.

Figure 4B shows the results. Damage in ventral temporal model regions (ATL, pFG and LOC) significantly predicted greater impairment for animals than artifacts, while damage in PMTG and IPL regions predicted the reverse pattern (Fig.4B-left). Importantly, the ATL effect was only carried by the HSVE simulations: SD simulations alone showed no relationship between lesion severity and category effect (Fig.4B-middle). The same pattern is observed in case-series studies of the corresponding syndromes for which data is available (Fig.4B-right). Thus the model explains both the canonical lesion-symptom results and their puzzling discrepancy with SD.

## Discussion

We have proposed a new neurocomputational model for the neural bases of semantic representation which, in building on contributions from several groups<sup>19,95,101</sup>, unifies domain-specific and domain-general approaches. The core and critical theoretical contribution is that initial connectivity, domain-general learning, and environmental

structure all jointly shape functional activation within the cortical semantic network, leading to graded category-specificity in some network components but domain-general processing in the ATL hub, within a network whose principal function is to support cross-modal inference. The model explains the neuroimaging and patient phenomena central to both domain-specific and domain-general theories, including (a) category-specific patterns of functional activation in sighted and congenitally-blind individuals, (b) patterns of impairment observed across four different neuropsychological syndromes, and (c) the anatomical bases of these patterns. It also exemplifies a general approach to functional specialization in cortex that we have termed connectivity-constrained cognition or  $C^3$ .

Our model reconciles and extends several competing perspectives in the literature. Like the sensory-functional hypothesis, category sensitivity arises from domain differences in the recruitment of action versus visual representations<sup>10,102</sup>; but we show that learning and connectivity can produce domain differences even in the absence of visual experience, and outside canonical action areas, addressing key criticisms of the sensory-functional view<sup>3</sup>. Like the distributed domain-specific hypothesis, category-sensitivity reflects network connectivity, with temporo-parietal pathways initially configured to facilitate vision-action relationships important for tool knowledge<sup>9</sup>. The model reconciles this perspective with the extensive evidence for domain-general representation in the ATL. An important account of optic aphasia relied on graded functional specialization arising from constraints on local connectivity<sup>8</sup>; our model extends this idea to incorporate long-range connectivity constraints. Like the correlated-structure view, category-selectivity arises partly from different patterns of overlap among animal versus artifact properties<sup>6</sup>, but in our model network connectivity also plays a critical role. Finally, this work extends the hub-and-spoke model under which the ATL constitutes a domain-general semantic hub for computing mappings amongst all surface modalities<sup>4</sup>. The model illustrates how domain-specific patterns can arise within the “spokes” of such a network, even while the ATL plays a critical domain-general role in semantic representation<sup>13</sup> (see SI-Discussion 2 for relationship to other models).

In emphasizing semantic representation we have not considered the fronto-parietal systems involved in semantic control<sup>103</sup>, nor does the model address open questions about lateralization, abstract and social concepts, or other conceptual distinctions amongst concrete objects. We therefore view the proposed model as establishing a crucial foundation rather than an end point. Nevertheless, the current work is unique in developing a neurocognitive model whose architecture is fully constrained, *a priori*, by systems-level neural data. The project illustrates how simulation models at this level of abstraction can provide an important conceptual bridge for relating structural and functional brain imaging and healthy and disordered cognitive functioning<sup>74</sup>. While interest in neural networks has recently rekindled in machine learning<sup>104</sup>, their original promise as tools for bridging minds and brains<sup>105</sup> has remained largely untested. We have shown that the convergent use of network simulation models with the other tools of cognitive neuroscience can produce new insights with the potential to resolve otherwise pernicious theoretical disputes. We further believe the  $C^3$  approach we have sketched, in which network models are used to illuminate how connectivity, learning, and environmental structure give joint rise to cognitive function, can be similarly useful in other cognitive domains.



## Methods

### ALE analysis

We followed the standard literature search procedure from previous ALE studies<sup>25,106</sup> and found 49 papers describing 73 independent studies (31 for animal and 42 for artifact; for study selection, see SI-Methods 1) up to July, 2013 and reporting a total of 270 foci (103 for animal and 167 for artifact). The ALE meta-analysis was carried out with the software package gingerALE v2.3<sup>107,108</sup>. The ALE analysis strictly followed the steps proposed by Price et al.<sup>106</sup> and Eickhoff et al.<sup>72,107,108</sup>, and coordinates in MNI space were used for ALE analysis and reports. Main effects of animal and artefact concepts (concordance of foci showing greater activations for animal vs. baseline and artefact vs. baseline) are reported in Table S1. Next we combined the resulting ALE animal and artifact maps and tested for brain regions commonly activated by both categories (conjunction analysis) and showing reliably different activations for the two categories of interest (contrast analysis) as reported in the main text.

### Connectivity analysis

Diffusion-weighted images were collected from 24 right-handed healthy subjects (11 female; mean age = 25.9) at University of Manchester, UK<sup>88</sup>. All participants are right-handed as determined by the Edinburgh Handedness Inventory<sup>109</sup>. Inclusion and exclusion criteria were stated in previous studies<sup>88,110</sup>, and no randomization or blinding was needed. Informed consents were obtained for all subjects.

**Image acquisition**—Imaging data were acquired on a 3T Philips Achieva scanner (Philips Medical Systems, Best, Netherlands), using an 8 element SENSE head coil. Diffusion weighted imaging was performed using a pulsed gradient spin echo echo-planar sequence with TE=59 ms, TR≈1500 ms (cardiac gated), G=62 mTm<sup>-1</sup>, half scan factor=0.679, 112×112 image matrix reconstructed to 128×128 using zero padding, reconstructed resolution 1.875×1.875 mm, slice thickness 2.1 mm, 60 contiguous slices, 61 non-collinear diffusion sensitization directions at b=1200 smm<sup>-2</sup> (δ=29.8ms, δ=13.1ms), 1 at b=0, and SENSE acceleration factor=2.5. A high-resolution T1-weighted 3D turbo field echo inversion recovery scan (TR≈2000 ms, TE=3.9 ms, TI=1150ms, flip angle 8°, 256×205 matrix reconstructed to 256×256, reconstructed resolution 0.938×0.938 mm, slice thickness 0.9 mm, 160 slices, SENSE factor=2.5), was also acquired for the purpose of high precision anatomical localization of seed regions for tracking. Distortion correction to remediate signal loss in ventral ATL was applied using the same method reported in other studies<sup>88,110</sup>.

**ROI definition**—The ROIs were chosen to reside in the white-matter underlying the peaks identified in the ALE-meta analysis, or from regions reported in the relevant literature. Specifically, ROIs in lateral pFG, medial pFG, MOG, pMTG, IPL (IPL\_1) and SPL (SPL\_1) in the left hemisphere were chosen from the ALE meta-analysis as regions showing reliable category-specific activation patterns. The ATL ROI was chosen from an fMRI study<sup>111</sup> that reported cross-modal activation for conceptual processing in the ATL. Due to the uncertainty of tempo-parietal connectivity, we also included a second IPL seed (IPL\_2) whose

coordinates were chosen from a study in which TMS to this region slowed naming of tools but not animals<sup>112</sup>. Likewise we included a second SPL ROI (SPL\_2) reported by Mahon et al.<sup>63</sup> as a peak showing preferential activation for artifact stimuli in both sighted and congenitally blind participants. To assess the caudal-going trajectory of the ILF, we placed an additional seed in the inferior temporal white matter at the anterior-most extent of the artifact peak revealed by the ALE meta-analysis. As reported in the main text, this streamline branched superiorly up into parietal cortex, intersecting the streamline from the SPL seeds. To determine whether a single tract might connect SPL, medial pFG and ATL, we placed a final waypoint seed at this intersection. For more details about ROI definitions, see SI-Methods 2.

**Probabilistic tracking procedure**—We restricted our analysis to the left hemisphere, and following similar procedure of previous study<sup>110</sup>, a sphere with a diameter of 6mm centered on the seed coordinate for each ROI was then drawn in the MNI template (see Table S2 for the exact coordinates; details in SI). Finally, the ROIs defined in a common space were converted into the native brain space of each individual.

For each voxel within a seed ROI sphere, 15,000 streamlines were initiated for unconstrained probabilistic tractography using the PICo (Probabilistic Index of Connectivity) method<sup>110,113</sup>. Step size was set to 0.50 mm. Stopping criteria for the streamlines were set so that tracking terminated if pathway curvature over a voxel was greater than 180, or the streamline reached a physical path limit of 500 mm. In the native-space tracking data from each seed region for each individual, ROI masks were overlaid and a maximum connectivity value (ranging from 0 to 15,000) was obtained for the seed region and each of the other ROIs, resulting in a matrix of streamline-based connectivity. A standard two-level threshold approach was applied to determine high likelihood of connection in this matrix<sup>110</sup>. At each individual level, three thresholds, 1% (lenient), 2.5% (standard), and 5% (stringent) were used to investigate the probable tracts in a wider range. At the group level, only connections present in at least half ( $\geq 12$ ) subjects were considered highly probable across subjects (for more details of thresholding, see in SI-Methods 3). A group-averaged tractography image was then obtained by averaging the normalized individual data<sup>110</sup>.

### Computer simulations of fMRI data

The model architecture shown in Figure 3B (main text) was implemented using the Light Efficient Network Simulator (LENS) software<sup>114</sup>. The model included four visible layers directly encoding model analogs of visual, verbal (names and descriptions), praxic, and functional properties of objects. Each visible layer was reciprocally connected with its own modality-specific hidden layer, providing model analogs to the posterior fusiform (pFG, visual hidden units), superior temporal gyrus (STG, verbal hidden units), inferior parietal lobule (IPL, function hidden units), and superior parietal lobule (SPL, praxic hidden units). The model also included two further hidden layers corresponding to the ventral ATL and the posterior MTG. Hidden layers were connected with bidirectional connections matching the results of the tractography analysis, as shown in Figure 3B. A spatial gradient of learning rate on visuo-praxic connections of units in the pFG layer along an anatomical lateral-to-

medial axis was implemented<sup>12</sup> (details see SI-Methods 4), to capture the observation that medial pFG is more strongly connected to parietal regions than is lateral pFG<sup>19</sup>. All units employed a sigmoidal activation function and were given a fixed bias of -2 so that, in the absence of input from other units, they tended to adopt a low activation state. Units updated their activation states continuously using a time integration constant of 30. Model implementation and training environment files can be downloaded online (see Data Availability).

**Training environment**—A model environment was constructed to contain visual, verbal, function/action and praxic representations for 24 different exemplars of animals and 24 different exemplars of tools, with each domain organized into 4 basic categories, each containing 6 exemplars (for representational schemes of training exemplars, see Table S3 and SI-Methods 5). In total, there were 48 training exemplars. Visual and verbal representations for each item in this set were generated stochastically in accordance with the constraints identified by Rogers et al.<sup>5</sup> in their analysis of verbal attribute-listing norms and line drawings of objects. Thus (a) items in different domains shared few properties; (b) items within the same category shared many properties; (c) animals from different categories shared more properties than did artifacts from different categories; and (d) animals had more properties overall than did artifacts. Each item was also given a unique name as well as a label common to all items in the same category.

Praxis representations were also constructed for each item, taking the form of distributed patterns over the 10 units in the visible praxic layer<sup>12</sup>. For all animal items, these units were turned off. For artifacts, distributed patterns were created that covaried with, but were not identical to, the item's corresponding visual pattern, as a model analog of vision-to-action affordances. Function representations simply duplicated the praxic patterns across the 10 visible units for function features.

**Model training procedures**—The model was trained to generate, given partial information about an item as input, all of the item's associated properties, including its name, verbal description, visual, function and praxic features, similar to our previous work<sup>12</sup> (details see SI-Methods 5). Weights were updated using a variant of the backpropagation learning algorithm suited to recurrent neural networks, using a base learning rate of 0.01 and a weight decay of 0.0005 without momentum<sup>15</sup>. 'Congenitally blind' model variants were trained with the same parameters on the same patterns, but without visual experience: visual inputs were never applied to the model, and visual units were never given targets. All models were trained exhaustively for 100k epochs at which point they generated correct output (details, see SI-Methods 5) across all visible units for the great majority (>94%) of inputs. For each model population (sighted/blind), 15 different subjects were simulated with different model training runs, each initialized with a different set of weights sampled from a uniform random distribution with mean 0 and range  $\pm 0.1$  (for model performance after training, see Table S4).

**Simulating functional brain imaging studies**—The brain imaging studies simulated involved two tasks: picture viewing, in which participants made a semantic judgment from a picture of a familiar item, and name comprehension, in which they made a semantic

judgment from the spoken name of a familiar item. To simulate the picture viewing task in sighted model variants, the visual feature pattern corresponding to a familiar item was applied to visual input units and the trained model cycled until it reached a steady state. To simulate name comprehension in both sighted and congenitally blind variants, a single unit corresponding to the item's name was given excitatory external input, and the model again cycled until it reached a steady state. In both tasks, after settling, the activation of each model unit was recorded and taken as an analog of the mean activity difference from baseline for a population of neurons at a single voxel. This value was then distorted with Gaussian noise ( $\mu = 0$ ,  $\sigma^2 = 0.1$ ) to reflect the error in signal estimation intrinsic to brain imaging methods. The response of each unit was then averaged across items in each condition (Animal vs. Artifact) and then spatially smoothed with a Gaussian kernel ( $\mu = 0$ ,  $\sigma^2 = 1$ ) encompassing two adjacent units. A group-level contrast was performed to find the peak activation for both Animal and Artifacts concepts using the averaged data across the 15 model subjects. An ROI analysis was then performed on activation value of the peak unit averaged together with two neighboring units on either side.

### Computer simulations of patient data

Following simulation of functional imaging data, we assessed whether the model could explain patterns of impaired semantic cognition and their neuroanatomical basis in four disorders of semantic representation. Here, we provide basic information of the phenotype of each disorder and model simulation procedure (details of pathology and motivation in SI-Discussion 3). The model architecture and training environment were the same as in the simulations of brain imaging data, except that pattern frequencies were adapted to ensure that the names of animal and artifact items appeared as inputs and targets with equal frequency (see Fig. S3).

(1) Semantic dementia (SD) is a neurodegenerative disorder associated with gradual thinning of cortical grey matter and associated white-matter fibers, centered in the ATL78, and produces a robust, progressive and yet selective deterioration of semantic knowledge for all kinds of concepts, across all modalities of reception and expression76,5,116. We simulated SD by removing an increasing proportion of all weights entering, leaving, or internal to the ATL hidden layer uniformly from 0.1 to 1.0 with an increment of 0.1. At each level of damage, the model was tested without allowing it to relearn/reorganize.

(2) Herpes Simplex Viral Encephalitis (HSVE) is a disease that produces rapid bilateral necrosis of gray and white matter, generally encompassing the same regions affected in SD, but patients with semantic impairments from HSVE, has been found with greater damage in temporal white matter especially in the lateral axis94. HSVE patients often show less semantic impairment overall, with greater deficits of knowledge for animals than for manmade objects79,94. The main paper considers two potential explanations, each associated with a different model of HSVE pathology. The model first captures differences in the time-course of SD vs. HSVE: whereas the former progresses slowly over a course of years, the latter develops rapidly and is then halted by anti-viral medication after which patients often show at least some recovery of function. Weights were removed from the ATL layer as in the SD simulation, but the damaged model was then retrained before assessment

on the naming task (for motivation, see SI-Discussion 3.2). Retraining employed the same parameters used in the last cycle of the initial training, namely, learning rate =  $10^{-3}$  and weight decay =  $10^{-6}$ . The main text reports data following 3k epochs of retraining when the model performance had largely stabilized (see Fig. S4 for recovery trajectory). The second further assessed the potential contribution of differential white-matter damage across the lateral/medial axis of the ATL in HSVE94. To simulate this, a proportion of ATL connections selecting uniformly with probability  $p$  were removed in a first pass (as in the SD simulation), then a second removal of connections was applied to weights between ATL and lateral pFG units (units 0-9). In a 30% lesion, for instance, 30% of all ATL connections entering or leaving each ATL unit were removed, and then 30% of the original connections between ATL and lateral pFG units were additionally removed. Thus, when the global lesion severity equaled or exceeded 50%, all connections between ATL and lateral pFG were removed. Finally, the model was retrained as in the homogeneous variant of HSVE and performance on the retrained model was assessed (see Fig. S5). We also demonstrated that without relearning, the HSVE variant showed little evidence for category-specific impairment (just as with SD simulations), but the HSVE+ variant showed more severe impairment in the animal category (see Fig. S6).

(3) Temporo-parietal tumor resection (TPT). Campanella and colleagues<sup>95</sup> presented the first relatively large-scale case-series study of artifact-category impairment in a group of 30 patients who had undergone surgical removal of temporal-lobe tumors. The group exhibited significantly worse knowledge of nonliving things compared to animals, with difference scores in naming accuracy ranging from 2%-21%. Voxel-based lesion-symptom mapping (VLSM) revealed that the magnitude of the category effect was predicted by pathology in posterior MTG, inferior parietal cortex, and the underlying white matter. To simulate this pathology in the model we removed connections between and within IPL and MTG model regions uniformly from 0.1 to 1.0 with an increment of 0.1.

(4) Category-specific visual agnosia (VA). Finally, a long tradition of research suggests that forms of associative visual agnosia arising from damage to occipitotemporal regions can have a greater impact on recognition of living than nonliving things<sup>117,118</sup>. The deficit is specific to vision, and more evident in naming latency rather than accuracy at milder impairment<sup>96</sup>. To capture disordered visual perception, we removed a proportion of the weights projecting from the visual input layer (LOC) to the visual hidden layer (pFG). We used a smaller range from 0.025 to 0.25 in increments of 0.025 in order to preserve sufficient visual inputs to the system.

**Assessment of model performance**—For each disorder, model performance was assessed on simulated picture naming. For each item, the corresponding visual features were given positive input, and the activations subsequently generated over units encoding basic-level names were inspected to assess performance. Naming performance was scored as correct if the target name unit was (a) the most active of all basic name units and (b) was activated above 0.5; otherwise it was scored as incorrect. In visual agnosia at mild impairment, the category-specific impairment can be observed in response time so that we computed naming latency as the number of update cycles (ticks) required for the target unit to reach an activation of 0.5 (for correct naming trials only). For comparison to standardized

human latency data in VA, the model latency was standardized by computing  $(N_j - N_0)/N_0$ , where  $N_j$  is the number of ticks used for the model to produce a response at the  $j$ th level of lesion severity and  $N_0$  was the number of ticks for naming without any lesion in the model. Therefore, the raw latency measure is adjusted by baseline response latency differences between categories that exist in the performance of the intact models. Note that in figures of the main text, the severity was recomputed as the overall naming accuracy collapsing animal and artifact categories. See Table S5 for naming accuracy and latency at different levels of lesion severity measured as percentage of affected connections.

### Data availability

Program scripts and source data that support the data analysis of this project are available in online public repositories, and more details are available upon request. See [https://github.com/halleycl/ChenETAL\\_NatHumanBehav\\_SI-Online-materials](https://github.com/halleycl/ChenETAL_NatHumanBehav_SI-Online-materials) and <https://app.box.com/v/ChenETAL-NatHumanBehav-SI>.

### Supplementary Information

Refer to Web version on PubMed Central for supplementary material.

### Acknowledgments

This research was supported by a programme grant from the Medical Research Council (MRC, UK, MR/J004146/1) to MALR and by a University Fellowship from UW-Madison to LC. We also want to thank Lauren Cloutman for assisting the tractography analysis and Ryo Ishibashi for assisting the ALE analysis.

### References

1. Fernandino L, et al. Predicting brain activation patterns associated with individual lexical concepts based on five sensory-motor attributes. *Neuropsychologia*. 2015; 76:17–26. [PubMed: 25863238]
2. Caramazza A, Shelton JR. Domain-specific knowledge systems in the brain: The animate- inanimate distinction. *J Cogn Neurosci*. 1998; 10:1–34. [PubMed: 9526080]
3. Caramazza A, Mahon BZ. The organization of conceptual knowledge: The evidence from category-specific semantic deficits. *Trends Cogn Sci*. 2003; 7:354–361. [PubMed: 12907231]
4. Patterson K, Nestor PJ, Rogers TT. Where do you know what you know? The representation of semantic knowledge in the human brain. *Nat Rev Neurosci*. 2007; 8:976–987. [PubMed: 18026167]
5. Rogers TT, et al. The structure and deterioration of semantic memory: a computational and neuropsychological investigation. *Psychol Rev*. 2004; 111:205–235. [PubMed: 14756594]
6. Tyler LK, Moss HE, Durrant-Peatfield MR, Levy JP. Conceptual Structure and the Structure of Concepts: A Distributed Account of Category-Specific Deficits. *Brain Lang*. 2000; 75:195–231. [PubMed: 11049666]
7. Chen L, Rogers TT. Revisiting domain-general accounts of category specificity in mind and brain. *Wiley Interdiscip Rev Cogn Sci*. 2014; 5:327–44. [PubMed: 26308567]
8. Plaut DC. Graded modality-specific specialisation in semantics: A computational account of optic aphasia. *Cogn Neuropsychol*. 2002; 19:603–639. [PubMed: 20957556]
9. Mahon BZ, Anzellotti S, Schwarzbach J, Zampini M, Caramazza A. Category-Specific Organization in the Human Brain Does Not Require Visual Experience. *Neuron*. 2009; 63:397–405. [PubMed: 19679078]
10. Warrington EK, Shallice T. Category specific semantic impairments. *Brain*. 1984; 107:829–854. [PubMed: 6206910]

11. Cree GS, McRae K. Analyzing the factors underlying the structure and computation of the meaning of chipmunk, cherry, chisel, cheese, and cello (and many other such concrete nouns). *J Exp Psychol Gen.* 2003; 132:163–201. [PubMed: 12825636]
12. Chen L, Rogers TT. A Model of Emergent Category-specific Activation in the Posterior Fusiform Gyrus of Sighted and Congenitally Blind Populations. *J Cogn Neurosci.* 2015; 27:1981–1999. [PubMed: 26042499]
13. Pobric G, Jefferies E, Lambon Ralph MA. Category-specific versus category-general semantic impairment induced by transcranial magnetic stimulation. *Curr Biol.* 2010; 20:964–8. [PubMed: 20451381]
14. Sadtler PT, et al. Neural constraints on learning. *Nature.* 2014; 512:423–426. [PubMed: 25164754]
15. Gomez J, et al. Functionally defined white matter reveals segregated pathways in human ventral temporal cortex associated with category-specific processing. *Neuron.* 2015; 85:216–227. [PubMed: 25569351]
16. Mahon BZ, Caramazza A. What drives the organization of object knowledge in the brain? *Trends Cogn Sci.* 2011; 15:97–103. [PubMed: 21317022]
17. Plaut DC, Behrmann M. Complementary neural representations for faces and words: A computational exploration. *Cogn Neuropsychol.* 2011; 28:251–275. [PubMed: 22185237]
18. Martin A, Chao LL. Semantic memory and the brain: structure and processes. *Curr Opin Neurobiol.* 2001; 11:194–201. [PubMed: 11301239]
19. Mahon BZ, et al. Action-related properties shape object representations in the ventral stream. *Neuron.* 2007; 55:507–520. [PubMed: 17678861]
20. Hickok G, Poeppel D. The cortical organization of speech processing. *Nat Rev Neurosci.* 2007; 8:393–402. [PubMed: 17431404]
21. Kellenbach ML, Brett M, Patterson K. Actions speak louder than functions: the importance of manipulability and action in tool representation. *J Cogn Neurosci.* 2003; 15:30–46. [PubMed: 12590841]
22. Chouinard, Pa, Goodale, Ma. Category-specific neural processing for naming pictures of animals and naming pictures of tools: an ALE meta-analysis. *Neuropsychologia.* 2010; 48:409–418. [PubMed: 19800353]
23. Pobric G, Jefferies E, Lambon Ralph MA. Anterior temporal lobes mediate semantic representation: Mimicking semantic dementia by using rTMS in normal participants. *Proc Natl Acad Sci U S A.* 2007; 104:20137–20141. [PubMed: 18056637]
24. Acosta-Cabronero J, et al. Atrophy, hypometabolism and white matter abnormalities in semantic dementia tell a coherent story. *Brain.* 2011; 134:2025–2035. [PubMed: 21646331]
25. Chouinard PA, Goodale MA. Category-specific neural processing for naming pictures of animals and naming pictures of tools: An ALE meta-analysis. *Neuropsychologia.* 2010; 48:409. [PubMed: 19800353]
26. Hwang K, et al. Category-specific activations during word generation reflect experiential sensorimotor modalities. *Neuroimage.* 2009; 48:717–725. [PubMed: 19559802]
27. Smith CD, et al. Differences in functional magnetic resonance imaging activation by category in a visual confrontation naming task. *J Neuroimaging.* 2001; 11:165–170. [PubMed: 11296587]
28. Grossman M, et al. The neural basis for category-specific knowledge: an fMRI study. *Neuroimage.* 2002; 15:936–948. [PubMed: 11906234]
29. Martin A, Haxby JV, Lalonde FM, Wiggs CL, Ungerleider LG. Discrete Cortical Regions Associated with Knowledge of Color and Knowledge of Action. *Science (80- ).* 1995; 270:102–105.
30. Tyler LK, et al. Do semantic categories activate distinct cortical regions? Evidence for a distributed neural semantic system. *Cogn Neuropsychol.* 2003; 20:541–559. [PubMed: 20957584]
31. Damasio H, Grabowski TJ, Tranel D, Hichwa RD, Damasio AR. A neural basis for lexical retrieval. *Nature.* 1996; 380:499–505. [PubMed: 8606767]
32. Cappa SF, Perani D, Schnur T, Tettamanti M, Fazio F. The effects of semantic category and knowledge type on lexical-semantic access: a PET study. *Neuroimage.* 1998; 8:350–359. [PubMed: 9811553]

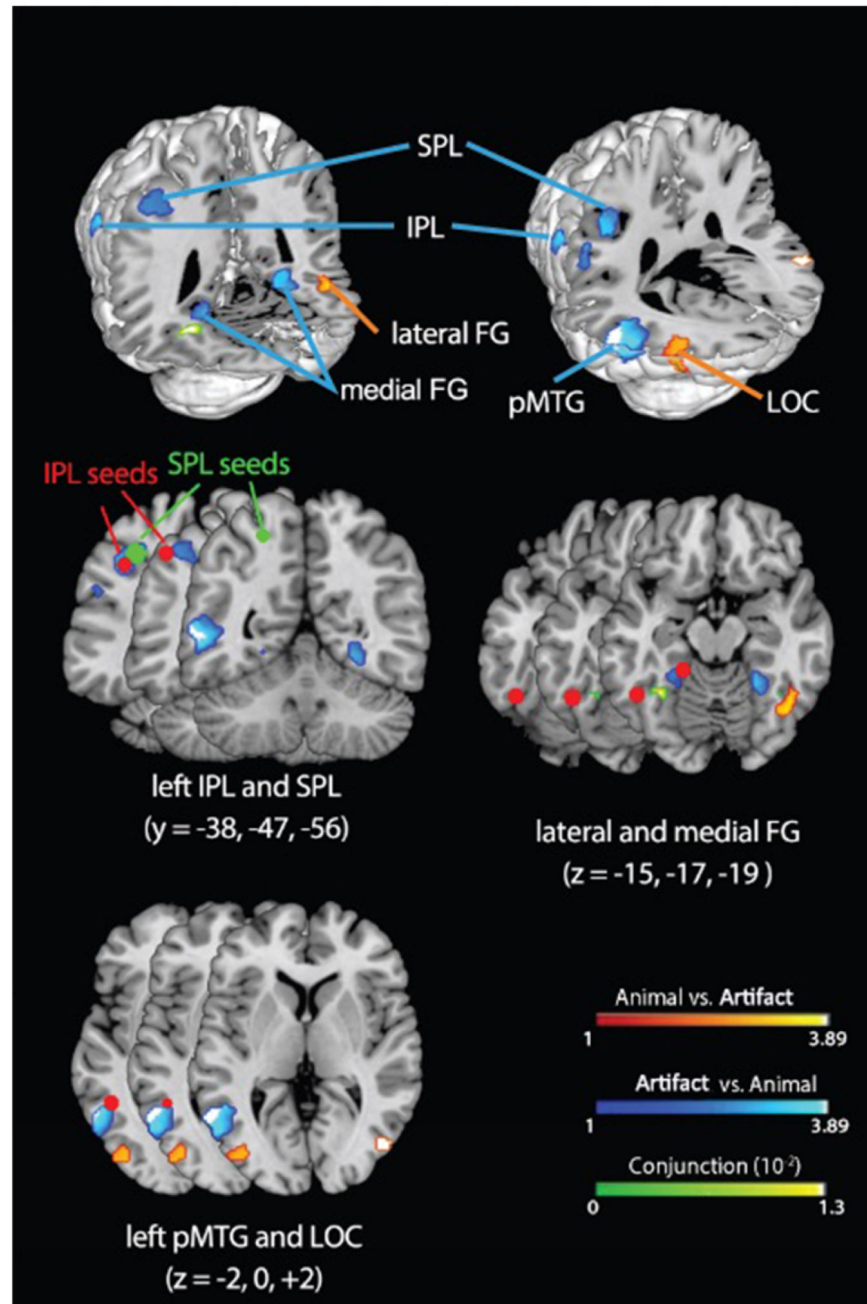
33. Mechelli A, Sartori G, Orlandi P, Price CJ. Semantic relevance explains category effects in medial fusiform gyri. *Neuroimage*. 2006; 30:992–1002. [PubMed: 16343950]
34. Noppeney U, Josephs O, Kiebel S, Friston KJ, Price CJ. Action selectivity in parietal and temporal cortex. *Brain Res Cogn Brain Res*. 2005; 25:641. [PubMed: 16242924]
35. Phillips JA, Noppeney U, Humphreys GW, Price CJ. Can segregation within the semantic system account for category-specific deficits? *Brain*. 2002; 125:2067–2080. [PubMed: 12183352]
36. Kroliczak G, Frey SH. A common network in the left cerebral hemisphere represents planning of tool use pantomimes and familiar intransitive gestures at the hand-independent level. *Cereb Cortex*. 2009; 19:2396–2410. [PubMed: 19181695]
37. Grossman M, et al. Category-specific semantic memory: Converging evidence from bold fMRI and Alzheimer's disease. *Neuroimage*. 2013; 68:263–274. [PubMed: 23220494]
38. Laine M, Rinne JO, Hiltunen J, Kaasinen V, Sipilä H. Different brain activation patterns during production of animals versus artefacts: A PET activation study on category-specific processing. *Cogn Brain Res*. 2002; 13:95–99.
39. Boronat CB, et al. Distinctions between manipulation and function knowledge of objects: evidence from functional magnetic resonance imaging. *Cogn Brain Res*. 2005; 23:361–373.
40. Gorno-Tempini M-L. Category differences in brain activation studies: where do they come from? *Proc R Soc London Ser B Biol Sci*. 2000; 267:1253–1258.
41. Gerlach C, Law I, Paulson OB. When action turns into words. Activation of motor-based knowledge during categorization of manipulable objects. *J Cogn Neurosci*. 2002; 14:1230–1239. [PubMed: 12495528]
42. Rogers TT, Hocking J, Mechelli A, Patterson K, Price C. Fusiform activation to animals is driven by the process, not the stimulus. *J Cogn Neurosci*. 2005; 17:434–45. [PubMed: 15814003]
43. Lewis JW, Brefczynski JA, Phinney RE, Janik JJ, DeYoe EA. Distinct cortical pathways for processing tool versus animal sounds. *J Neurosci*. 2005; 25:5148–5158. [PubMed: 15917455]
44. Canessa N, et al. The different neural correlates of action and functional knowledge in semantic memory: An fMRI study. *Cereb Cortex*. 2008; 18:740–751. [PubMed: 17621607]
45. Joseph JE, Gathers AD, Piper GA. Shared and dissociated cortical regions for object and letter processing. *Cogn Brain Res*. 2003; 17:56–67.
46. Gerlach C, et al. Brain activity related to integrative processes in visual object recognition: bottom-up integration and the modulatory influence of stored knowledge. *Neuropsychologia*. 2002; 40:1254–1267. [PubMed: 11931928]
47. Gerlach C, Law I, Gade A, Paulson OB. Categorization and category effects in normal object recognition: A PET study. *Neuropsychologia*. 2000; 38:1693–1703. [PubMed: 11099727]
48. Devlin JT, Rushworth MFS, Matthews PM. Category-related activation for written words in the posterior fusiform is task specific. *Neuropsychologia*. 2005; 43:69–74. [PubMed: 15488907]
49. Noppeney U, Price CJ, Penny WD, Friston KJ. Two distinct neural mechanisms for category-selective responses. *Cereb Cortex*. 2006; 16:437–445. [PubMed: 15944370]
50. Whatmough C, Chertkow H, Murtha S, Hanratty K. Dissociable brain regions process object meaning and object structure during picture naming. *Neuropsychologia*. 2002; 40:174–186. [PubMed: 11640940]
51. Grafton ST, Fadiga L, Arbib MA, Rizzolatti G. Premotor cortex activation during observation and naming of familiar tools. *Neuroimage*. 1997; 6:231–236. [PubMed: 9417966]
52. Chao LL, Haxby JV, Martin a. Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. *Nat Neurosci*. 1999; 2:913–919. [PubMed: 10491613]
53. Chao LL, Martin A. Representation of Manipulable Man-Made Objects in the Dorsal Stream. *Neuroimage*. 2000; 12:478–484. [PubMed: 10988041]
54. Goldberg RF, Perfetti CA, Schneider W. Perceptual knowledge retrieval activates sensory brain regions. *J Neurosci*. 2006; 26:4917–4921. [PubMed: 16672666]
55. Bai HM, et al. Functional MRI mapping of category-specific sites associated with naming of famous faces, animals and man-made objects. *Neurosci Bull*. 2011; 27:307–318. [PubMed: 21934726]



56. Folstein JR, Palmeri TJ, Gauthier I. Category learning increases discriminability of relevant object dimensions in visual cortex. *Cereb Cortex*. 2013; 23:814–823. [PubMed: 22490547]
57. Martin, a, Wiggs, CL., Ungerleider, LG., Haxby, JV. Neural correlates of category-specific knowledge. *Nature*. 1996; 379:649–652. [PubMed: 8628399]
58. Perani D, et al. Word and picture matching: a PET study of semantic category effects. *Neuropsychologia*. 1999
59. Handy TC, Grafton ST, Shroff NM, Ketay S, Gazzaniga MS. Graspable objects grab attention when the potential for action is recognized. *Nat Neurosci*. 2003; 6:421–427. [PubMed: 12640459]
60. Gerlach C, Law I, Paulson OB. Structural similarity and category-specificity: A refined account. *Neuropsychologia*. 2004; 42:1543–1553. [PubMed: 15246291]
61. Wadsworth HM, Kana RK. Brain mechanisms of perceiving tools and imagining tool use acts: a functional MRI study. *Neuropsychologia*. 2011; 49:1863–1869. [PubMed: 21419144]
62. Okada T, et al. Naming of animals and tools: a functional magnetic resonance imaging study of categorical differences in the human brain areas commonly used for naming visually presented objects. *Neurosci Lett*. 2000; 296:33. [PubMed: 11099827]
63. Mahon BZ, Schwarzbach J, Caramazza A. The Representation of Tools in Left Parietal Cortex Is Independent of Visual Experience. *Psychol Sci*. 2010; 21:764–771. [PubMed: 20483823]
64. Creem-Regehr SH, Lee JN. Neural representations of graspable objects: are tools special? *Cogn Brain Res*. 2005; 22:457–469.
65. Mruczek REB, von Loga IS, Kastner S. The representation of tool and non-tool object information in the human intraparietal sulcus. *J Neurophysiol*. 2013; 109:2883–96. [PubMed: 23536716]
66. Zannino GD, et al. Visual and semantic processing of living things and artifacts: An fMRI Study. *J Cogn Neurosci*. 2010; 22:554–570. [PubMed: 19301993]
67. Chao LL, Weisberg J, Martin A. Experience-dependent modulation of category-related cortical activity. *Cereb Cortex*. 2002; 12:545–551. [PubMed: 11950772]
68. Anzellotti S, Mahon BZ, Schwarzbach J, Caramazza A. Differential activity for animals and manipulable objects in the anterior temporal lobes. *J Cogn Neurosci*. 2011; 23:2059–2067. [PubMed: 20815733]
69. Moore CJ, Price CJ. A functional neuroimaging study of the variables that generate category-specific object processing differences. *Brain*. 1999; 122:943–962. [PubMed: 10355678]
70. Tranel D, Martin C, Damasio H, Grabowski TJ, Hichwa R. Effects of noun–verb homonymy on the neural correlates of naming concrete entities and actions. *Brain Lang*. 2005; 92:288–299. [PubMed: 15721961]
71. Grabowski TJ, Damasio H, Damasio AR. Premotor and prefrontal correlates of category-related lexical retrieval. *Neuroimage*. 1998; 7:232–243. [PubMed: 9597664]
72. Eickhoff SB, Bzdok D, Laird AR, Kurth F, Fox PT. Activation likelihood estimation meta-analysis revisited. *Neuroimage*. 2012; 59:2349–2361. [PubMed: 21963913]
73. Martin A. The representation of object concepts in the brain. *Annu Rev Psychol*. 2007; 58:25–45. [PubMed: 16968210]
74. Ueno T, Saito S, Rogers TT, Lambon Ralph MA. Lichtheim 2: Synthesizing aphasia and the neural basis of language in a neurocomputational model of the dual dorsal-ventral language pathways. *Neuron*. 2011; 72:385–96. [PubMed: 22017995]
75. Visser M, Jefferies E, Lambon Ralph MA. Semantic processing in the anterior temporal lobes: A meta-analysis of the functional neuroimaging literature. *J Cogn Neurosci*. 2010; 22:1083–1094. [PubMed: 19583477]
76. Adlam ALR, et al. Semantic dementia and fluent primary progressive aphasia: two sides of the same coin? *Brain*. 2006; 129:3066–3080. [PubMed: 17071925]
77. Shimotake A, et al. 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*. 2015; 25:3802–17. [PubMed: 25491206]
78. Acosta-Cabronero J, et al. Atrophy, hypometabolism and white matter abnormalities in semantic dementia tell a coherent story. *Brain*. 2011; 134:2025–2035. [PubMed: 21646331]

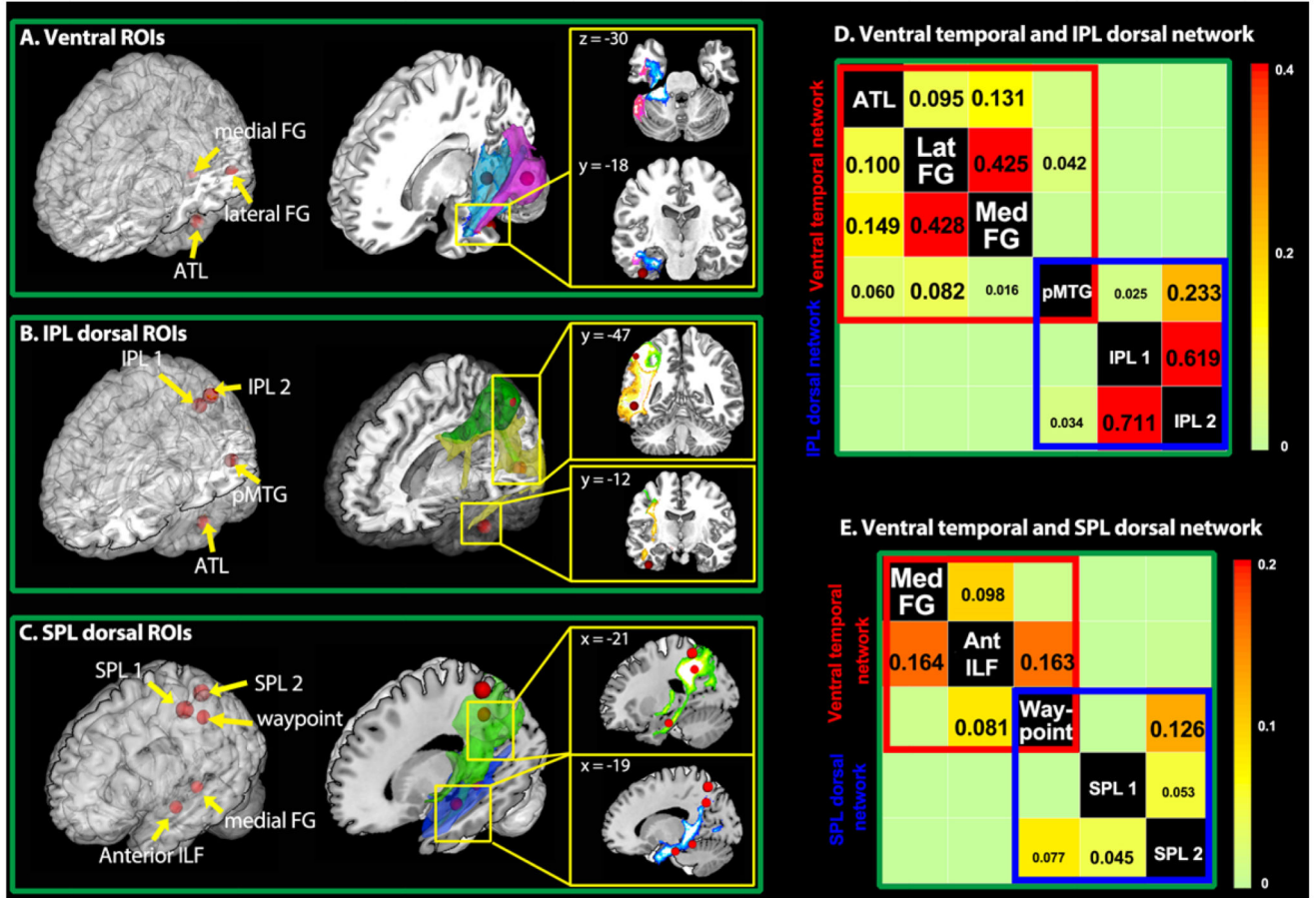
79. Lambon Ralph MA, Lowe C, Rogers TT. Neural basis of category-specific semantic deficits for living things: Evidence from semantic dementia, HSVE and a neural network model. *Brain*. 2007; 130:1127–1137. [PubMed: 17438021]
80. Binkofski F, Buxbaum LJ. Two action systems in the human brain. *Brain Lang*. 2013; 127:222–229. [PubMed: 22889467]
81. Humphreys GW, Riddoch MJ. Features, objects, action: The cognitive neuropsychology of visual object processing, 1984–2004. *Cogn Neuropsychol*. 2006; 23:156–183. [PubMed: 21049326]
82. Humphreys GF, Lambon Ralph MA. Fusion and Fission of Cognitive Functions in the Human Parietal Cortex. *Cereb Cortex*. 2015; 25:3547–3560. [PubMed: 25205661]
83. Jung J, Cloutman LL, Binney RJ, Ralph MAL. The structural connectivity of higher order association cortices reflects human functional brain networks. *Cortex*. 2016
84. Embleton KV, Haroon HA, Morris DM, Lambon Ralph MA, Parker GJM. Distortion correction for diffusion-weighted MRI tractography and fMRI in the temporal lobes. *Hum Brain Mapp*. 2010; 31:1570–1587. [PubMed: 20143387]
85. Materials and methods are available as supplementary materials on Science Online.
86. Binney RJ, Embleton KV, Jefferies E, Parker GJM, Lambon Ralph MA. 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*. 2010; 20:2728–38. [PubMed: 20190005]
87. Kravitz DJ, Saleem KS, Baker CI, Ungerleider LG, Mishkin M. The ventral visual pathway: an expanded neural framework for the processing of object quality. *Trends Cogn Sci*. 2013; 17:26–49. [PubMed: 23265839]
88. Binney RJ, Parker GJM, Lambon Ralph MA. Convergent connectivity and graded specialization in the rostral human temporal lobe as revealed by diffusion-weighted imaging probabilistic tractography. *J Cogn Neurosci*. 2012; 24:1998–2014. [PubMed: 22721379]
89. Schmahmann, JD., Pandya, D. *Fiber pathways of the brain*. Oxford University Press; 2009.
90. Bajada CJ, Lambon Ralph MA, Cloutman LL. Transport for language south of the Sylvian fissure: The routes and history of the main tracts and stations in the ventral language network. *Cortex*. 2015; 69:141–51. [PubMed: 26070011]
91. Bedny M, Caramazza A, Pascual-Leone A, Saxe R. Typical neural representations of action verbs develop without vision. *Cereb Cortex*. 2012; 22:286–93. [PubMed: 21653285]
92. Warrington EK, McCarthy R. Category specific access dysphasia. *Brain*. 1983; 106:859–878. [PubMed: 6652466]
93. Gotts S, Plaut DC. The impact of synaptic depression following brain damage: A connectionist account of “access/refractory” and “degraded-store” semantic impairments. *CABN*. 2002; 2:187–213. [PubMed: 12775185]
94. Noppeney U, et al. Temporal lobe lesions and semantic impairment: a comparison of herpes simplex virus encephalitis and semantic dementia. *Brain A J Neurol*. 2007; 130:1138–1147.
95. Campanella F, D’Agostini S, Skrap M, Shallice T. Naming manipulable objects: anatomy of a category specific effect in left temporal tumours. *Neuropsychologia*. 2010; 48:1583–97. [PubMed: 20144630]
96. Roberts, D. *Exploring the link between visual impairment and pure alexia*. University of Manchester; 2009.
97. Humphreys GW, Forde EME. Category specificity in mind and brain? *Behav Brain Sci*. 2001; 24:497–509.
98. Laiacona M, Capitani E, Barbarotto R. Semantic category dissociations: A longitudinal study of two cases. *Cortex*. 1997; 33:441–461. [PubMed: 9339328]
99. Pietrini V, et al. Recovery from herpes simplex encephalitis: selective impairment of specific semantic categories with neuroradiological correlation. *J Neurol Neurosurg Psychiatry*. 1988; 51:1284–1293. [PubMed: 3225585]
100. Damasio H, Tranel D, Grabowski T, Adolphs R, Damasio A. Neural systems behind word and concept retrieval. *Cognition*. 2004; 92:179–229. [PubMed: 15037130]

101. Plaut DC, Behrmann M. Complementary neural representations for faces and words: a computational exploration. *Cogn Neuropsychol*. 2011; 28:251–275. [PubMed: 22185237]
102. Farah MJ, McClelland JL. A computational model of semantic memory impairment: Modality-specificity and emergent category-specificity. *J Exp Psychol Gen*. 1991; 120:339–357. [PubMed: 1837294]
103. Badre D, Wagner A. Semantic retrieval, mnemonic control, and prefrontal cortex. *Behav Cogn Neurosci Rev*. 2002; 1:206–218. [PubMed: 17715593]
104. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015; 521:436–444. [PubMed: 26017442]
105. McClelland, JL., Rumelhart, DE., Hinton, GE. *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. Rumelhart, DE., McClelland, JL., the PDP Research Group. , editors. Vol. 1. MIT Press; 1986. p. 3-44.
106. Price CJ, Devlin JT, Moore CJ, Morton C, Laird AR. Meta-analyses of object naming: Effect of baseline. *Hum Brain Mapp*. 2005; 25:70–82. [PubMed: 15846820]
107. Eickhoff SB, et al. Coordinate-based activation likelihood estimation meta-analysis of neuroimaging data: a random-effects approach based on empirical estimates of spatial uncertainty. *Hum Brain Mapp*. 2009; 30:2907–2926. [PubMed: 19172646]
108. Turkeltaub PE, et al. Minimizing within-experiment and within-group effects in activation likelihood estimation meta-analyses. *Hum Brain Mapp*. 2012; 33:1–13. [PubMed: 21305667]
109. Oldfield RC. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*. 1971; 9:97–113. [PubMed: 5146491]
110. Cloutman LL, Binney RJ, Drakesmith M, Parker GJM, Lambon Ralph MA. The variation of function across the human insula mirrors its patterns of structural connectivity: Evidence from in vivo probabilistic tractography. *Neuroimage*. 2012; 59:3514–3521. [PubMed: 22100771]
111. Visser M, Lambon Ralph MA. Differential contributions of bilateral ventral anterior temporal lobe and left anterior superior temporal gyrus to semantic processes. *J Cogn Neurosci*. 2011; 23:3121–3131. [PubMed: 21391767]
112. Pobric G, Jefferies E, Lambon Ralph MA. Category-Specific versus Category-General Semantic Impairment Induced by Transcranial Magnetic Stimulation. *Curr Biol*. 2010; 20:964–968. [PubMed: 20451381]
113. Parker GJM, Alexander DC. Probabilistic anatomical connectivity derived from the microscopic persistent angular structure of cerebral tissue. *Philos Trans R Soc B Biol Sci*. 2005; 360:893–902.
114. Rohde DLT. LENS: the light, efficient network simulator. Technical Report CMU-CS-99-164. 1999
115. Rumelhart, DE., Hinton, GE., Williams, RJ. *Learning representations by back-propagating errors*. MIT Press; Cambridge, MA, USA: 1988.
116. Garrard P, Carroll E. Lost in semantic space: a multi-modal, non-verbal assessment of feature knowledge in semantic dementia. *Brain*. 2006; 129:1152–1163. [PubMed: 16585052]
117. Dixon MJ, Bub DN, Arguin M. The interaction of object form and object meaning in the identification performance of a patient with category-specific visual agnosia. *Cogn Neuropsychol*. 1997; 14:1085–1130.
118. Dixon MJ, Bub DN, Chertkow H, Arguin M. Object identification deficits in dementia of the Alzheimer type: combined effects of semantic and visual proximity. *J Int Neuropsychol Soc*. 1999; 5:330–345. [PubMed: 10349296]
119. Zevin JD, Seidenberg MS. Simulating consistency effects and individual differences in nonword naming: A comparison of current models. *J Mem Lang*. 2006; 54:145–160.

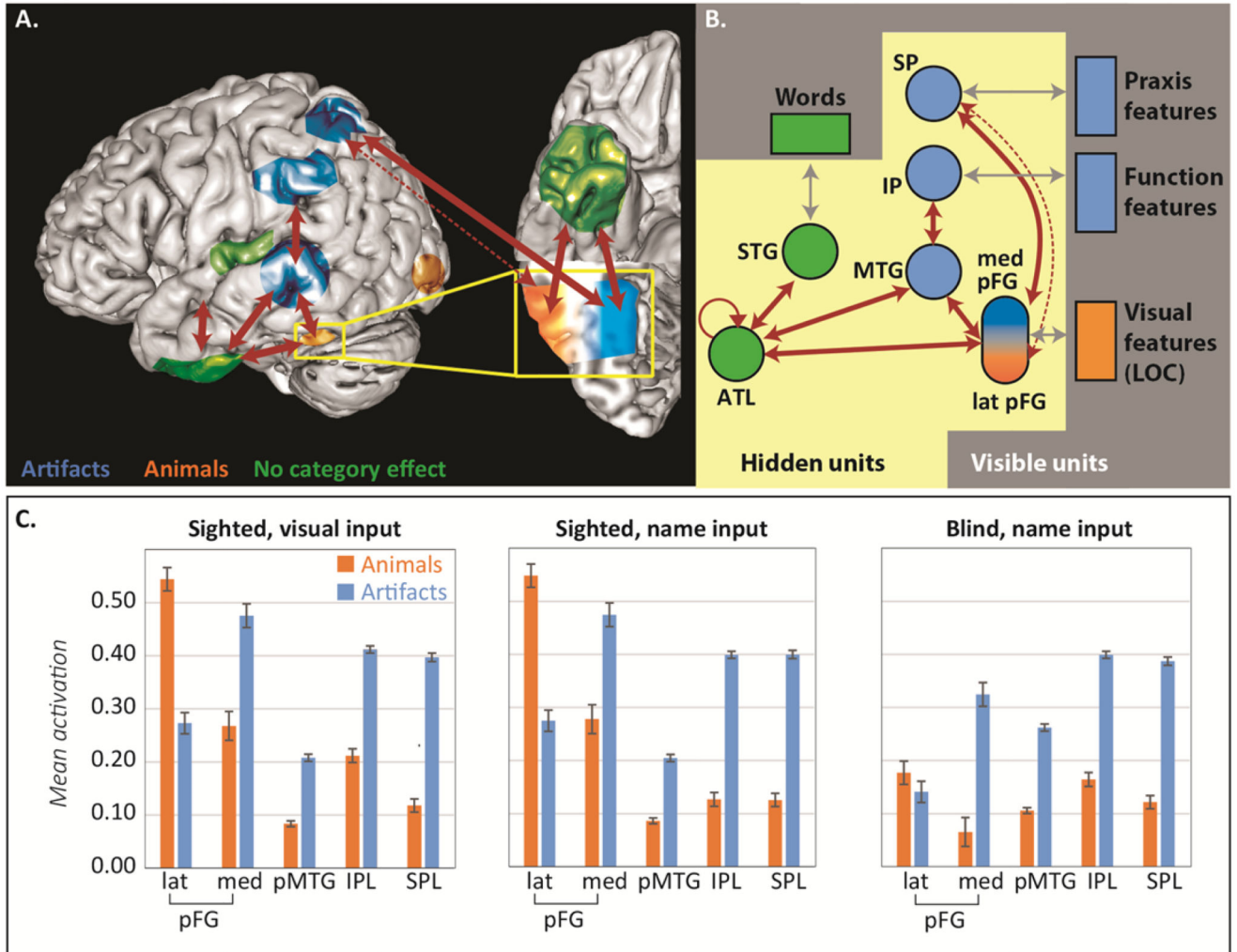


**Figure 1.**

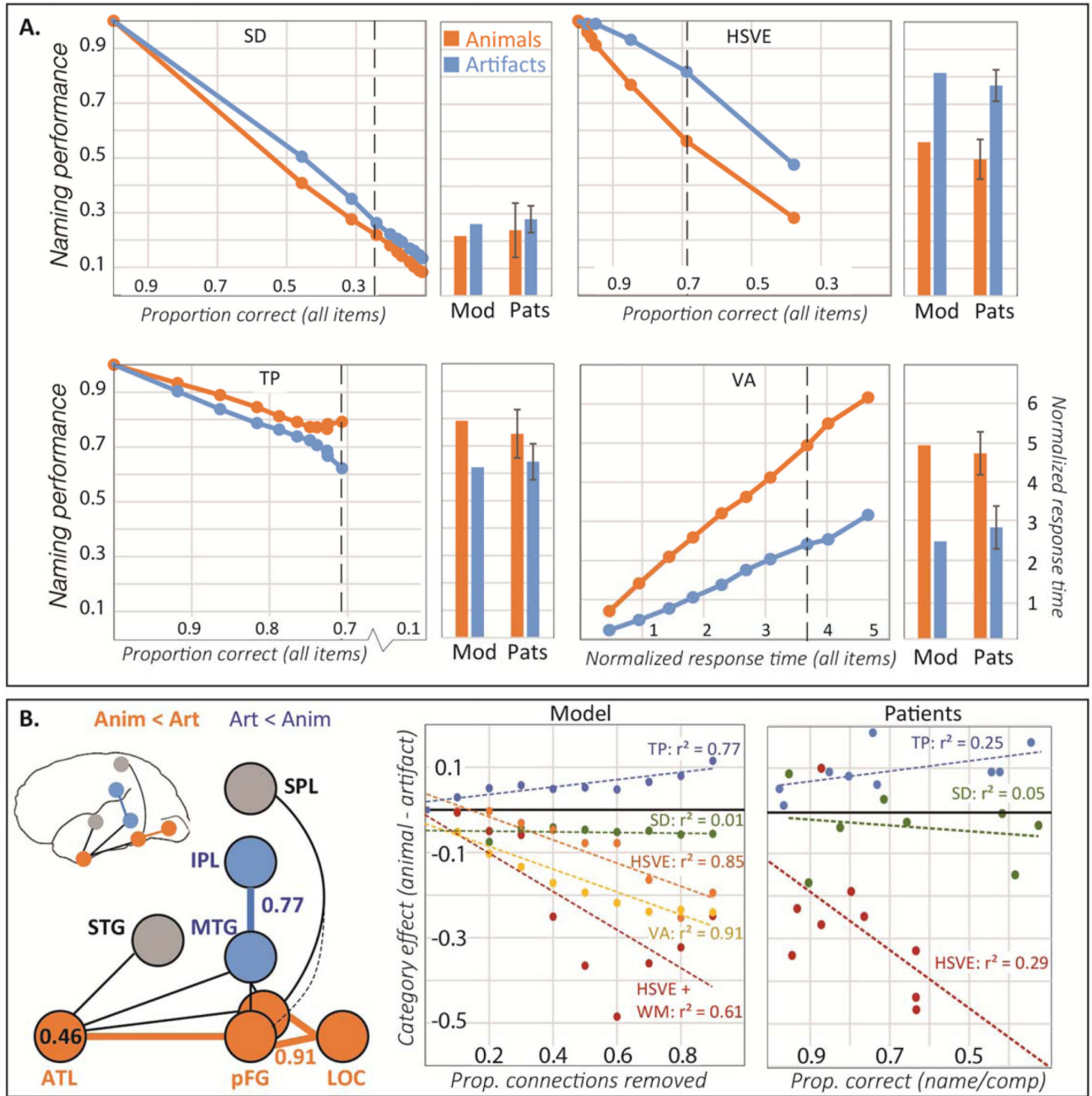
ALE analysis showing regions that systematically respond more to animals than artifacts (orange), more to artifacts than animals (blue), or equally to both (green). Red dots indicate seed points from activation likelihood estimation (ALE) analysis and literature review. IPL = inferior parietal lobe, SPL = superior parietal lobe, pFG = posterior fusiform gyrus, pMTG = posterior middle temporal gyrus, LOC = lateral occipital complex.



**Figure 2.** Tractography results. Red spheres indicate seed points from activation likelihood estimation (ALE) analysis and literature review. **(A)**. Streams from medial (blue) and lateral (pink) pFG project to ATL. **(B)**. Streams from pMTG (yellow) project to ATL and IPL, while IPL streams (green) project to pMTG but not ATL. **(C)**. Streams from inferior ATL white matter (blue) pass by medial pFG and branch superiorly, where they intersect SPL streamlines (green). The waypoint seed was placed at this intersection. **(D-E)**. Matrices showing significant connectivity of temporal regions with IPL regions via the pMTG and with SPL regions via the tract identified by the waypoint seed. Numbers indicate group-averaged probability estimates (0-1) from seed (column) to target (row) regions.



**Figure 3.** Model architecture and fMRI data simulations. **(A).** Schematic showing ALE and connectivity results. Red arrows indicate significant connectivity in tractography while colors indicate semantic category effects in the ALE analysis. The dotted arrow indicates that connectivity diminishes from medial to lateral pFG. **(B).** Architecture of the corresponding neural network model. Boxes indicate layers that directly encode features of objects (visible units) and circles indicate model analogs of cortical regions of interest where representations are learned (hidden units). For visible units, blue indicates more active features for animals than artifacts while orange indicates the reverse. For hidden units, circle color indicates expected category effects using the same scheme as panel A. Red arrows indicate model connections that correspond to tractography results; gray arrows indicate connections that mediate activation between visible and hidden units. **(C).** Mean unit activation for animals or artifacts in each model region of interest, for visual and word inputs of the “sighted” model (left and middle) and for word inputs in the “blind” model (right).



**Figure 4.** Results of patient simulations. **(A).** Line plots show model naming accuracy for animals and artifacts at ten increasing levels of damage for each disorder plotted against overall accuracy (all items). Dashed vertical lines indicate the damage level that most closely matches mean overall accuracy in the corresponding patient group. HSVE data are for the homogeneous damage model (HSVE); data for the asymmetric damage model (HSVE+) appear in Supplementary Figure S5. Barplots show accuracy by category for the model at this level compared to patient means/standard errors reported in 79,95,96. **(B).** Lesion-symptom

mapping results. Left: Layers/connections where lesion size predicts increasing artifact (blue) or animal (orange) disadvantage. Middle: Correlation between lesion size and category effect in each simulated patient group. Right: Category effect size in naming plotted against overall impairment as measured by word comprehension (SD and HSVE) or overall naming (TPT) in case-series studies of real patients.