

**A model of emergent category-specific activation in the  
posterior fusiform gyrus of sighted and congenitally blind  
populations**

Journal:	<i>Journal of Cognitive Neuroscience</i>
Manuscript ID:	JOCN-2014-0456.R1
Manuscript Type:	Original
Date Submitted by the Author:	n/a
Complete List of Authors:	Chen, Lang; University of Wisconsin-Madison, Department of Psychology Rogers, Timothy; University of Wisconsin-Madison, Department of Psychology;
Keywords:	Computational modeling, Memory: Long-term memory, Neuroimaging, Temporal cortex, Neurological disorders

Category specificity in posterior fusiform

1

**A model of emergent category-specific activation in the posterior fusiform gyrus of  
sighted and congenitally blind populations**

Lang Chen<sup>1,2,\*</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>Stanford Cognitive and Systems Neuroscience Laboratory, 1070 Arastradero Rd. Suite 220, Palo Alto, CA 94304, USA

Running title: Category specificity in posterior fusiform

\*Corresponding author:

Lang Chen  
Department of Psychology, University of Wisconsin-Madison, 1202 West Johnson Street, Madison, WI 53705, USA  
[lchen32@wisc.edu](mailto:lchen32@wisc.edu)

**Abstract**

Theories about the neural bases of semantic knowledge tend between two poles, one proposing that distinct brain regions are innately dedicated to different conceptual domains and the other suggesting that all concepts are encoded within a single network. Category-sensitive functional activations in the fusiform cortex of the congenitally blind have been taken to support the former view, but also raise several puzzles. We use neural network models to assess a hypothesis that spans the two poles: the interesting functional activation patterns reflect the base connectivity of a domain-general semantic network. Both similarities and differences between sighted and congenitally blind groups can emerge through learning in a neural network, but only in architectures adopting real anatomical constraints. Surprisingly, the same constraints suggest a novel account of a quite different phenomenon: the dyspraxia observed in patients with semantic impairments from anterior temporal pathology. The work suggests one reason why the cortical semantic network may be structured the way it is, which accords with existing proposals in some ways but differs in others.

Introduction

It is well known that knowledge of different conceptual categories (e.g., animals, tools, faces, landmarks, etc.) can be differentially impaired with brain damage (e.g., Hillis & Caramazza, 1991; Lambon Ralph, Patterson, Garrard, & Hodges, 2003; Warrington & Shallice, 1984) and can elicit different cortical responses in healthy adults (Chao, Haxby, & Martin, 1999; Hwang, Palmer, Basho, Zadra, & Müller, 2009; Martin & Chao, 2001; Martin, Wiggs, Ungerleider, & Haxby, 1996), but the importance of such findings for theories about the neural bases of conceptual knowledge remains unclear. For some, category sensitivity provides evidence of dissociable and innately constrained *domain-specific* knowledge systems in the mind and brain (Caramazza & Shelton, 1998; Kanwisher, 2010; New, Cosmides, & Tooby, 2007). Yet items from different conceptual domains vary in many of their sensory and motor characteristics, raising the alternative possibility that category sensitivity arises through learning about sensory-motor structure within a single *domain-general* knowledge system (for a review, see Chen & Rogers, 2014). The controversy reflects two fundamentally different views of the mechanisms that support everyday human knowledge.

Functional magnetic resonance imaging (fMRI) studies of the congenitally blind have recently offered new evidence important to the issue (Bedny, Caramazza, Pascual-Leone, & Saxe, 2012; He et al., 2013; Mahon, Schwarzbach, & Caramazza, 2010). In a seminal study, Mahon and colleagues (2009) compared activations in the posterior fusiform gyrus (pFG) of sighted and congenitally blind adults while they made semantic judgments from spoken names of animals or artifacts. Both groups showed an activation gradient across the lateral/medial axis, with manmade objects, relative to animals, evoking more activation medially and less laterally in the pFG, consistent with the well-documented pattern in sighted populations for pictorial stimuli (Chao et al., 1999; Chouinard & Goodale, 2010; Mahon et al., 2007) and, less frequently, for spoken and written words (Devlin, Rushworth, & Matthews, 2005; Mahon et al., 2009). Thus category-sensitive activation patterns can arise in the ventral visual stream in the absence of visual experience. This result appears to seriously challenge domain-general views of semantics that have attributed posterior fusiform organization to

Category specificity in posterior fusiform

4

the influence of visual experience (e.g., Rogers, Hocking, Mechelli, Patterson, & Price, 2005).

To explain the pattern, Mahon et al (2009; 2007) proposed that functional specialization in ventral temporal cortex is constrained, not solely by patterns of visual input, but also by innate connectivity between ventral temporal cortex and other areas. Blind and sighted participants show similar category effects because they have the same initial connectivity, which is configured to support distinct category-specific knowledge networks. In this sense, the data are offered as support for the hypothesis that evolutionary pressures have led to independent domain-specific knowledge systems in the brain (Caramazza and Shelton, 1998).

This *distributed domain-specific* (DDS) view can be viewed as advancing two separate proposals: first, that functional specificity in cortex is partly shaped by native patterns of connectivity, and second, that such connectivity produces a domain-specific organization of the cortical semantic network. In several elegant follow-up studies, Mahon and colleagues have provided substantial evidence for the former proposal (Almeida, Mahon, & Caramazza, 2010; Bedny et al., 2012; He et al., 2013), and this work is of a piece with a broader movement in the literature to understand how cortical network connectivity shapes functional specialization across many cognitive domains (Bouhali et al., 2014; Martin, 2006; Plaut & Behrmann, 2011). We believe it represents a significant advance in the effort to understand the organization of the cortical semantic network.

The second proposal of the DDS hypothesis, however, raises some non-trivial puzzles. First, Mahon et al. (2007) observed both similarities and differences between sighted and blind participants; the DDS hypothesis clearly explains the similarities but not the differences. Second, the mechanisms that give rise to the proposed domain-specific organization of cortex remain underspecified, beyond the general hypothesis that connectivity is important. The authors describe connectivity as being “concentrated around” distinct conceptual domains (Mahon and Caramazza, 2009) or “at the grain of a limited number of evolutionarily relevant domains of knowledge” (Caramazza & Mahon, 2003), but it is not clear what specific commitments these statements entail. Third, it is difficult to see how the DDS hypothesis might explain phenomena that challenge any domain-specific approach to

Category specificity in posterior fusiform semantic representation. One such phenomenon is the existence of domain-general semantic impairments resulting from circumscribed cortical pathology, such as that observed in semantic dementia (SD)—a progressive dementing illness that progressively erodes semantic knowledge for all conceptual domains and all modalities of reception and expression (Adlam et al., 2006; Hodges, Patterson, Oxbury, & Funnell, 1992; Lambon Ralph, Lowe, & Rogers, 2007) These deficits are caused by cortical thinning and hypometabolism circumscribed within ventro-lateral anterior temporal cortex (Guo et al., 2013). It is difficult to see how localized pathology of this kind would produce domain-general knowledge impairments if semantic knowledge is supported by multiple independent domain-specific networks. Relatedly, while items from distinct semantic domains certainly differ in many of their properties, they can also share important characteristics. Thus animals and tools are physical objects, subject in similar ways to forces of gravity, momentum and inertia; hawks and gliders can hang suspended in the air for similar reasons; cows are living animals but are used for food and to make leather. If animals, foods, and manmade objects are supported by distinct knowledge systems, how might knowledge about a living cow generalize to the butchered cow? Domain-specific approaches to knowledge representation account for differences in knowledge structures across domains, but struggle to explain cross-domain knowledge generalization (see Rogers & McClelland, 2004; Rogers & McClelland, 2008; Rogers & Plaut, 2002 for discussion).

The current paper considers whether one important contribution of the DDS hypothesis—the proposal that ventral temporal representations are partly shaped by native connectivity—can explain the interesting imaging data in sighted and blind populations, within a domain-general semantic knowledge system. Our approach builds on the “semantic hub” theory that was developed in part to explain the domain-general knowledge impairments observed in SD (Lambon Ralph et al., 2007; Patterson, Nestor, & Rogers, 2007; Rogers et al., 2004). In common with many others, this view proposes that conceptual knowledge arises from learned associations among sensory, motor, and linguistic representations localized in different brain regions and communicating via pathways in a distributed cortical network. While many pathways exist in the network, the theory further proposes

Category specificity in posterior fusiform 6  
that there exists a single cross-modal “hub” located in the anterior temporal lobes (ATL) that is  
important for mediating interactions among many different surface modalities, for all domains of  
knowledge. By virtue of such mediation, the hub comes to encode representations that express  
conceptual similarity structure and so support conceptual generalization (Rogers et al., 2004). A key  
assumption of the approach is that the same “hub” neurons contribute to the representation of all  
conceptual domains, so that they can express both differences and commonalities across domains.  
Consequently knowledge of all conceptual domains is disrupted with neuropathology to the hub, as  
observed in SD (Lambon Ralph et al., 2007; Noppeney et al., 2007)

We here elaborate this view by adopting a central thesis of the DDS hypothesis, also presaged in  
earlier work by Plaut (2002). Specifically, we propose that different pathways in the network vary in  
the strength or effectiveness of their *connectivity*. Regions that are highly effectively interconnected  
will exert strong mutual influences and so will respond similarly to given inputs. The activation  
patterns generated within such network will thus reflect, not only the effects of learning in a  
structured environment, but also the base connectivity of the network. Specifically, category-  
sensitive activations in pFG arise because the medial aspect of this region is more effectively  
connected to parietal regions that support object-associated manipulation knowledge (Mahon, Kumar,  
& Almeida, 2013). In contrast to the DDS proposal, however, we further hypothesize that such  
patterns can arise within a system that remains domain general insofar as critical elements contribute  
to semantic processing for all conceptual domains. Such an account would thus reconcile the  
challenges faced by both domain-general and domain-specific approaches to semantic knowledge.

Though the hypothesis itself is straightforward, its assessment is not. Existing neuro-  
computational models of semantic knowledge have been connected only loosely (Lambon Ralph et  
al., 2007; Rogers et al., 2004) or not at all (Cree, McNorgan, & McRae, 2006; Farah & McClelland,  
1991; Tyler, Moss, Durrant-Peatfield, & Levy, 2000) to the anatomy of the cortical semantic network.  
Though useful for understanding patterns of behavior in healthy and disordered populations, such  
models cannot shed light on the anatomical distribution of functional activation patterns evoked by

Category specificity in posterior fusiform semantic tasks; indeed, no prior computational model has been applied to the understanding of fMRI data in semantic tasks. Second, our working hypothesis requires us to specify how visual, verbal, and praxic representations interact within the semantic network, yet prior semantic models have mainly focused on interactions between visual and verbal representations. Third, the introduction of praxic representations must allow for an account of the dyspraxic patterns commonly observed in patients with semantic knowledge impairments (e.g., Hodges, Bozeat, Lambon Ralph, Patterson, & Spatt, 2000), yet no such model has previously been advanced.

We addressed these challenges by investigating three different extrapolations of the ATL-hub model of semantic processing (Lambon Ralph et al., 2007; Patterson et al., 2007; Rogers et al., 2004). The hub model aligns coarsely with the anatomy of the temporal lobe insofar as phonological representations of words are supported by the superior temporal gyrus (Catani, Jones, & ffytche, 2005; Davis, Di Betta, Macdonald, & Gaskell, 2009; Graves, Desai, Humphries, Seidenberg, & Binder, 2010), visual representations of objects are supported by posterior infero-temporal cortex (Goodale & Milner, 1992; Rogers et al., 2005), and these regions are reciprocally connected with anterior temporal cortex where different sensory streams converge (Binney, Parker, & Ralph, 2012). The three models presented here, however, differ in their proposals about how praxic representations encoded in left lateral parietal cortex (Binkofski & Buxbaum, 2013; Buxbaum, Kyle, Grossman, & Coslett, 2007) interact with the rest of the network, with two motivating factors.

The first is neuropsychological: along with other modalities of reception and expression, knowledge about object praxis is seriously degraded in patients with SD. In particular, SD patients show impairments acting on common objects or matching them for shared praxis or function, but spared knowledge of action affordances in their use of novel tools and in components of praxis directed toward everyday objects (Bozeat, Lambon Ralph, Patterson, & Hodges, 2002; Hodges et al., 2000). The preservation of systematic information (such as visuo-praxic affordances) with degradation of more idiosyncratic information (such as unaffordanced elements of praxis) is a hallmark of SD in visual and verbal tasks (Patterson et al., 2006); consequently prior articulations of



Category specificity in posterior fusiform 8  
the “hub” theory have proposed that praxic representations interact directly with the ATL hub so that  
praxic knowledge degrades in the same way as visual and verbally-expressed knowledge with ATL  
pathology (Rogers & Patterson, 2007).

The second motivation stems from studies of connectivity between pFG and lateral parietal (LP)  
regions that support object-directed praxis (dorso-dorsal stream) and long-term object use  
representations (ventro-dorsal stream; Binkofski & Buxbaum, 2013). Contra the proposal from  
neuropsychology, human diffusion-weighted tractography does not reveal direct anatomical  
connectivity between ATL and LP (Binney et al., 2012; Caspers et al., 2011; Wakana, Jiang, Nagae-  
Poetscher, van Zijl, & Mori, 2004), but there is considerable evidence for an indirect pathway  
projecting through the medial pFG. Mahon and colleagues (2014; 2013; 2007) have reported  
significant functional connectivity between medial pFG and parietal areas including the  
supramarginal gyrus and a more superior region (Mahon, Kumar and Almeida, 2013). Tractography  
in non-human primates shows long projections from lateral parietal cortex into the ventral temporal  
lobe (Seltzer & Pandya, 1994; Zhong & Rockland, 2003), and human white-matter tractography  
suggests that the dorsal aspect of the inferior fronto-occipital fasciculus (IFOF), which runs medially  
along the temporal lobe, branches dorsally at a point near pFG to terminate in superior parietal  
radiations (Wakana et al., 2004). These observations are all consistent with the view that medial pFG  
connects with parietal regions that support object-directed action. Moreover, both medial and lateral  
FG send white-matter projections to the ATL (Binney et al., 2012).

The three models we consider (see Figure 1) all learn associations among visual, verbal, and  
praxic representations, but vary in the architecture through which these associations are encoded. The  
*hub-only* model incorporates a direct pathway between praxic representations and the ATL hub; the  
*fully-connected* model includes this pathway as well as a path connecting visual and praxic regions  
directly; and the *anatomically-constrained* model includes the visual-praxic pathway suggested by  
connectivity studies without direct connection between praxic and hub regions. To capture  
lateral/medial differences in pFG connectivity, visuo-praxic connections in the fully-connected and

Category specificity in posterior fusiform anatomically-constrained models were anatomically graded so that “medial” visual units interacted more effectively with praxic units than did “lateral” visual units.

In what follows, we first show that the similarities and differences in functional activations evoked in sighted and blind individuals emerge in just one of the architectures. We then consider how well each architecture explains dyspraxic patterns arising from ATL atrophy, with results that validate conclusions drawn from the simulation of imaging results. The simulations illustrate a mechanism by which category-sensitive patterns can arise from network connectivity as suggested by the DDS hypothesis, but also suggest a somewhat different explanation as to why the system adopts a particular pattern of connectivity in the first place.

**Experiment 1: Simulating category-sensitivity in healthy and congenitally blind groups**

The goal of the first experiment was to assess the ability of each hypothesis to account for the category-sensitive patterns of functional activation in the pFG of sighted and congenitally blind populations.

**Methods**

*Model architectures*

The three model architectures shown in Figure 1 were implemented using the light efficient network simulator (LENS) (Rohde, 1999). Each model included three visible layers that directly encoded, as distributed patterns of activation over units, pre-specified representations corresponding to the visual appearance of an object, the object’s name and verbal description, and its associated praxis. 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) and lateral parietal cortex (LP, praxic hidden units). An additional hidden layer served as the ATL hub (ATL, cross-modal units). All units employed a sigmoidal activation

Category specificity in posterior fusiform function and were given a fixed bias of -3 so that, in the absence of input from other units, they tended to adopt a low activation state.

The models varied in how the hidden layers connected. For the hub-only and fully-connected models, all modality-specific hidden layers were reciprocally connected with the ATL units. For the fully-connected and anatomically-constrained models, pFG units were reciprocally connected with LP units. To capture the proposed gradation in connectivity effectiveness along this pathway, units in the pFG layer were situated along an anatomical lateral-to-medial axis. The error derivatives (and hence the strength of influence on weight changes) on visuo-praxic connections was scaled according to the visual unit's position along this axis, with medial-most units having a larger scaling factor than other connections in the model and the rate diminishing for increasingly lateral units according to a sigmoid function:

$$S(i) = 2 / (1 + e^{\frac{n}{2} - i})$$

where  $S(i)$  is the scaling parameter applied on error derivatives,  $n$  is the number of units in vision layer, and  $i$  (the unit location) ranges from 1 to 20 on the medial-lateral axis. Across units the mean learning rate was equal to that on other connections in the model.

### *Training environment*

A model environment was constructed containing visual, verbal, and praxic representations for 24 different animals and 24 different tools, with each domain organized into 4 basic categories, each containing 6 exemplars. Visual representations for each item were generated stochastically in accordance with the constraints identified by Rogers et al. (2004) with animals sharing more visual properties overall than tools,  $M_{animal} = 7.58$ ,  $SD = 1.38$  and  $M_{tool} = 4.17$ ,  $SD = 1.20$  (McRae, de Sa, & Seidenberg, 1997). Thus (1) items in different domains shared few properties, (2) items within the same category shared many properties, (3) animals from different categories shared more properties than did artifacts from different categories, and (4) animals had more visual properties than did artifacts. Each item was also given a unique name represented by a single unit and a set of 10

Category specificity in posterior fusiform features as verbal descriptors which were designed to capture linguistic environment for describing animal and tool objects. The number of verbal features was largely comparable between categories (see Table 1).

Praxic representations were also constructed for each item, taking the form of distributed patterns over the 10 visible units in the visible praxic layer. For all animal items, all praxic units were simply turned off, capturing the intuition that most animals are not associated with rich praxic information. For artifacts, praxic representations were generated for each tool by copying the item’s visual properties and flipping the state of each unit with a small probability ( $p=0.2$ ), producing a model analog of partial visual affordance: each visual tool feature reliably predicted a particular praxic feature across items, but this prediction was violated for some items. Consequently, the visual and praxic features of tool category had a correlation of .65 overall.

*Tasks and training procedures*

Given partial information about an item as input, the models were trained to activate the item’s associated properties, including its name, verbal description, and visual and praxic features. Model inputs could be single names, verbal descriptions, visual images, or praxic features. Inputs were applied to visible units by providing these with direct excitatory input, and units throughout the network were updated successively over time in random order. Time was discretized with an integration constant of 30, so that each unit updated its state 30 times in one cycle. After 8 update cycles, target values were applied to visible units and weights were updated using backpropagation with a learning rate of 0.01 (scaled as noted above for visuo-praxic connections) and weight decay of 0.0005 without momentum (Rumelhart, Hinton, & Williams, 1988) . While fully-sighted model variants were trained with all varieties of input and outputs, 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.

The response to a given input was counted as correct if, across all visible units, target properties were active above 0.5 and non-target below 0.5. All models, both sighted and blind, were trained for

Category specificity in posterior fusiform

12

100k epochs at which point they generated correct output by this criterion for the great majority of inputs. Therefore, the sighted and blind variants of this network were equated on the total amount of non-visual experience.<sup>1</sup>

### *Simulation of functional brain imaging results*

The brain imaging studies involved tasks in which participants made a semantic judgment from either a picture or the spoken name of a familiar item. To simulate these tasks, input corresponding to either an item's visual appearance (for picture comprehension) or its name (for word comprehension) was applied to the model, which then cycled until unit activations stopped changing. Given that each unit was associated with a large negative bias and its initial activation was rather low, after settling, unit activations were recorded and thus taken as an analog of the BOLD response at a single voxel. This value was then distorted with Gaussian noise ( $\mu = 0$ ,  $\delta^2 = 0.1$ ) to reflect the error in signal estimation intrinsic to brain imaging methods.

For each architecture and model population (sighted/blind) 15 different model training runs were conducted, each initialized with a different random set of weights, to simulate 15 different subjects. Our analysis then followed Mahon et al.'s (2009) procedure. The response of each unit was averaged across items in each condition (Tool/Animal) and then spatially smoothed with a Gaussian kernel ( $\mu = 0$ ,  $\delta^2 = 1$ ) encompassing two adjacent units. A group-level contrast was performed to find the peak activation for both Animal and Tool concepts using the averaged data across the 15 model subjects. An ROI analysis was then performed on the activation value of the peak unit averaged with two neighboring units on either side. This procedure was used to generate the results shown in Figure 2.

To assess the gradient of the category effect along the lateral/medial axis of the model analog to posterior fusiform, we again replicated the approach of Mahon et al. (2009). At each unit and for each model subject, we computed the t statistic contrasting activation for Animals versus Tools. Across model subjects, we then computed confidence limits on this t statistic at each unit location

---

<sup>1</sup> We also considered "blind" models given more extensive training with non-visual patterns, but their behavior was qualitatively very similar to the "blind" models reported here so we do not discuss these further.

Category specificity in posterior fusiform  
along the axis. This analysis was used to generate Figure 3.

**Results and Discussions**

On model analogs of the tasks of interest, both sighted and blind models for all three architectures generated the correct pattern of activation over output units with greater than 90% accuracy across all model training runs. We then assessed whether the category-sensitive activation patterns reported in pFG from prior work arise in the visual hidden layer of each model, using the approach of Mahon et al. (2009) described earlier. Results are shown in Figure 2 and Table 2.

In the hub only architecture, unsurprisingly, no category effect was observed for fully sighted models in the lateral peak for either task. The medial peak showed reliably greater activation for animals than for artifacts in these simulations, likely reflecting the greater number of visual features contained in the visual patterns representing animals. In the blind models, pFG activation was low for both categories. Overall the pattern was very different from that observed in the imaging data.

Both fully-connected and anatomically-constrained models replicated the category-sensitive pattern for words and images in fully sighted participants, but the fully-connected model did not replicate the pattern observed in the congenitally blind: greater activation arose in medial relative to lateral units, with no category effect. In contrast, the anatomically constrained model replicated the pattern well: both lateral and medial peaks were moderately active, and the category effect was observed in medial but not lateral peaks.

We next considered how category-sensitive activation patterns varied across the lateral-to-medial axis of the pFG layer. As shown in Figure 3, Mahon et al. (2009) found that the contrast of animal to tool activation shifted gradually along this axis. In sighted individuals, greater activation for animals gives way to the reverse pattern as one moves from lateral toward medial pFG voxels (Figure 3D for pictorial stimuli and 3E for word stimuli). The congenitally blind participants showed a similar gradient with an important qualitative difference: no category effect was observed in lateral pFG, but this null effect gradually gave way to a reliable effect for tools in medial pFG (Figure 3F).

We computed analogous curves for the anatomically-constrained and fully-connected model variants (left and right columns of Figure 3). Qualitatively similar patterns were observed for both sighted models, and the gradients in each correlated strongly with those in the imaging study ( $r^2$  between 0.75 and 0.89 for all comparisons). In the blind models, however, the fully-connected architecture failed to capture the key pattern: only a very small effect was observed in the medial-most extent of the pFG layer, and the model gradient correlated only weakly with the observed gradients ( $r^2 = 0.249$ ). In contrast, the anatomically-constrained model again provided a good match to the data: no category effect was observed in the lateral-most units; a strong effect was observed in the medial-most units; and the model gradient correlated strongly ( $r^2 = 0.697$ ) with the observed gradient.

Note that the ATL hub plays the same functional role in all three networks: it mediates interactions between visual and verbal representations, with all units contributing to the representation of all items. That is, it serves as a domain-general cross-modal hub. The simulations thus illustrate that category-sensitive patterns in pFG can emerge within an interactive neural network that remains domain-general, depending upon the base connectivity of the network.

The contrasting patterns arise from differences in the way activation propagates in the three networks. The blind models must learn associations among names of objects, their verbal descriptions, and their praxis. When all surface representations communicate directly with the ATL units, the hub can directly encode all of these associations. The pFG units never get input and never drive outputs, and so can remain largely silent without affecting system behavior. When praxic representations do not connect directly to the ATL hub, the pFG units indirectly mediate these interactions and so become active. The same forces that cause category-sensitive patterns in the sighted models then come into play: greater connectivity effectiveness between LP and medial pFG units, coupled with richer praxis for artifacts, cause the medial pFG units to participate more strongly in the representations of artifacts (for similar thoughts, see Almeida, Fintzi, & Mahon, 2013; Mahon et al., 2007).



The model also, however, explains why lateral pFG regions show greater activation for animals in sighted individuals but not in the congenitally blind. In the sighted models, medial pFG units contribute more to the representation of artifacts, so the remaining units must “pick up the slack” in communicating visual information about animals to the ATL hub. Consequently the lateral units come to more strongly activate for animals, even though lateral and medial pFG units are all equally well connected with the ATL hub. In the blind models, there is no visual information about animals to communicate to the hub—the pFG units only serve to communicate interactions between the hub and the praxic representations. Thus pFG units that are strongly connected to the ATL hub but not to the praxic representations—the lateral units—come to respond equally strongly to items from both domains.

**Experiment 2: Simulating dyspraxia from ATL pathology**

Though results from Experiment 1 provide an account of the brain imaging data, they also raise a puzzle. The direct connection between praxis and the ATL hub was motivated by the dyspraxic pattern arising from ATL pathology in SD (Bozeat et al., 2002; Hodges et al., 2000). If praxic representations do not communicate directly with the ATL hub, and do receive direct connections from pFG, what explains this phenomenon? Answering this question was the goal of Experiment 2.

The key data stem from studies of performance on two different kinds of object-use tasks in SD. In the object-use task, participants demonstrate the use of an everyday object, and features of the elicited action are coded as correct or incorrect based on comparison to a control group. Praxic features are considered affordanced if, across many objects, they reliably covary with a particular visual property, and as unaffordanced otherwise, and separate accuracy scores are tabulated for these property types (Bozeat et al., 2002). The central findings are that patients with SD show increasing impairment generating the correct praxis (Figure 4A), but that unaffordanced features are much more vulnerable than afforded features (Figure 5A).



Category specificity in posterior fusiform

16

The second task requires participants to solve simple mechanical puzzles by selecting the correct novel tool from amongst 4 options and using it to extract a block from its container. Participants are scored on whether they (1) select the correct novel tool and (2) use it successfully to extract the block. Despite their impaired praxic knowledge for everyday objects, patients with SD perform normally at this task (Figure 7).

Thus the goal of Experiment 2 was to understand why ATL pathology in SD causes (1) an overall loss of praxic knowledge for everyday objects, but with (2) sparing of affordanced elements of praxis and (3) a spared ability to use novel objects in mechanical-puzzle tasks. As noted earlier, previous work has suggested that dyspraxia in SD arises because praxic representations interact directly with the ATL hub, yet Experiment 1 suggested that the functional imaging results of interest will only arise when praxic representations communicate indirectly with the hub, via the medial pFG—a hypothesis that is also consistent with recent studies of functional and structural connectivity in cortex. We therefore assessed the ability of each architecture to explain dyspraxia in SD by simulating the effects of the disorder on model analogs of the object-use and mechanical-puzzle tasks.

## Methods

The sighted variants of the same trained models from Experiment 1 were used in this investigation.

*Simulating the object-use task.* To simulate the object-use task, the visual patterns for familiar items were applied across visual input units and the pattern generated across the praxis layer was inspected. Overall performance for each item was scored as correct if all praxic features were on the correct side of 0.5 and incorrect otherwise (for similar approach, see Harm & Seidenberg, 2004).<sup>2</sup> Additionally, the praxic features associated with each item were classified as either affordanced or unaffordanced, according to the criteria of Bozeat et al. (2002). Specifically, praxic features that, in the model training environment, strongly correlated with an item's visual properties were deemed

---

<sup>2</sup> This scoring protocol is very commonly used in neural network modeling; however other common protocols produced a qualitatively similar pattern of behavior.

Category specificity in posterior fusiform 17  
“affordanced,” while praxic properties not so correlated were deemed “unaffordanced.” We then  
computed, across all familiar items and at each level of damage, the proportion of affordanced and  
unaffordanced features correctly activated above a threshold of 0.5.

*Simulating the mechanical puzzle task.* To simulate the visual appearance of novel tools we  
generated new visual input patterns over features common to artifacts. Target praxic patterns were  
also created for each tool, consisting of the praxic features that covaried with the item’s visual  
features in the training environment. One praxic feature for each novel item was deemed important  
for solving each puzzle. To simulate tool selection in a particular puzzle, the visual inputs for 4 novel  
tools, only one possessing the critical praxic feature, were applied to the model in succession. The  
model was considered to have chosen whichever tool most strongly activated the critical feature. Use  
of the selected item was subsequently scored as correct if the model successfully activated all target  
features and no non-target features above a threshold of 0.5 in the praxic output layer. Fourteen trials  
were constructed using the 20 novel tools by arranging them in different groups of 4.

*Simulation of ATL pathology in SD.* To simulate ATL pathology observed in SD, the models  
were subjected to increasingly severe “lesions” created by randomly removing an increasing  
proportion of units in the ATL hub layer. We investigated removal of 5% - 30% of units in 5%  
increments, producing 6 levels of impairments to the cross-modal representations (Devlin,  
Gonnerman, Andersen, & Seidenberg, 1998). In each administration of damage to each model, we  
assessed model performance on analogs of the object-use and mechanical puzzle tasks.

**Results and Discussions**

Model performance was compared to the patient data reported in both Hodges et al. (2000) and  
Bozeat et al. (2002) and is reported in Table 3. Figure 4 shows patient and model performance for  
each architecture on the object-use task. Correct praxis always declined steadily with increasing  
disease severity, even in the anatomically-constrained architecture where praxic representations are  
disconnected from the hub. Much larger effects were observed in the hub-only and fully-connected

Category specificity in posterior fusiform

18

architectures, however: serious dyspraxia arose when only 5% of hub units were removed, contrasting with the patient data where comparably poor praxis is observed only in late stages of the disease.

Figure 5 shows patient and model performance for affordanced vs. unaffordanced praxic features plotted against overall severity. Affordanced features fared better than unaffordanced features in all cases, but with qualitative differences across architectures: both property types declined equally rapidly in hub-only models, while unaffordanced properties were much more vulnerable in the anatomically constrained models. The fully-connected model showed an intermediate pattern, but with significant loss for both feature types.

We quantified these observations by computing the difference in mean accuracy for affordanced versus unaffordanced properties at each level of damage and correlating this with disease severity. Bozeat et al. (2002) observed that the advantage for affordanced properties grew larger with increasingly severe semantic impairment, with a correlation of  $r^2 = 0.86$ ,  $p < 0.01$  across patients. This result was closely matched in the anatomically-constrained model ( $r^2 = 0.88$ ,  $p < 0.01$ ), but not the other two architectures (both  $r$ 's  $\sim 0.5$ , n.s.; Figure 6).

Finally, Figure 7 shows patient and model data in the simulation of tool-selection and tool-use in the mechanical puzzle task. Patients performed as well as controls on average in both aspects of this task. The same was true of the anatomically-constrained model, but the other two architectures both showed significant impairment with ATL pathology, averaged across all levels of damage.

Thus while all three architectures exhibited increasing dyspraxia with greater ATL pathology, only the anatomically constrained model captured the preserved knowledge of affordanced praxis for familiar objects, and the preserved ability to select and use novel objects.

These results raise two questions. First, why does ATL pathology disrupt vision-to-praxis mappings for familiar items even in the anatomically constrained model, where praxic representations do not communicate with the damaged hub? The pattern arises because the network is interactive. Activation from visual input flows toward both the hub and the praxic representations,

Category specificity in posterior fusiform and then echoes back to the visual hidden units. Distorted feedback from the damaged hub causes the visual patterns to differ from those arising in the healthy model. The differences propagate to the praxic units, causing some to deactivate and others to incorrectly activate. Greater pathology in ATL produces greater distortion of visual representations, hence greater disruption to praxic patterns generated through the intact direct pathway.

Second, why are afforded properties (and novel items that exploit afforded properties) only spared in the anatomically constrained model? Affordances arise from the systematic co-occurrence of visual and praxic features, correlations that do not always align with conceptual structure. Objects with straight cylindrical handles—hammers, dustpans, spatulas, swords—afford a similar grasp but need not be conceptually related. Moreover their semantic neighbors—nails, brooms, whisks, shields—may lack either or both properties. Thus affordances may cross-cut semantic category structure in important ways.

Weights in a direct path between visual object and praxis representations can support robust co-activation of correlated visual-praxic pairs. When this mapping must go through the ATL hub, these correlations are less robustly encoded—representations in the hub express conceptual structure, and so cannot exploit visual-praxic correlations that cross-cut this structure. Thus when hub representations degrade, knowledge of both afforded and unafforded praxic properties are rapidly lost. In the anatomically-constrained model, disrupted feedback from the ATL can distort visual representations, but because the direct path from vision to praxis robustly encodes affordances, this distortion mainly affects knowledge of unafforded properties. The ability to retrieve afforded properties of both familiar and novel items remains unaffected, even with severe ATL pathology.

**General Discussion**

The simulations show that the category-sensitive activation patterns seen in both sighted and blind populations can arise from learning in an interactive neural network when (a) there exists a direct but graded pathway between pFG and LP such that medial aspects of pFG connect more

Category specificity in posterior fusiform

20

effectively than lateral aspects (Mahon et al., 2007; 2013), and (b) there is no direct pathway from LP to the ATL hub, as suggested by classical anatomy and recent diffusion tractography studies (Binney et al., 2012; Caspers et al., 2011; Wakana et al., 2004). A neural network model adopting this connectivity explained both the similarities and differences between sighted and blind populations, while two models adopting different connectivity amongst the same pools of units did not capture the full pattern of imaging results. The results suggest that category-sensitive activation reflect joint effects of learning and base connectivity in a cortical network that is not innately organized by conceptual domain.

Perhaps surprisingly, the same model also captured the dyspraxic pattern caused by anterior temporal pathology in SD. To explain dyspraxia in SD, prior work has suggested that praxic representations communicate directly with the ATL hub (Bozeat et al., 2002; Hodges et al., 2000; Lambon Ralph et al., 2007; Patterson et al., 2007). Though models adopting this architecture did show serious praxic impairments with simulated ATL damage, they failed to capture preserved knowledge of affordanced features of familiar items or the preserved ability to select and use novel tools in simple mechanical puzzles (Bozeat et al., 2002; Hodges et al., 2000). These patterns suggest that knowledge of systematic visuo-praxic relations remains robust when concepts degrade in SD. The anatomically-constrained model explains why: correlations between visual and praxic properties are encoded by weights in the direct path from vision to praxis, which remain unaffected by pathology in SD.

These results are important both in providing the first model of dyspraxia in SD and also in validating the conclusions drawn from the imaging simulations. A critic might reasonably view the anatomically constrained model as having been “tailored” to produce category-sensitive activation in pFG layer, but the same cannot be said of the dyspraxia results: the anatomically-constrained model directly contradicts prior proposals about the architecture needed to explain dyspraxia in SD. The fact that this model best captures this very different set of data thus validates the hypothesis about network connectivity developed to explain the imaging data.

Finally, the simulations suggest a novel hypothesis about the forces that constrain initial connectivity in the semantic network. Conceptual and affordance structure, though both important for aspects of behavior, are independent. Conceptual structure promotes knowledge generalization, as when the carpenter realizes she can use either a hammer or a nailgun to build a wall, despite the different praxis these objects engage. Affordance knowledge is critical for interactions with objects generally, and with novel items particularly: the first time one encounters a nailgun, one must infer that the item is to be gripped and pointed differently from the hammer, despite sharing the hammer’s function. Conceptual structure, arguably, does not cohere well with affordance structure. A network in which there exist separate pathways for extracting knowledge of affordances versus concepts may best promote both kinds of generalization.

*Implications for domain-specific approaches to knowledge representation*

As noted in the introduction, our proposal adopts many elements central to the DDS view developed by Mahon and Caramazza (2009, 2011). It also resonates with domain-specific proposals more generally in invoking initial constraints to explain category-sensitivity in brain activation. Thus our model might be viewed as an implementation or extension of the DDS hypothesis, and as generally consistent with domain-specific approaches. Our model also differs in important respects, however, from both the DDS and other domain-specific views articulated in different writings. Specifically, on our view:

Different network components are interactive and not functionally independent. To explain the apparent double-dissociation of knowledge for different semantic categories, domain-specific approaches often propose that separate knowledge systems in the mind and brain are functionally independent: the healthy functioning of one system does not depend on the intact functioning of other systems (e.g. Hillis and Caramazza, 1991; Caramazza & Shelton, 1998). Such an argument relies on the modularity assumption adopted in classical cognitive neuropsychology (e.g., Caramazza, 1984). Our model proposes interactive rather than modular processing, with computations in different network components mutually interdependent. In the current context, interactivity explains

Category specificity in posterior fusiform

22

why anterior temporal pathology causes praxic impairments for familiar items. More generally, interactive processing is a core claim of the domain-general approach to semantic representation and processing (Rogers et al., 2004; Rogers & McClelland, 2014). Double-dissociations can arise in interactive networks, but do not transparently reveal the functional organization of the system (Farah & McClelland, 1991; Plaut, 1995). On our reading, the question of functional independence is unclear in the DDS. On one hand, (Mahon & Caramazza, 2008) argue that domain-specific organization does not preclude interactive processing, but elsewhere they argue that category-specific knowledge impairments arise from selective damage to category-specific knowledge systems (Mahon & Caramazza, 2009, 2011), a view that seems to hew closely to the modular assumptions of classical cognitive neuropsychology.

Processing for all domains depends upon a central cross-modal hub. Some domain-specific approaches emphasize that different cortical regions are dedicated to different conceptual domains and/or different modalities, with no single region important for all kinds of concepts across all modalities (Caramazza & Shelton, 1998; Damasio, Grabowski, Tranel, Hichwa, & Damasio, 1996). This makes it difficult to understand neuropsychological syndromes like SD, in which circumscribed pathology in the ATL erodes conceptual knowledge for all domains across all modalities. Articulations of the DDS hypothesis have remained silent on this issue, and appear to make no mention of this well-studied syndrome (see e.g. Mahon and Caramazza, 2009). Though category sensitivity emerges in our model in the visuo-praxic pathway, the network still adopts a cross-modal hub important for conceptual representation across all domains and so offers an account of such findings as investigated extensively in prior work (Bozeat et al., 2002; Hodges et al., 2000).

Learning and experience are important. Some domain-specific accounts of knowledge acquisition downplay the role of learning and experience in advocating for innate knowledge modules, especially in studies of conceptual development (Carey & Spelke, 1994). Cognitive neuropsychology and cognitive neuroscience have generally been more open to the likelihood that neural organization is highly plastic, but the precise role of learning in domain-specific approaches to



Category specificity in posterior fusiform knowledge in the brain remains unclear. In different articulations of the DDS hypothesis, the authors clearly allow a considerable role for learning and experience (e.g., Mahon & Caramazza, 2011), but also suggest their view is compatible with developmental approaches that eschew such influences (e.g., Mahon & Caramazza, 2009). In our model, initial constraints on connectivity and the effects of learning and experience jointly shape emergent category sensitivity, explaining both the similarities and differences between sighted and congenitally blind groups. We also note that initially-specified connectivity does not preclude the possibility that white-matter connectivity itself can be partly shaped by experience (Scholz, Klein, Behrens, & Johansen-Berg, 2009; Tang et al., 2010), further highlighting the potentially important role of experience.

Category effects are overlapping, graded and fuzzy. Because category effects in our model arise from learning and experience they are not all-or-nothing but overlapping and graded. Within the model, artifacts that are not frequent targets of praxis will not strongly drive the visuo-praxic system. For individuals who interact physically with animals (e.g. farmers), these animals may drive the visuo-praxic system and so elicit similar patterns to manipulable manmade objects. Moreover units in medial FG that activate more strongly for tools may nevertheless still contribute somewhat to the representation of animals and vice versa, similar to Plaut’s (2002) account of graded functional specialization in optic aphasia. This perspective seems very different from a domain-specific view under which different neural systems are natively tailored to only process a particular conceptual domain and do not contribute to other domains. Whether our view is consistent with the DDS hypothesis in this respect depends upon issues noted in the introduction—in particular, on what is intended by the proposal that initial connectivity is “concentrated around” different conceptual domains. If the claim is that a given tract only propagates information about items in a given domain, this seems clearly in opposition to our proposal. Likewise if the claim is that initial connectivity expresses innate knowledge of property co-occurrences—for instance, knowledge that things with faces likely can move on their own and have emotional states, which might be reflected by innate connectivity between face, motion, and limbic regions—this also differs from our proposal, under



Category specificity in posterior fusiform

24

which such co-occurrences are learned. If, however, the claim is simply that different conceptual domains come to rely more on some tracts than others by virtue of the surface representations they engage, this seems more consistent with our proposal, though less consistent with earlier domain-specific views.

Gross connectivity may be innately specified, but knowledge is not. Domain-specific approaches in cognitive development propose that infants are born with knowledge that allows them to construe items as belonging to distinct core conceptual domains, such as physical objects, animals, agents, or social beings (for discussion, see Carey & Spelke, 1994; Mahon & Caramazza, 2009). Under our proposal, the gross connectivity of the cortical network is initially specified but this does not reflect innate knowledge about distinct domains and their properties. Indeed, prior to learning our model treats all items as similar to one another. The proposed constraints on initial connectivity instead promote, with experience, efficient acquisition of different kinds of internal representations, specifically conceptual and affordance-based representations. To us it is unclear whether the initial constraints on connectivity proposed by the DDS hypothesis constitute innate knowledge as proposed in nativist approaches to development, or whether these constraints operate only in conjunction with learning and experience as in our proposal.

Initial constraints distinguish conceptual from affordance structure. Finally, initial connectivity in our model promotes acquisition of two different kinds of knowledge structures—conceptual structures and affordance structures. Contra the domain-specific view, these do not align with different semantic categories. Conceptual structures describe the degree to which different items are similar in kind, and thus promote conceptual generalization, for all semantic domains. Affordance structures describe systematic relations between an item's visual appearance and the kinds of actions that it can support. Because tools are frequently associated with praxis, they will strongly drive this pathway, but the affordance system should also support action directed toward items in other semantic domains. For instance, while faces are indeed important properties of animals and must contribute to their semantic representation, they can also be targets of praxis, as anyone who brushed

Category specificity in posterior fusiform 25  
crumbs from a child’s cheek or kissed their partner can attest. Clearly actions are also directed  
toward foods and toward other body parts—categories that, under some domain-specific views, are  
supported by distinct neural systems (Caramazza & Shelton, 1998). Thus while our proposal accords  
with the DDS in suggesting that initial connectivity supports acquisition of different kinds of  
knowledge, the particular structures invoked are very different: not distinct semantic domains, but  
distinct similarity structures.

In sum, we believe our proposal occupies a useful middle ground between strict domain-specific  
views and blank-slate domain-general views, similar to both the DDS hypothesis and to recent  
proposals in other cognitive domains (Plaut & Behrmann, 2011; Ueno, Saito, Rogers, &  
Lambon Ralph, 2011; Zevin & Seidenberg, 2006). We further believe, however, that it differs from  
the DDS hypothesis in several respects and so is best viewed as an alternative proposal rather than as  
a model implementation of that view.

*Implications for the ATL “hub” theory*

Despite invoking separate pathways for visual-praxic and visuo-conceptual interactions, our  
model remains consistent with the ATL hub theory. The theory’s central claim is that a cross-modal  
hub is needed to learn the conceptual structure that governs knowledge generalization for all  
conceptual domains. We have argued that behavioral interactions with objects require knowledge of a  
different kind of structure—affordance structure—that can often cross-cut conceptual structure.  
Where different behaviors depend upon cross-cutting statistical structure, it makes sense that  
different pathways should be involved in encoding those structures. This proposal is broadly similar  
to arguments developed in the context of reading (Harm & Seidenberg, 2004; Plaut, McClelland,  
Seidenberg, & Patterson, 1996) and spoken language processing (Ueno et al., 2011), where the  
systematic structure existing amongst perceptual/motor representations of words cross-cuts the  
conceptual structure expressed by word meaning. To exploit both forms of structure, contemporary  
models likewise propose separate pathways for encoding the different mappings.

*Implications for other theories of category-specificity*

Finally, we note that the current proposals about category-sensitive functional activations in the brain are similar in spirit to the well-known sensory/functional hypothesis about category specific semantic impairment observed in some forms of neuropathology. This view, first articulated in Warrington and Shallice's classic work (1984) and subsequently developed in computer simulations by Farah and McClelland (1991), proposed that knowledge of different semantic domains might rely differentially on knowledge of perceptual versus functional properties of items in the environment. This view was appealing partly because it explained the interesting neuropsychological patterns with reference to a division of labor in the brain that was already well-established—specifically, the anatomical division between sensory and motor systems. There was no need to postulate additional cortical and cognitive modules dedicated to particular semantic kinds to explain the selective disruption of particular categories of knowledge. The sensory-functional view came under fire for a variety of reasons in the 1990's (see Caramazza & Shelton, 1998), but there is now considerable evidence from both patient and functional imaging work for an elaborated variant of this perspective, under which there are not merely two but many different kinds of sensory and motor properties represented in different regions throughout cortex (see Chen & Rogers, 2014). The current work illustrates how differential reliance on different kinds of sensory-motor properties, coupled with additional constraints arising from patterns of connectivity amongst the regions that code these properties, can explain a range of phenomena in both brain imaging and semantic disorders. We view the work as highly consistent with the original motivations for the sensory-functional hypothesis.

#### *Limitations and future directions*

Our simulations illustrate how base network connectivity can produce category-sensitive activation patterns in both sighted and blind populations, within a network that is not innately organized by conceptual domain. In so doing it also provides a novel account of dyspraxia in semantic dementia. We view these contributions as promising, but the model is also limited in several respects.

First, the literature contains several additional claims of category- and modality-sensitive activation patterns in sighted and blind groups (Bedny et al., 2012; Mahon et al., 2010). Some results are highly provocative, such as the finding that the “visual word-form area” of congenitally blind individuals becomes active during tactile reading (Reich, Szwed, Cohen, & Amedi, 2011). In focusing on mappings among visual, verbal, and praxic representations of objects, our model omits many other kinds of information and so cannot directly address these new findings. An important question for future work will be whether a better elaborated model that incorporates haptic and orthographic representations can extend the same principles to these and other related findings (Yildirim & Jacobs, 2013).

Second, our simulations motivate further study of connectivity between ventral temporal and parietal regions. We have argued that medial pFG connects to praxic representations in the parietal cortex, but object-directed action also involves knowledge of the functions to which objects are put. Like conceptual structure, functional structure varies independently of praxis (Boronat et al., 2005; Buxbaum, Schwartz, & Carew, 1997; Schwartz et al., 1998)—whereas the typewriter and piano engage similar movements toward different ends, the typewriter and pen engage different movements toward similar ends (Kellenbach, Brett, & Patterson, 2003). Functional properties critically constrain conceptual representations of artifacts—so much so that artifacts are often viewed as changing in kind when they change in function (Ahn, 1998; Keil, 1992). Thus, whereas praxic representations may not strongly constrain the representations that arise in the ATL hub, function representations likely do.

Knowledge about praxis and function can doubly dissociate with brain damage, and evoke different activation patterns within parietal cortex (Buxbaum & Saffran, 2002; Schwartz et al., 1998). Buxbaum (Binkofski & Buxbaum, 2013; Boronat et al., 2005; Buxbaum & Saffran, 2002) thus proposes a division of labor within lateral parietal cortex, with a “dorso-dorsal” stream mainly supporting praxic knowledge and an “ventro-dorsal” stream mainly supporting functional knowledge. Both superior and inferior aspects of the lateral parietal lobe appear to engage in tool concept

Category specificity in posterior fusiform

28

processing (Almeida et al., 2010; Anzellotti, Mahon, Schwarzbach, & Caramazza, 2011; Canessa et al., 2008; Noppeney, Price, Penny, & Friston, 2006) and show category-specific activation in naming (Chouinard & Goodale, 2010). If inferior regions like supra-marginal gyrus mainly encode tool-related function, while praxis is mainly supported by more superior regions, an important test for our hypothesis concerns the nature of the connectivity between medial pFG and superior versus inferior parietal regions. Our framework suggests that medial pFG should be more effectively connected to praxic (superior parietal) than functional (inferior) representations.

On our reading the literature here is somewhat mixed. Mahon has emphasized functional connectivity between medial pFG and supra-marginal gyrus in two studies (Garcea & Mahon, 2014; Mahon et al., 2007), but also reports significant non-zero connectivity from medial pFG to more superior parietal cortex in a third (2013, Figures 2 and 4B). To our knowledge there is no evidence from any method for direct white-matter projections between medial pFG and supra-marginal gyrus. Animal studies (Seltzer & Pandya, 1994; Zhong & Rockland, 2003) have found a long projection from the inferior parietal lobe into various temporal regions, though with little evidence of a lateral/medial gradient. In humans, both dissection studies (Martino, Brogna, Robles, Vergani, & Duffau, 2010) and diffusion tensor imaging (Wakana et al., 2004) have shown white-matter pathways that branch up from medial infero-posterior temporal cortex to terminate in superior parietal regions. We believe this supports the face validity of our proposal, though this is clearly an avenue for further research.

Third, theoretical approaches to knowledge representation in the brain have been strongly influenced by patterns of spared and impaired semantic memory across different neuropsychological groups. The current model builds on an existing account of semantic impairments in SD and herpes simplex viral encephalitis, but some semantic deficits—most notably cases in which knowledge of manmade objects is more impaired than knowledge of animals—have not been explored in this framework. A model that incorporates both praxic and functional representations, and whose architecture is constrained by what is known about the relevant connectivity of cortex, may allow us to better understand the full spectrum of semantic knowledge impairments observed in

Category specificity in posterior fusiform neuropsychology. One interesting direction for the future research could be investigating how semantic impairment of different neuropsychological groups may or may not be modulated by their premorbid experience (i.e., congenitally blind patients). Although these individual cases would be rather difficult to encounter, a recent study has demonstrated that this approach can be extremely fruitful for answering important theoretical questions (Jefferies, Rogers, & Lambon Ralph, 2011).

Finally, our training environment consists of artificial representations that capture important aspects of category structure within and across modalities. Though widely used in simulation research (Lambon Ralph et al., 2007; Plaut, 1997; Rogers et al., 2004) such patterns are highly abstracted from the real sensory-motor structure of the world. The development of models that can operate on representations derived from real images (Gale, Done, & Frank, 2001; Plaut & Behrmann, 2011) may provide further insight into the relation between learning, connectivity, and representational structure.

**Conclusion**

We have developed an approach to knowledge representation in the brain under which activation patterns generated by words or pictures are influenced both by learning and by the native connectivity of the network architecture. The approach explains similarities and differences in activation patterns observed in sighted and congenitally blind participants, and further accounts for the dyspraxia arising from anterior temporal pathology in SD. The work provides a new perspective on the interaction of nature and nurture in conceptual knowledge acquisition, suggesting that evolution may have shaped the base connectivity of the cortical knowledge network so as to promote efficient learning of both conceptual and affordance structure.

**Acknowledgements**

This work was conducted with generous support from the University of Wisconsin-Madison in the form of a University Fellowship to the first author and a Romnes Award to the second author. We



Category specificity in posterior fusiform

30

wish to thank Bradford Mahon and Matthew Lambon Ralph for kindly sharing their empirical data with us.

## References

- Adlam, A. L. R., Patterson, K., Rogers, T. T., Nestor, P. J., Salmond, C. H., Acosta-Cabronero, J., et al. (2006). Semantic dementia and fluent primary progressive aphasia: two sides of the same coin? *Brain*, 129(11), 3066-3080.
- Ahn, W.-k. (1998). Why are different features central for natural kinds and artifacts?: The role of causal status in determining feature centrality. *Cognition*, 69(2), 135-178.
- Almeida, J., Fintzi, A. R., & Mahon, B. Z. (2013). Tool manipulation knowledge is retrieved by way of the ventral visual object processing pathway. *Cortex*, 49(9), 2334-2344.
- Almeida, J., Mahon, B. Z., & Caramazza, A. (2010). The Role of the Dorsal Visual Processing Stream in Tool Identification. *Psychological Science*, 21(6), 772-778.
- Anzellotti, S., Mahon, B. Z., Schwarzbach, J., & Caramazza, A. (2011). Differential activity for animals and manipulable objects in the anterior temporal lobes. *Journal of Cognitive Neuroscience*, 23(8), 2059-2067.
- Bedny, M., Caramazza, A., Pascual-Leone, A., & Saxe, R. (2012). Typical Neural Representations of Action Verbs Develop without Vision. *Cerebral Cortex*, 22(2), 286-293.
- Binkofski, F., & Buxbaum, L. J. (2013). Two action systems in the human brain. *Brain and Language*, 127(2), 222-229.
- Binney, R. J., Parker, G. J., & Ralph, M. A. L. (2012). Convergent connectivity and graded specialization in the rostral human temporal lobe as revealed by diffusion-weighted imaging probabilistic tractography. *Journal of Cognitive Neuroscience*, 24(10), 1998-2014.
- Boronat, C. B., Buxbaum, L. J., Coslett, H. B., Tang, K., Saffran, E. M., Kimberg, D. Y., et al. (2005). Distinctions between manipulation and function knowledge of objects: evidence from functional magnetic resonance imaging. *Cognitive Brain Research*, 23(2-3), 361-373.
- Bouhali, F., de Schotten, M. T., Pinel, P., Poupon, C., Mangin, J.-F., Dehaene, S., et al. (2014). Anatomical connections of the visual word form area. *The Journal of Neuroscience*, 34(46), 15402-15414.
- Bozeat, S., Lambon Ralph, M. A., Patterson, K., & Hodges, J. R. (2002). When objects lose their meaning: What happens to their use? *Cognitive, Affective, & Behavioral Neuroscience*, 2(3), 236-251.
- Buxbaum, L. J., Kyle, K., Grossman, M., & Coslett, B. (2007). Left inferior parietal representations for skilled hand-object interactions: evidence from stroke and corticobasal degeneration. *Cortex*, 43(3), 411-423.
- Buxbaum, L. J., & Saffran, E. M. (2002). Knowledge of object manipulation and object function: dissociations in apraxic and nonapraxic subjects. *Brain and Language*, 82(2), 179-199.
- Buxbaum, L. J., Schwartz, M. F., & Carew, T. G. (1997). The role of semantic memory in object use. *Cognitive Neuropsychology*, 14(2), 219-254.
- Canessa, N., Borgo, F., Cappa, S. F., Perani, D., Falini, A., Buccino, G., et al. (2008). The different neural correlates of action and functional knowledge in semantic memory: an fMRI study. *Cerebral Cortex*, 18(4), 740-751.
- Caramazza, A. (1984). The logic of neuropsychological research and the problem of patient classification in aphasia. *Brain and Language*, 21(1), 9-20.
- Caramazza, A., & Mahon, B. Z. (2003). The organization of conceptual knowledge: The evidence from category-specific semantic deficits. *Trends in Cognitive Sciences*, 7(8), 354-361.
- Caramazza, A., & Shelton, J. R. (1998). Domain-specific knowledge systems in the brain: The animate-inanimate distinction. *Journal of Cognitive Neuroscience*, 10(1), 1-34.
- Carey, S., & Spelke, E. (1994). Domain-specific knowledge and conceptual change. *Mapping the*

mind: Domain specificity in cognition and culture, 169-200.

Caspers, S., Eickhoff, S. B., Rick, T., von Kapri, A., Kühlen, T., Huang, R., et al. (2011). Probabilistic fibre tract analysis of cytoarchitectonically defined human inferior parietal lobule areas reveals similarities to macaques. *Neuroimage*, 58(2), 362-380.

Catani, M., Jones, D. K., & ffytche, D. H. (2005). Perisylvian language networks of the human brain. *Annals of Neurology*, 57(1), 8-16.

Chao, L. L., Haxby, J. V., & Martin, A. (1999). Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. *Nature Neuroscience*, 2(10), 913.

Chen, L., & Rogers, T. T. (2014). Revisiting domain - general accounts of category specificity in mind and brain. *Wiley Interdisciplinary Reviews: Cognitive Science*, 5(3), 327-344.

Chouinard, P. A., & Goodale, M. A. (2010). Category-specific neural processing for naming pictures of animals and naming pictures of tools: An ALE meta-analysis. *Neuropsychologia*, 48(2), 409.

Cree, G. S., McNorgan, C., & McRae, K. (2006). Distinctive features hold a privileged status in the computation of word meaning: Implications for theories of semantic memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(4), 643-658.

Damasio, H., Grabowski, T. J., Tranel, D., Hichwa, R. D., & Damasio, A. R. (1996). A neural basis for lexical retrieval. *Nature*, 380(6574), 499-505.

Davis, M. H., Di Betta, A. M., Macdonald, M. J., & Gaskell, M. G. (2009). Learning and consolidation of novel spoken words. *Journal of Cognitive Neuroscience*, 21(4), 803-820.

Devlin, J. T., Gonnerman, L. M., Andersen, E. S., & Seidenberg, M. S. (1998). Category-Specific Semantic Deficits in Focal and Widespread Brain Damage: A Computational Account. *Journal of Cognitive Neuroscience*, 10(1), 77-94.

Devlin, J. T., Rushworth, M. F. S., & Matthews, P. M. (2005). Category-related activation for written words in the posterior fusiform is task specific. *Neuropsychologia*, 43(1), 69-74.

Farah, M. J., & McClelland, J. L. (1991). A computational model of semantic memory impairment: Modality specificity and emergent category specificity. *Journal of Experimental Psychology: General*, 120(4), 339-357.

Gale, T. M., Done, D. J., & Frank, R. J. (2001). Visual crowding and category specific deficits for pictorial stimuli: A neural network model. *Cognitive Neuropsychology*, 18(6), 509-550.

Garcea, F. E., & Mahon, B. Z. (2014). Parcellation of left parietal tool representations by functional connectivity. *Neuropsychologia*, 60(0), 131-143.

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, 15(1), 20-25.

Graves, W. W., Desai, R., Humphries, C., Seidenberg, M. S., & Binder, J. R. (2010). Neural Systems for Reading Aloud: A Multiparametric Approach. *Cerebral Cortex*, 20(8), 1799-1815.

Guo, C. C., Gorno-Tempini, M. L., Gesierich, B., Henry, M., Trujillo, A., Shany-Ur, T., et al. (2013). Anterior temporal lobe degeneration produces widespread network-driven dysfunction. *Brain*, 136(10), 2979-2991.

Harm, M. W., & Seidenberg, M. S. (2004). Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. *Psychological Review*, 111(3), 662-720.

He, C., Peelen, M. V., Han, Z., Lin, N., Caramazza, A., & Bi, Y. (2013). Selectivity for large nonmanipulable objects in scene-selective visual cortex does not require visual experience. *Neuroimage*, 79, 1-9.

Hillis, A. E., & Caramazza, A. (1991). Mechanisms for accessing lexical representations for output: Evidence from a category-specific semantic deficit. *Brain and Language*, 40(1), 106-144.

Hodges, J. R., Bozeat, S., Lambon Ralph, M. A., Patterson, K., & Spatt, J. (2000). The role of conceptual knowledge in object use Evidence from semantic dementia. *Brain*, 123(9), 1913-1925.

Hodges, J. R., Patterson, K., Oxbury, S., & Funnell, E. (1992). Semantic dementia progressive fluent



- aphasia with temporal lobe atrophy. *Brain*, 115(6), 1783-1806.
- Hwang, K., Palmer, E. D., Basho, S., Zadra, J. R., & Müller, R.-A. (2009). Category-specific activations during word generation reflect experiential sensorimotor modalities. *Neuroimage*, 48(4), 717-725.
- Jefferies, E., Rogers, T. T., & Lambon Ralph, M. A. (2011). Premorbid expertise produces category-specific impairment in a domain-general semantic disorder. *Neuropsychologia*, 49(12), 3213-3223.
- Kanwisher, N. (2010). Functional specificity in the human brain: A window into the functional architecture of the mind. *Proceedings of the National Academy of Sciences of the United States of America*, 107(25), 11163-11170.
- Keil, F. C. (1992). *Concepts, kinds, and cognitive development*: MIT Press.
- Kellenbach, M. L., Brett, M., & Patterson, K. (2003). Actions speak louder than functions: The importance of manipulability and action in tool representation. *Journal of Cognitive Neuroscience*, 15(1), 30-46.
- Lambon Ralph, M. A., Lowe, C., & Rogers, T. T. (2007). Neural basis of category-specific semantic deficits for living things: Evidence from semantic dementia, HSVE and a neural network model. *Brain*, 130, 1127-1137.
- Lambon Ralph, M. A., Patterson, K., Garrard, P., & Hodges, J. R. (2003). Semantic dementia with category specificity: A comparative case-series study. *Cognitive Neuropsychology*, 20(3-6), 307-326.
- Mahon, B. Z., Anzellotti, S., Schwarzbach, J., Zampini, M., & Caramazza, A. (2009). Category-Specific Organization in the Human Brain Does Not Require Visual Experience. *Neuron*, 63(3), 397-405.
- Mahon, B. Z., & Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Physiology-Paris*, 102(1-3), 59-70.
- Mahon, B. Z., & Caramazza, A. (2009). Concepts and categories: A cognitive neuropsychological perspective. *Annual Review of Psychology*, 60(1), 27-51.
- Mahon, B. Z., & Caramazza, A. (2011). What drives the organization of object knowledge in the brain? *Trends in Cognitive Sciences*, 15(3), 97-103.
- Mahon, B. Z., Kumar, N., & Almeida, J. (2013). Spatial Frequency Tuning Reveals Interactions between the Dorsal and Ventral Visual Systems. *Journal of Cognitive Neuroscience*, 1-10.
- Mahon, B. Z., Milleville, S. C., Negri, G. A. L., Rumiat, R. I., Caramazza, A., & Martin, A. (2007). Action-related properties shape object representations in the ventral stream. *Neuron*, 55(3), 507-520.
- Mahon, B. Z., Schwarzbach, J., & Caramazza, A. (2010). The Representation of Tools in Left Parietal Cortex Is Independent of Visual Experience. *Psychological Science*, 21(6), 764-771.
- Martin, A. (2006). Shades of Déjerine—Forging a Causal Link between the Visual Word Form Area and Reading. *Neuron*, 50(2), 173-175.
- Martin, A., & Chao, L. L. (2001). Semantic memory and the brain: structure and processes. *Current Opinion in Neurobiology*, 11(2), 194-201.
- Martin, A., Wiggs, C. L., Ungerleider, L. G., & Haxby, J. V. (1996). Neural correlates of category-specific knowledge. *Nature*, 379(6566), 649-652.
- Martino, J., Brogna, C., Robles, S. G., Vergani, F., & Duffau, H. (2010). Anatomic dissection of the inferior fronto-occipital fasciculus revisited in the lights of brain stimulation data. *Cortex*, 46(5), 691-699.
- McRae, K., de Sa, V. R., & Seidenberg, M. S. (1997). On the nature and scope of featural representations of word meaning. *Journal of Experimental Psychology: General*, 126(2), 99-130.
- New, J., Cosmides, L., & Tooby, J. (2007). Category-specific attention for animals reflects ancestral priorities, not expertise. *Proceedings of the National Academy of Sciences*, 104(42), 16598-

- 16603.
- Noppeney, U., Patterson, K., Tyler, L. K., Moss, H., Stamatakis, E. A., Bright, P., et al. (2007). Temporal lobe lesions and semantic impairment: a comparison of herpes simplex virus encephalitis and semantic dementia. *Brain*, 130(4), 1138-1147.
- Noppeney, U., Price, C. J., Penny, W. D., & Friston, K. J. (2006). Two distinct neural mechanisms for category-selective responses. *Cerebral Cortex*, 16(3), 437-445.
- 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(12), 976-987.
- Patterson, K., Ralph, M. A. L., Jefferies, E., Woollams, A., Jones, R., Hodges, J., et al. (2006). "Presemantic" cognition in semantic dementia: Six deficits in search of an explanation. *Journal of Cognitive Neuroscience*, 18(2), 169-183.
- Plaut, D. C. (1995). Double dissociation without modularity: Evidence from connectionist neuropsychology. *Journal of Clinical and Experimental Neuropsychology*, 17(2), 291-321.
- Plaut, D. C. (1997). Structure and function in the lexical system: Insights from distributed models of word reading and lexical decision. *Language & Cognitive Processes*, 12, 765-806.
- Plaut, D. C., & Behrmann, M. (2011). Complementary neural representations for faces and words: A computational exploration. *Cognitive Neuropsychology*, 28(3-4), 251-275.
- Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*, 103(1), 56-115.
- Reich, L., Szwed, M., Cohen, L., & Amedi, A. (2011). A ventral visual stream reading center independent of visual experience. *Current Biology*, 21(5), 363-368.
- Rogers, T. T., Hocking, J., Mechelli, A., Patterson, K., & Price, C. (2005). Fusiform Activation to Animals is Driven by the Process, Not the Stimulus. *Journal of Cognitive Neuroscience*, 17(3), 434-445.
- Rogers, T. T., Lambon Ralph, M. A., Garrard, P., Bozeat, S., McClelland, J. L., Hodges, J. R., et al. (2004). Structure and deterioration of semantic memory: A neuropsychological and computational investigation. *Psychological Review*, 111(1), 205-235.
- Rogers, T. T., & McClelland, J. (2004). *Semantic cognition: A parallel distributed processing approach*. Cambridge, Massachusetts: MIT Press.
- Rogers, T. T., & McClelland, J. L. (2008). Précis of semantic cognition: A parallel distributed processing approach. *Behavioral and Brain Sciences*, 31(06), 689-714.
- Rogers, T. T., & McClelland, J. L. (2014). Parallel Distributed Processing at 25: Further Explorations in the Microstructure of Cognition. *Cognitive Science*.
- Rogers, T. T., & Patterson, K. (2007). Object categorization: Reversals and explanations of the basic-level advantage. *Journal of Experimental Psychology: General*, 136(3), 451-469.
- Rogers, T. T., & Plaut, D. C. (2002). Connectionist perspectives on category-specific deficits. *Category-specificity in brain and mind*, 251-290.
- Rohde, D. L. T. (1999). LENS: the light, efficient network simulator. Pittsburgh, PA: Carnegie Mellon University, Department of Computer Science.
- Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1988). *Learning representations by back-propagating errors*: MIT Press, Cambridge, MA, USA.
- Scholz, J., Klein, M. C., Behrens, T. E. J., & Johansen-Berg, H. (2009). Training induces changes in white-matter architecture. *Nat Neurosci*, 12(11), 1370-1371.
- Schwartz, M. F., Buxbaum, L. J., Montgomery, M. W., Fitzpatrick-DeSalme, E., Hart, T., Ferraro, M., et al. (1998). Naturalistic action production following right hemisphere stroke. *Neuropsychologia*, 37(1), 51-66.
- Seltzer, B., & Pandya, D. N. (1994). Parietal, temporal, and occipital projections to cortex of the superior temporal sulcus in the rhesus monkey: A retrograde tracer study. *The Journal of Comparative Neurology*, 343(3), 445-463.
- Tang, Y.-Y., Lu, Q., Geng, X., Stein, E. A., Yang, Y., & Posner, M. I. (2010). Short-term meditation

## Category specificity in posterior fusiform

34

- induces white matter changes in the anterior cingulate. *Proceedings of the National Academy of Sciences*, 107(35), 15649-15652.
- Tyler, L. K., Moss, H. E., Durrant-Peatfield, M. R., & Levy, J. P. (2000). Conceptual Structure and the Structure of Concepts: A Distributed Account of Category-Specific Deficits. *Brain and Language*, 75(2), 195-231.
- Ueno, T., Saito, S., Rogers, Timothy T., & Lambon Ralph, Matthew A. (2011). Lichtheim 2: Synthesizing Aphasia and the Neural Basis of Language in a Neurocomputational Model of the Dual Dorsal-Ventral Language Pathways. *Neuron*, 72(2), 385-396.
- Wakana, S., Jiang, H., Nagae-Poetscher, L. M., van Zijl, P. C. M., & Mori, S. (2004). Fiber Tract-based Atlas of Human White Matter Anatomy1. *Radiology*, 230(1), 77-87.
- Warrington, E. K., & Shallice, T. (1984). Category specific semantic impairments. *Brain*, 107, 829-854.
- Yildirim, I., & Jacobs, R. A. (2013). Transfer of object category knowledge across visual and haptic modalities: Experimental and computational studies. *Cognition*, 126(2), 135-148.
- Zevin, J. D., & Seidenberg, M. S. (2006). Simulating consistency effects and individual differences in nonword naming: A comparison of current models. *Journal of Memory and Language*, 54(2), 145-160.
- Zhong, Y.-M., & Rockland, K. S. (2003). Inferior Parietal Lobule Projections to Anterior Inferotemporal Cortex (Area TE) in Macaque Monkey. *Cerebral Cortex*, 13(5), 527-540.

Table 1. Descriptive statistical data (Mean (SD)) for visual, praxic, verbal representations for animal and tool items in training environment

	Visual (n=30)	Praxic (n=10)	Verbal (n=10)	Name (n=48)	Total (n=98)
Animal	7.58 (1.38)	0.00 (0.00)	5.04 (1.16)	1.00 (0.00)	13.63 (1.66)
Tool	4.17 (1.20)	4.38 (1.38)	4.21 (0.83)	1.00 (0.00)	13.75 (2.36)

Category specificity in posterior fusiform

36

Table 2. Descriptive and statistical data for the simulations of category-specific activation in neuroimaging data

		Hub-only model		Fully-connected model		Anatomically-constrained model	
		$F(1,14) = 7.395^*$		$F(1,14) = 49.640^{***}$		$F(1,14) = 72.211^{***}$	
Visual Viewing (Sighted)	Cat × ROIs interaction	Lateral Mean (SD)	Medial Mean (SD)	Lateral Mean (SD)	Medial Mean (SD)	Lateral Mean (SD)	Medial Mean (SD)
	Animal	.275 (.077)	.377 (.087)	.497 (.101)	.227 (.079)	.459 (.101)	.201 (.069)
	Artifact	.315 (.106)	.266 (.062)	.315 (.111)	.444 (.066)	.324 (.083)	.442 (.067)
	<i>t-value</i>	- .946	3.572**	3.578**	-7.793***	3.997**	-8.388***
	Cat × ROIs interaction	$F(1,14) = 8.804^{**}$		$F(1,14) = 48.711^{***}$		$F(1,14) = 65.730^{***}$	
Name Comprehension (Sighted)		Lateral Mean (SD)	Medial Mean (SD)	Lateral Mean (SD)	Medial Mean (SD)	Lateral Mean (SD)	Medial Mean (SD)
	Animal	.273 (.078)	.381 (.088)	.497 (.103)	.997 (.077)	.463 (.101)	.206 (.071)
	Artifact	.381 (.097)	.261 (.066)	.314 (.108)	.972 (.065)	.321 (.080)	.431 (.068)
	<i>t-value</i>	-1.063	3.679**	3.625**	-7.687***	4.281**	-8.030***
	Cat × ROIs interaction	$F(1,14) = .384$		$F(1,14) = 3.286$		$F(1,14) = 39.547^{***}$	
Name Comprehension (Blind)		Lateral Mean (SD)	Medial Mean (SD)	Lateral Mean (SD)	Medial Mean (SD)	Lateral Mean (SD)	Medial Mean (SD)
	Animal	.017 (.012)	.020 (.014)	.023 (.012)	.126 (.094)	.199 (.062)	.127 (.090)
	Artifact	.016 (.012)	.016 (.010)	.023 (.012)	.147 (.108)	.203 (.061)	.308 (.085)
	<i>t-value</i>	.619	1.146	-0.163	-1.786	-.314	-8.154***

Note:  $^*p < .05$ ,  $^{**}p < .01$ ,  $^{***}p < .001$ .

Category specificity in posterior fusiform

37

Table 3 Averaged accuracy data for the simulations of degenerated praxic knowledge in semantic dementia

Familiar tool use		Level of severity (proportion of hub units lesioned)						Overall accuracy	
		5%	10%	15%	20%	25%	30%	Lesion	Control
Tool use (item)	hub-only	0.51	0.22	0.20	0.09	0.11	0.07	0.20	1
	full-conn	0.56	0.39	0.31	0.25	0.19	0.16	0.31	1
	anat-cons	0.84	0.75	0.50	0.58	0.46	0.47	0.60	1
Affordanced praxic features	hub-only	0.89	0.81	0.74	0.66	0.71	0.64	0.74	1
	full-conn	0.92	0.90	0.86	0.84	0.82	0.82	0.86	1
	anat-cons	0.98	0.97	0.92	0.93	0.91	0.92	0.94	1
Unaffordanced praxic features	hub-only	0.80	0.61	0.63	0.43	0.56	0.45	0.58	1
	full-conn	0.79	0.71	0.66	0.61	0.61	0.63	0.67	1
	anat-cons	0.93	0.91	0.78	0.77	0.72	0.74	0.81	1
Mechanical puzzle		Level of severity (proportion of hub units lesioned)						Overall accuracy	
		5%	10%	15%	20%	25%	30%	Lesion	Control
Novel tool selection	hub-only	0.73	0.65	0.57	0.49	0.53	0.45	0.57	0.82
	full-conn	0.81	0.77	0.75	0.70	0.67	0.66	0.73	0.81
	anat-cons	0.82	0.82	0.78	0.79	0.73	0.78	0.79	0.81
Novel tool use	hub-only	0.80	0.73	0.72	0.60	0.76	0.65	0.71	0.79
	full-conn	0.85	0.84	0.81	0.83	0.79	0.80	0.82	0.89
	anat-cons	0.83	0.85	0.83	0.85	0.82	0.80	0.83	0.84

Note: hub-only for the hub only model; full-conn for the fully connected model; anat-cons for the anatomically-constrained model.



## Figure Captions

Figure 1. Illustrations for the “hub” model of semantic processing. (A) A hypothetical cortical semantic network for object recognition, use, and naming tasks. This simplified network involves anterior temporal lobe (ATL), lateral parietal cortex (LP), posterior fusiform gyrus (pFG) and the perisylvian language system (mostly in superior temporal gyrus, STG). Solid lines delineate identified anatomical pathways (Binney et al., 2012). The dark-blue dashed lines and red dotted line delineate two hypotheses about how praxic representations interact with the network. (B) Model architectures for three hypotheses about network connectivity. The colored ovals are hidden layers which are analogous to critical brain regions illustrated in Figure 1A and used to simulate brain activation patterns. The dark-blue dashed lines depict the lateral/medial gradient in connectivity effectiveness between visual and praxic representations. This structure differentiated the hub-only from fully-connected and anatomically-constrained models. The red dotted line illustrates the direct connection between praxic representations and the ATL hub, included in hub-only and fully-connected models.

Figure 2. Mean activation in lateral and medial ROIs of the visual layer for three models. (A-C) Hub only models. (D-F) Fully connected models. (G-I) Anatomically constrained models. Top: Sighted variants performing picture viewing task (inputs from visual modality); Middle: Sighted variants performing name comprehension task (inputs from verbal modality); Bottom: Blind variants performing name comprehension task (inputs from verbal modality). Error bars are standard errors.

Figure 3. Contrast-weighted t-values of animal vs. tool from the lateral to the medial aspect in model simulations and neuroimaging data. Positive t-values indicate greater activation for animals (red) and negative t-values indicate greater activation for tools (blue). Numbers between panels indicate proportion of shared variance ( $r^2$ ) between datasets ( $*p < .05$ ,  $**p < .01$ ). (A-C) Anatomically constrained model; (D-F) empirical data generated by averaging the t-values reported in Mahon et al. (2009) across left and right hemispheres; (G-I) fully connected model. Top: Visual viewing (picture comprehension) in sighted individuals/models; Middle: word comprehension in sighted individuals/models; Bottom: word comprehension in blind individuals/models.

Figure 4. Effect of lesioning hub units on the activation of praxis in sighted models. (A) Accuracy on familiar tool use as a function of the severity of semantic deficit measured by word-picture matching (WPM) in patients with SD (data from Hodges et al., 2000, Figure 3). (B-D) Overall accuracy generating correct praxis for familiar items as a function of the proportion of deleted hub units in the hub-only model (B), the fully connected model (C), and the anatomically constrained model (D).

Figure 5. The effect of lesioning hub units on the activation of affordanced (gray) or unaffordanced (black) praxic properties from visual inputs for familiar tool items. Linear regression lines are also shown in the corresponding colors. (A) Patient data from Bozeat et al., 2002; (B) the hub only model; (C) the fully connected model; (D) the anatomically constrained model.

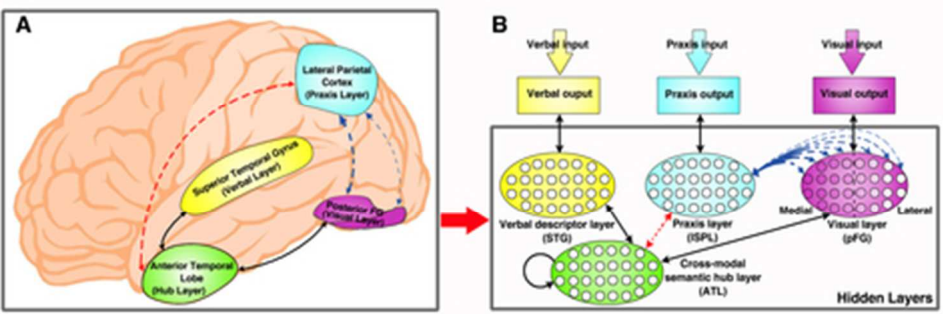
Figure 6. Correlation between disease severity and retrieval of affordanced versus unaffordanced praxic features. Regression lines for each property type are shown with  $r^2$  and  $p$  values in the top-right. (A) Patient data replotted from original data of Bozeat et al., (2002); (B) the hub only model; (C) the fully connected model; (D) the anatomically constrained model.

Figure 7. The effect of lesioning hub units on accuracy of the mechanical puzzle task in patients and sighted models. The overall accuracy ( $M \pm SEM$ ) on the mechanical puzzle task in tool selection (A) and tool use (B) aspects from Bozeat et al. (2002), Hodges et al. (2000) and three models. Patient

data from Bozeat et al. were replotted from original data and those from Hodges et al. were replotted from Figure 1B (Hodges et al., 2000).

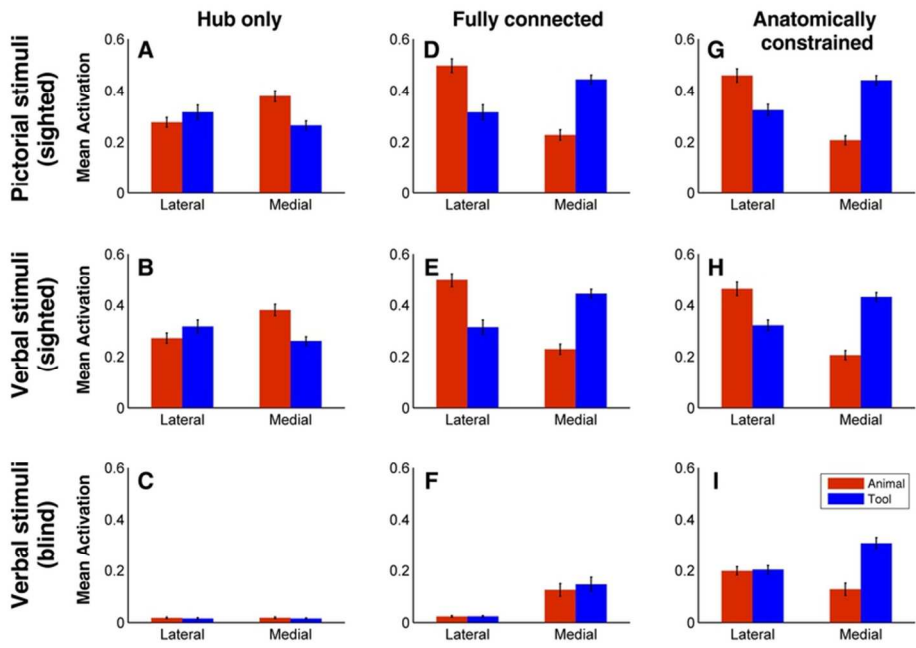
For Review Only





40x13mm (300 x 300 DPI)

or Review Only



79x51mm (300 x 300 DPI)

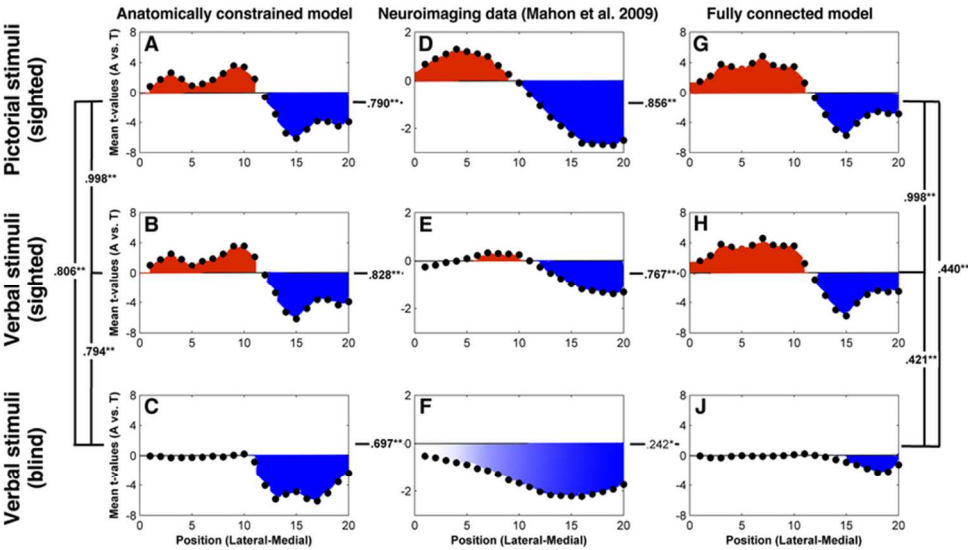


Figure 3. Contrast-weighted t-values of animal vs. tool from the lateral to the medial aspect in model simulations and neuroimaging data. Positive t-values indicate greater activation for animals (red) and negative t-values indicate greater activation for tools (blue). Numbers between panels indicate proportion of shared variance ( $r^2$ ) between datasets (\* $p < .05$ , \*\*  $p < .01$ ). (A-C) Anatomically constrained model; (D-F) empirical data generated by averaging the t-values reported in Mahon et al. (2009) across left and right hemispheres; (G-I) fully connected model. Top: Visual viewing (picture comprehension) in sighted individuals/models; Middle: word comprehension in sighted individuals/models; Bottom: word comprehension in blind individuals/models.  
77x44mm (300 x 300 DPI)

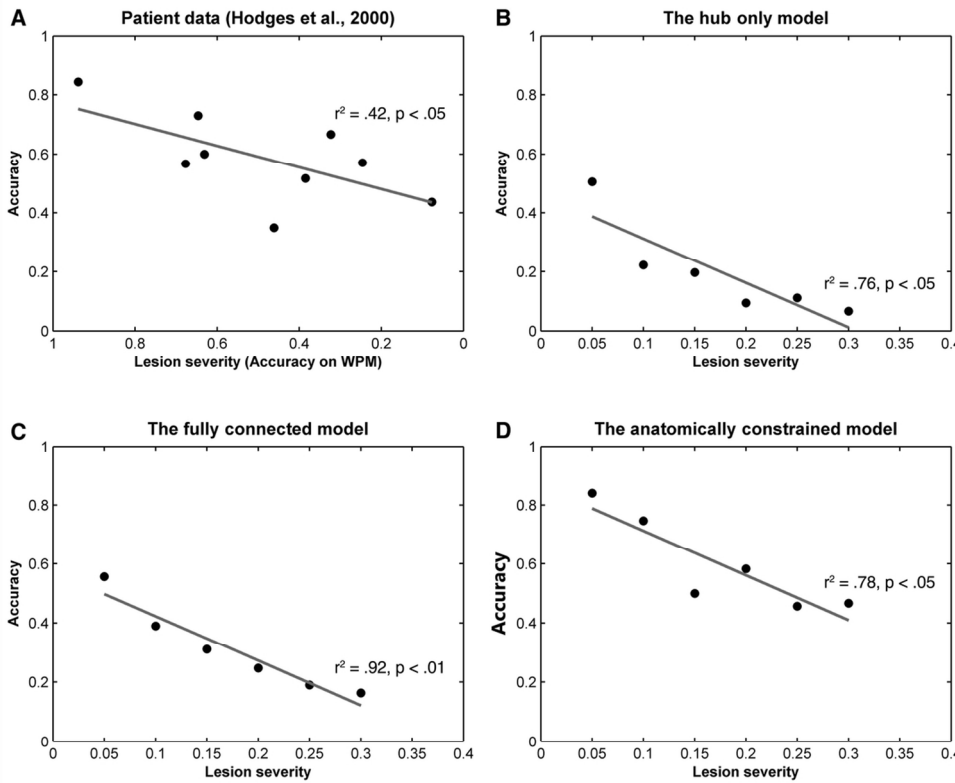


Figure 4. Effect of lesioning hub units on the activation of praxis in sighted models. (A) Accuracy on familiar tool use as a function of the severity of semantic deficit measured by word-picture matching (WPM) in patients with SD (data from Hodges et al., 2000, Figure 3). (B-D) Overall accuracy generating correct praxis for familiar items as a function of the proportion of deleted hub units in the hub-only model (B), the fully connected model (C), and the anatomically constrained model (D).

109x95mm (300 x 300 DPI)

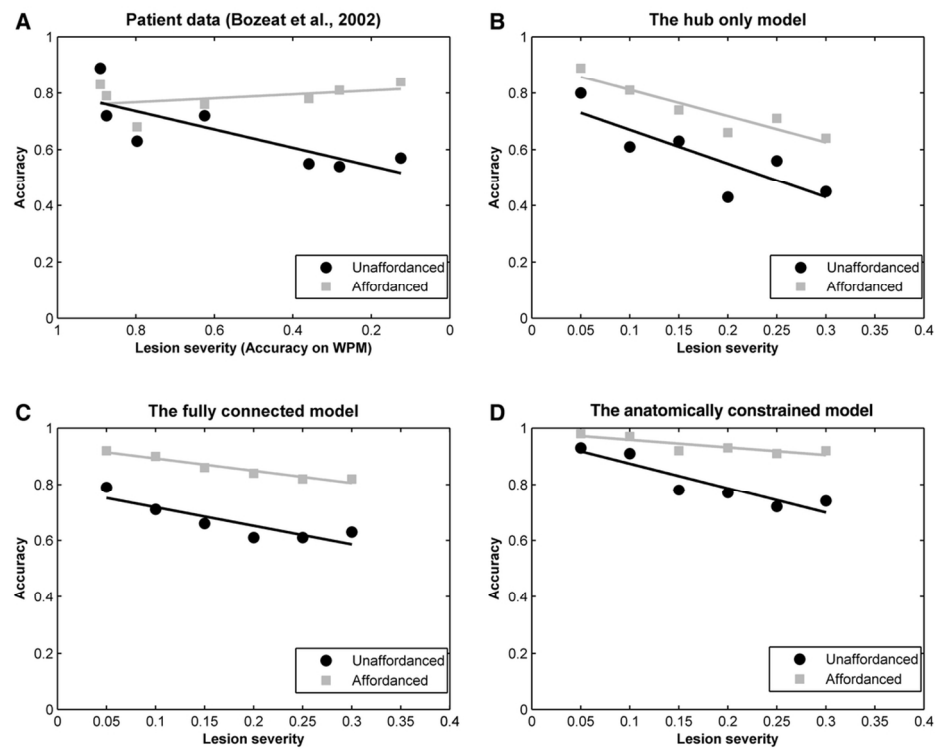


Figure 5. The effect of lesioning hub units on the activation of afforded (gray) or unaffordanced (black) praxic properties from visual inputs for familiar tool items. Linear regression lines are also shown in the corresponding colors. (A) Patient data from Bozeat et al., 2002; (B) the hub only model; (C) the fully connected model; (D) the anatomically constrained model.

105x86mm (300 x 300 DPI)

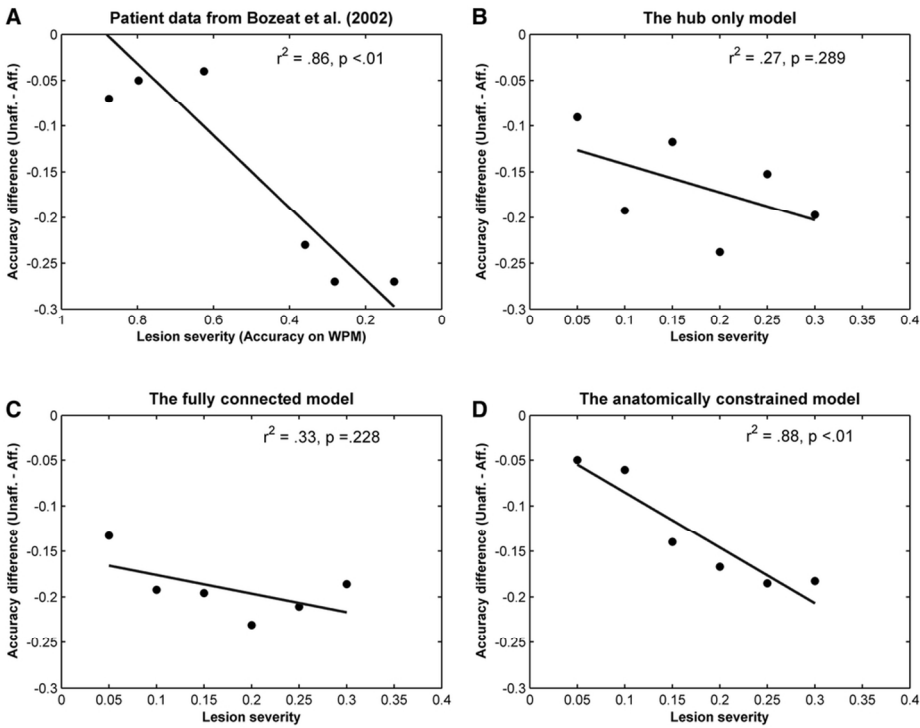


Figure 6. Correlation between disease severity and retrieval of afforded versus unaffordanced praxic features. Regression lines for each property type are shown with  $r^2$  and p values in the top-right. (A) Patient data replotted from original data of Bozeat et al., (2002); (B) the hub only model; (C) the fully connected model; (D) the anatomically constrained model.

95x71mm (300 x 300 DPI)

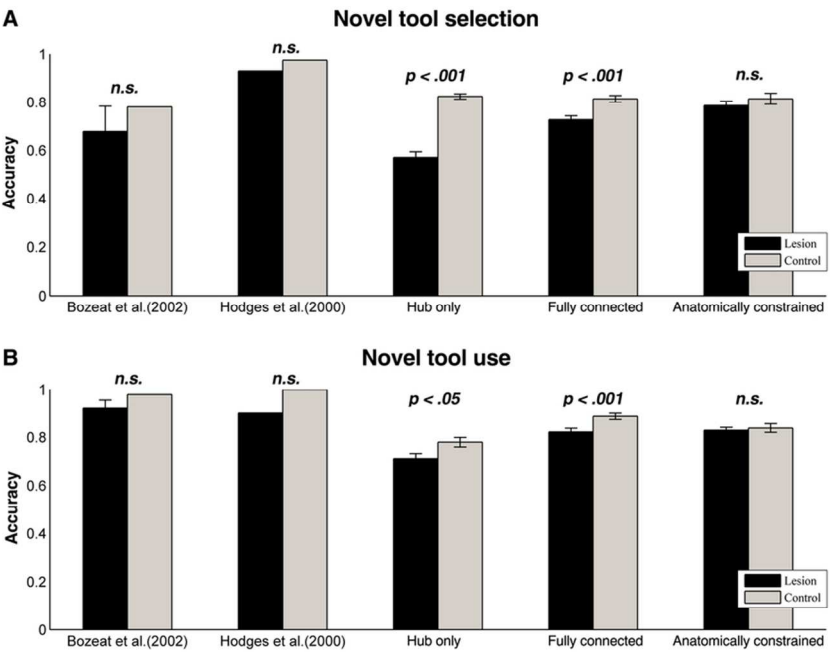


Figure 7. The effect of lesioning hub units on accuracy of the mechanical puzzle task in patients and sighted models. The overall accuracy (M±SEM) on the mechanical puzzle task in tool selection (A) and tool use (B) aspects from Bozeat et al. (2002), Hodges et al. (2000) and three models. Patient data from Bozeat et al. were replotted from original data and those from Hodges et al. were replotted from Figure 1B (Hodges et al., 2000).  
89x63mm (300 x 300 DPI)