Towards a distributed account of conceptual knowledge

Lorraine K. Tyler and Helen E. Moss

How is conceptual knowledge organized and represented? Are domains (such as living things) and categories (such as tools, fruit) represented explicitly or can domain and category structure emerge out of a distributed system? Taken at face value, evidence from brain-damaged patients and neuroimaging studies suggests that conceptual knowledge is explicitly structured in independent content-based stores. However, recent analyses of the fine-grained details of semantic impairments, combined with research using connectionist modelling, suggest a different picture – one in which concepts are represented as patterns of activation over multiple semantic properties within a unitary distributed system. Within this context, category-specific deficits emerge as a result of differences in the structure and content of concepts rather than from explicit divisions of conceptual knowledge in separate stores.

A fundamental human faculty is our ability to form DOMAINS and CATEGORIES (see Glossary) of knowledge, to partition objects in the world into meaningful sets such as living things, man-made objects, animals and tools. During the process of acquiring a new concept, we learn not only its meaning but also the domain (and category) to which it belongs. One of the most important distinctions that people seem to represent is between the broad domains of living and non-living things. This is a distinction that is observed in very young infants, in functional neuroimaging studies that show selective activation in cortical regions associated with concepts in different domains, and in brain-damaged patients with deficits restricted to a single domain (e.g. living things, non-living things) or category (e.g. animals, fruits, tools) of knowledge. Such impairments are typically referred to as 'category-specific semantic deficits', even though an entire domain of knowledge may be compromised rather than just a single category. We might assume that because domains of knowledge can be selectively impaired following brain damage and activated in imaging studies they must be explicitly and separately represented in the neural substrate. Whether or not the conceptual system is indeed organized in this way – as a set of distinct stores of knowledge or in a more distributed system – is a key topic in cognitive neuroscience because it relates to the fundamental issue of how cognitive systems in general are structured at both the functional and neural levels.

In fact, category-specific deficits have often been interpreted as showing that the first-order organizing principle in the conceptual system is *not* domain or category of knowledge but type of semantic property – that sensory properties (primarily visual) and non-sensory functional properties are represented in functionally and neuroanatomically separate stores[1]. Concepts in the living domain are more reliant on sensory properties whereas artifacts depend on functional properties, and therefore, these domains will be disproportionately affected by damage to the sensory and functional systems respectively. However, the 'sensory–functional view' has been challenged and two very different alternative approaches have recently been developed: (1) that the conceptual system is partitioned into neuroanatomically distinct content-specific stores[2,3]; (2) that category and domain structure are not explicitly represented but rather are an emergent property of the structure and content of semantic representations[4,5].

In this article we consider these different theoretical accounts in relation to behavioural data, connectionist modelling, lesion data and neuroimaging studies, and propose that conceptual structure accounts, which focus on the content and structure of concepts, provide a promising theoretical framework for understanding both the functional and neural organization of conceptual knowledge.

Domain/category-specific deficits

Patients with CATEGORY-SPECIFIC SEMANTIC DEFICITS show poorer performance for items in the impaired domain/category on a range of semantic tasks, including picture naming, word-picture matching, and generation of definitions. Various category-specific semantic dissociations have been reported including deficits for concrete words compared to abstract words[6,7] and vice versa[8], deficits for body parts[9,10], and colour names[11]. Our major focus here is, however, the contrast between the domains of living/non-living things, because this is the most widely studied dissociation. Typically, the impaired category is living things[12–20], or a specific category within the living-things domain, such as fruit and vegetables[21,22], or animals[2]. Only a handful of patients have been reported with the reverse pattern of degraded knowledge specifically for man-made objects[1,16,23–25]. Deficits for living things are most commonly associated with Herpes Simplex Encephalitis (HSE), and have often been linked to bilateral antero-medial and inferior temporal lobe lesions[26,27]. Deficits for artifacts are claimed to be

Lorraine K. Tyler*
Helen E. Moss
Dept of Experimental Psychology, University of Cambridge, Cambridge, UK CB2 3EB.
*e-mail: lktyler@ csl.psychol.cam.ac.uk

associated with extensive left-lateralized frontoparietal lesions[26].

Warrington, Shallice and McCarthy[1,20] initially suggested that dissociations between living and non-living things might arise from the differential contribution of sensory and functional semantic features in distinguishing among concepts within the two domains. They claimed that objects (e.g. *hammer* and *chisel*) are more reliably differentiated in terms of their function than their sensory properties, whereas for living things (e.g. *cabbage* and *cauliflower*) the reverse holds. On this account the conceptual system is fractionated into topographically distinct sub-stores, each of which can be independently affected by focal brain damage.

Reports of patients whose deficits cut across strict domain boundaries appear to support the sensory–functional account. Thus patients who have deficits for living things may also be impaired for certain non-living categories such as food and musical instruments[19,20]. The explanation is that concepts in the impaired non-living categories rely on sensory properties in the same way as living things and are therefore affected by damage to the sensory system. In addition, some patients with living-things deficits seem to have a greater impairment for perceptual than functional properties[12,14,19,28], as would be predicted on this account. Finally, neuroimaging studies have reported selective activation for animals in cortical regions involved in visual analysis, and activation for tools in neural regions which are also activated in action word retrieval[30].

Limitations of a sensory–functional account
However, there are several important problems with the sensory–functional account. The claimed associations and dissociations among categories have not always been observed; some patients with deficits for living things are not impaired on the supposedly similarly weighted categories of musical instruments[14], or food[15], and there are patients who show equal impairments for sensory and functional properties[2,17,31–33]. The few patients with artifact deficits have rarely been tested in enough detail to determine whether there is a greater deficit for functional properties as predicted by the sensory–functional account, but in two cases where relevant analyses are reported, the results show either no difference between sensory and functional properties[23], or an effect in the reverse direction[32]. Finally, the basic premise on which the sensory–functional hypothesis is predicated has also been questioned on the grounds that the contribution of functional information to the concepts of living things has been underestimated as the result of an overly narrow definition of what counts as a functional property. For example, animals have important biological functions, such as running, breathing, eating and so on, that are central to our conceptual knowledge of this domain (see Tyler and Moss[34] for a discussion) but which tend to be overlooked in analyses of their properties.

Category and domain-specific accounts
One response to these problems has been to argue that conceptual domain or category, rather than type of property (sensory/functional), is, after all, the first order organizing principle at both the functional and neuroanatomical level. Caramazza and Shelton[2,3] have argued this most forcefully, claiming that the conceptual system is partitioned into neuroanatomically distinct content-specific domains as a function of their role in the evolution of the human brain. Such domains include animals, plant life, conspecifics and possibly artifacts. In spite of the *prima facie* evidence in favour of domain-specific representation, caution is needed in accepting this strong claim for neural segregation. Although the lesion data show some broad correlations between bilateral antero-medial and inferior temporal lobe damage in patients and living-things deficits[26], there are certainly exceptions[2,16]. For example, Caramazza and Shelton report patient E.W., who has a selective deficit for animals in the context of a large area of encephalomalacia within the left posterior frontal and parietal lobes[2]. Living-things deficits have also been reported in patients with Alzheimer's dementia, where cortical damage is arguably patchy and widespread[4,35]. Also, voxel-based morphometric identification of the lesions in patients with HSE show considerable asymmetry in the extent to which anterior temporal cortex is compromised in different patients[36] with varying degrees of living-things deficits.

Evidence from neuroimaging
Neuroimaging studies, looking for evidence of neural specialization corresponding to categories or domains

Table 1. Domain-specific activations

Natural kinds minus artifacts	Ref.[a]	Artifacts minus natural kinds	Ref.
Right inferior parietal lobe (BA 40)	55	Left temporo-occipital junction (BA 37)	55
Bilateral anteromedial temporal lobe	55	Left posterior middle temporal gyrus	55
Left medial occipital gyrus (BA 18)	56	Left posterior middle gyrus (BA 21)	56
Left inferior temporal gyrus (BA 20)	56	Left inferior gyrus (BA 20)	56
Left fourth inferior temporal gyrus	56	Left lateral inferior frontal gyrus (BA 4/6)	30
Left calcarine sulcus	30	Left middle temporal gyrus (BA 21)	30
Left lingual gyrus (BA 18)	38	Right supramarginal gyrus	30
Left fusiform (BA 37)	38	Left anterior cingulate (BA 32)	30
Bilateral anterior temporal (BA 21/38)	37	Left inferior frontal cortex (BA 45)	38
Right posterior middle temporal cortex (BA 37/21)	37	Left posterior temporal cortex (BA 37/19)	37
		Left medial extrastriate cortex (BA 18)	37

[a]Tasks in the various studies were as follows. Ref. 55: subjects generated natural kinds or artifacts when presented with a category label, or words in response to an initial letter. Ref. 56: naming black and white pictures of animals and tools; the baseline task was to judge the orientation of unfamiliar faces. Ref. 30: silent naming of black and white line drawings of animals or tools, compared with passive viewing of non-objects. Ref. 38: same/different judgments to pairs of pictures (tools or animals); control conditions were visual texture discrimination and shape discrimination (meaningless shapes). Ref. 37: silent picture naming and word-picture matching (pictures were either coloured or black and white); the baseline task was false-font matching.

of knowledge, do not clearly support either the domain-specific or sensory–functional accounts. These studies typically test healthy subjects and use a variety of different tasks (silent naming, word-picture matching, category fluency) and materials (pictures, silhouettes, words). Although regions of activation specific to living or non-living things have been identified, these regions are not consistent across studies (see Table 1). For example, Moore and Price[37] reported activation in bilateral anterior temporal cortex and right posterior temporal cortex for living things, Martin *et al.*[30] found only a small area in the left calcarine sulcus that was differentially activated for living things, and Perani *et al.*[38] found peak activation in the left fusiform and lingual gyrus. Although these studies found additional processing in visual cortex for pictures of living things, this might have been because of their greater visual or structural complexity and so was unrelated to category/domain structure *per se*[37]. A similar pattern of inconsistent activations has also been found for non-living things[30,38].

The most striking aspect of the neuroimaging data is the extent to which living and non-living concepts activate common regions with only small and inconsistent differences between domains. These differences might be due to insufficient matching for factors such as frequency, familiarity, imageability and visual complexity. In addition, many studies report statistical values uncorrected for multiple comparisons and thus are liable to false positives. In a series of PET and fMRI studies, where we match items across categories and domains on the crucial variables of frequency, letter length, and visual complexity, we found no differences between categories or domains[39]. Using a lexical decision task (in which subjects decided whether a letter string formed a word) and a semantic categorization task (in which subjects saw three cue words presented sequentially and made a speeded decision about whether a fourth word belonged to the same

category), we found that conceptual knowledge activates a large network, primarily in the LH, involving the inferior and middle temporal gyri and the temporal pole, and no evidence for regional specialization as a function of either category or domain (see Figs 1 and 2).

The internal structure of categories and concepts
Explaining category-specific deficits in terms of damage to distinct conceptual stores does not in itself elucidate the structure and content of the concepts *within* those stores, and therefore cannot provide an adequate explanation for the detailed pattern of semantic deficits observed in patients. Brain damage does not selectively impair a specific type of knowledge or property in an all-or-none manner. Categories within domains are not always equally impaired – for example, musical instruments are sometimes categorized with living rather than non-living things[20]. Similarly, the vulnerability of concepts within categories varies as a function of many variables, including familiarity[40], homomorphy, value to perceiver and manipulability[41,42]. Moreover, some kinds of properties are more robust to damage than others; those that are true of many items within a domain are generally better preserved than those that are more specific[2,33], and properties that are densely correlated with each other are better preserved than those that are more weakly inter-related[5,43]. The patterns of deficits across many patients suggest that the nature and structure of concepts must play a crucial role in any theoretical account.

An alternative approach attempts to specify how the structure of concepts relates to the structure of categories and domains and to model apparently selective semantic impairments in a unitary distributed system with no explicit functional or neuroanatomical boundaries according to type of concept or property[4,5,43–45]. This approach assumes that concepts are represented within a unitary
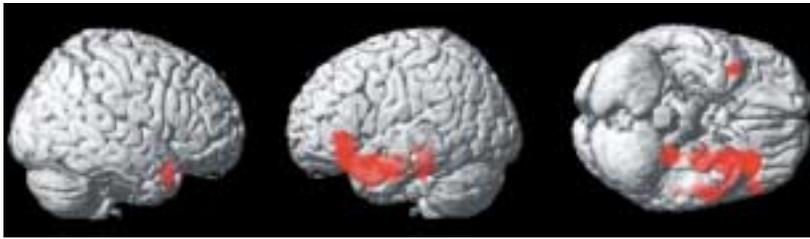
Fig. 1. Results of PET studies using written words in semantic categorization and lexical decision tasks. Brain areas commonly activated in the semantic conditions relative to their respective baselines are rendered in an MRI brain image in stereotaxic space. Red areas are activated regions significant at the cluster and/or voxel level after correcting for multiple comparisons. No voxels or clusters were significantly differentially activated for living things or non-living things; that is, there were no effects of domain (data from Ref. 39).

distributed conceptual system and that category-specific deficits emerge as a result of differences in the content and structure of concepts across categories, rather than from explicit divisions of conceptual knowledge in independent stores. These claims have recently been explored by considering concepts as patterns of activation over multiple semantic properties within a connectionist learning system[46–48].

Connectionist accounts of conceptual structure
Connectionist models of conceptual knowledge assume that domain and category structure is based on similarity, captured in the degree to which SEMANTIC PROPERTIES overlap. Thus, *lion* and *canary* belong to the same domain (living things) because they share domain-relevant properties (e.g. *legs*, *move*, *eyes*); however, they fall into different categories (*animals* versus *birds*) because some properties are shared only by concepts within one category (e.g. the wings of a bird) and not by members of the entire domain[4,5,43,48,49]. Several models have demonstrated that simple overlap of features leads to identifiable clusters in semantic space corresponding to different categories and/or domains[49,50]. This effect of 'lumpy semantic space' was predicted by Caramazza *et al.*[45] in their Organized Unitary Conceptual Hypothesis (OUCH), although this was not an implemented model. Although sharing the same basic principles, connectionist models differ in terms of specific architectures, whether the vector set is based on hypothetical or real properties, and, most

importantly, the contents and structure of the conceptual representations in the model, specifically the ways in which relationships between the features within a concept are captured.

There are two main classes of model: (1) Concepts are represented by activation over small, intuitively plausible, experimenter-generated feature sets, with no claims about the psychological validity of the representations. These models assume little structure in terms of the relations among the features in a concept[49,50]. (2) Feature structure is emphasized as the major determinant of higher-level structure. These accounts incorporate a set of theoretically and empirically derived claims about the structure of concepts across the living and non-living domains[4,5,43,44,48]. The key idea is that concepts in different categories and domains have different internal structures, in terms of the proportion of features of different kinds (e.g. perceptual versus functional) and in the CORRELATIONS among features, the relative DISTINCTIVENESS or sharedness of those features over members of a category and crucially, the many patterns of interaction among these variables. The details of one such model, the conceptual structure account[5,44], are described in Box 1.

The effects of brain damage
The conceptual structure account claims that concepts will be preserved, following brain damage, to the extent that they have numerous intercorrelated shared properties that support knowledge of the category or domain as a whole, as well as strong correlations among those more distinctive properties that are necessary for accurate identification and discrimination among similar members of a category. Overall, living things and artifacts differ along these dimensions. For living things, distinctive properties should be vulnerable to damage because they are weakly correlated with other properties of the concept, whereas shared properties, being numerous and densely intercorrelated, should be well-preserved. For artifacts, the shared–distinctive dissociation should be less marked, because distinctive properties are protected by form–function correlations and shared properties are fewer and less inter-correlated. These claims predict an interaction between distinctiveness and domain, with a greater disadvantage for distinctive compared to shared properties for living things than artifacts. Different tasks should show different degrees of dissociation, depending on the extent to which knowledge of distinctive information is required. Data supporting these predictions is reported in Moss *et al.*[33,51]

We also predict variation across categories within domains as a function of the structure of the concepts within each category. For example, according to our property norms, vehicles are like living things in that they have more numerous and
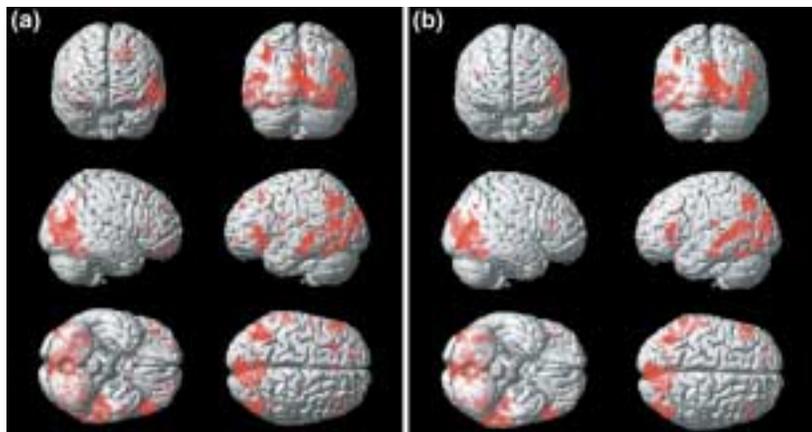
Fig. 2. Results of an fMRI study using pictures of living and non-living things in a semantic categorization task showing the neural regions significantly activated in response to (a) artifacts minus baseline, and (b) living-things minus baseline. All activations are masked with the appropriate contrasts and corrected for multiple comparisons. There were no areas that were significantly differentially activated for living things or non-living things. Data from Ref. 57.

Box 1. The conceptual structure account

This account combines theoretical insights and data from neuropsychology, modelling, developmental and experimental psychology with the aim of providing an integrated framework in which to model normal and disordered conceptual systems[a]. We claim that living things (most typically animals) have many shared properties (e.g. all mammals *breathe*, *have eyes*, *can see*, *eat*) that co-occur frequently and therefore are strongly correlated[b,c]. Living things also have distinctive properties that distinguish one category member from another (*having stripes* versus *having spots*) although these tend to be weakly

TRENDS in Cognitive Sciences

Fig. I. Model architecture. The number of units in each layer is indicated, and arrows show full connectivity between layers.

correlated with other properties and so are vulnerable to damage. (Note though, that this does *not* imply that highly distinctive members of a category will be especially vulnerable. In fact, highly distinctive members, such as *elephants* within the animal category, have many distinctive properties, and this fact increases the probability that at least some of them will be preserved in the face of random damage.) By contrast, artifacts have fewer properties, which tend to be more distinctive than those of living things. The conceptual structure account also incorporates the claim that specific perceptual properties become correlated with specific functions[d,e]. The nature of these form–function correlations distinguishes between living things and artifacts: artifacts have distinctive forms consistently associated with their functions (e.g. *blade–cut*) whereas for living things individual variations in form tend not to be functionally significant (e.g. *a lion's mane*)[a,b] . Even so, living things (like artifacts) do have form–function correlations but these involve shared properties (e.g. *eyes–see; legs–move*). We refer to these as biological functions[f,g]. A small-scale computational model based on these claims demonstrated how our assumptions can be captured in a unitary distributed conceptual system and how 'lesions' to such a model can simulate category-specific semantic deficits[a,f].

Table I. Global properties of the property norm set and the model vectors

	Property norms	Model vectors
Number of concepts	93	96
Highly distinctive features	78%	78%
Sparsity[a]	3.7%	4.6%

[a]Defined as the average proportion of features turned on for each vector.

We subsequently tested the validity of our theoretical claims in a property generation norm study. This confirmed that:
- Living things have (i) more, (ii) more correlated, and (iii) less distinctive properties, than do artifacts.
- Categories within domains differ; for example, vehicles are less typical of the artifact domain in having more properties overall and a higher ratio of shared to distinctive properties than tools.
- A higher proportion of correlations are distinctive for artifacts than for living things.

We used these property norms to develop a new computational model (see Fig. I), with concepts represented as vectors instantiating the statistics about conceptual structure derived from the property norms (Tables I and II)[h]. (See Refs i and j for similar approaches.) Vectors were presented to the auto-encoder network until they all could be recognized (indicated by the network reconstructing the input
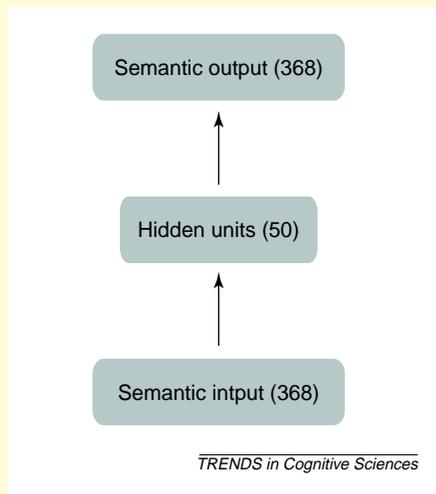
---

highly correlated properties than tools, and these properties are, on average, less distinctive (H.E. Moss *et al.*, unpublished data). This predicts that patients who have problems with animals should also have problems with vehicles in some situations, a pattern we have observed in patients (see Box 2) and which has been reported in at least one other patient in the literature[16] . Our model also predicts that fruit and vegetables will tend to be particularly vulnerable to damage, since they have relatively few distinctive properties and these are only weakly correlated. This is confirmed by the greater impairment on fruit than animals for several patients in the literature[13,17] although there are exceptions[14]. Other selective deficits for fruit/vegetables versus animals have been reported, but many of these are restricted to an anomia and so do not necessarily implicate the semantic system[21]. Similarly, highly specific deficits are sometimes reported for other categories such as body parts, but

it is debatable whether most of these are central semantic deficits rather than problems with lexical access or naming[10].

Organization of semantic space

None of the major models of conceptual knowledge (domain-specific, sensory–functional and conceptual structure) can currently account for all of the neuropsychological data. The strength of connectionist models is that they have the potential to account for seemingly complex patterns of impaired and preserved features which result in category-specific deficits, although they are not yet sufficiently well-formulated to account for all of the data. For example, our current model predicts that artifact deficits will only be seen when damage to the semantic system is particularly severe[5,23,44], and thus has difficulty accounting for patients who have greater difficulty with artifacts but in the context of a mild impairment[16,24].

Table II. Differences across domains: property norms and model vectors

	Property norms		Model vectors	
	Living things	Artifacts	Living things	Artifacts
Mean no. properties/concept	17.7	11.3	20	14
Mean distinctiveness of properties	0.64	0.73	0.22	0.32
No. of shared properties[b] per concept	13.7	7.5	15	6

[b]Defined as a property shared by three or more concepts; otherwise the property is distinctive.

pattern on the output layer); the squared error for each unit was <0.01. The model was then 'lesioned' to simulate brain damage: connections between layers were randomly removed in 10% increments. As predicted by the theory, distinctive properties of living things were more vulnerable to damage than the distinctive properties of artifacts, whereas the shared properties of living things were better preserved than in the case of artifacts, owing to the greater number of shared correlated properties (Fig. II).
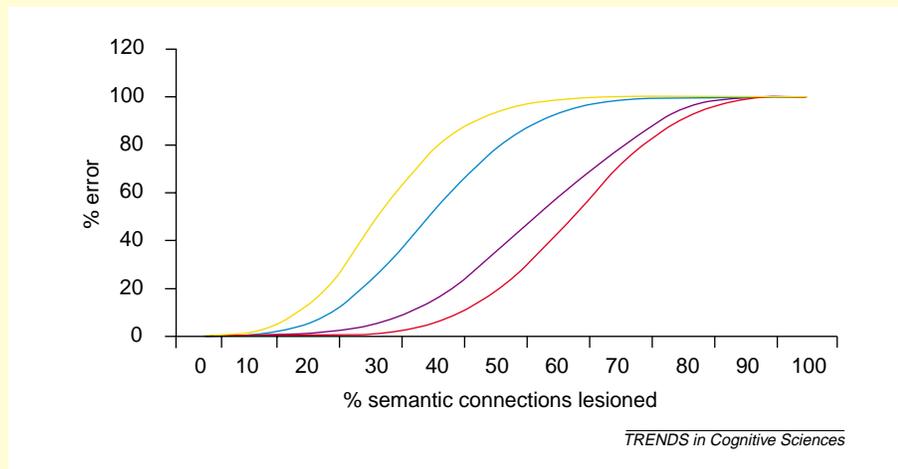

Fig. II. The results of 'lesioning' the model to simulate brain damage. As predicted, distinctive properties of living things (yellow) were more vulnerable to damage than the distinctive properties of artifacts (blue), whereas the shared properties of living things (red) were better preserved than those of artifacts (purple), owing to the greater number of shared correlated properties.

References

a Tyler, L.K. *et al.* (2000) Conceptual structure and the structure of concepts: A distributed account of category-specific deficits. *Brain Lang.* 75, 195–231

b Keil, F. (1986) The acquisition of living thing and artifact terms. In *Language Learning and Concept Acquisition: Foundational Issues* (Demoupoulous, W. and Marras, A., eds), pp. 133–153, Ablex

c Malt, B.C. and Smith, E. (1984) Correlated properties in natural categories. *J. Verbal Learn. Verbal Behav.* 23, 250–269

d Madole, K. *et al.* (1993) Developmental changes in infants' attention to function and form-function correlations. *Cognit. Dev.* 8, 189–209

e De Renzi, E. and Lucchelli, F. (1994) Are semantic systems separately represented in the brain? The case of living category impairment. *Cortex* 30, 3–25

f Durrant-Peatfield, M. *et al.* (1997) The distinctiveness of form and function in category structure: a connectionist model. In *Proc. 19th Annu. Conf. Cognit. Sci. Soc.* (Shafto, M.G. and Langley, P., eds), pp. 193–198, Erlbaum

g Tyler, L.K. and Moss, H.E. (1997) Functional properties of word meaning: studies of normal and brain-damaged patients. *Cognit. Neuropsychol.* 14, 511–545

h Greer, M. *et al.* The emergence of semantic categories from distributed featural representations. In *Proc. 23rd Annu. Conf. Cognit. Sci. Soc.*, Erlbaum (in press)

i McRae, K. *et al.* (1997) On the nature and scope of featural representations of word meaning. *J. Exp. Psychol. Gen.* 126, 99–130

j Devlin, J. *et al.* (1998) Category-specific semantic deficits in focal and widespread brain damage: a computational account. *J. Cogn. Neurosci.* 10, 77–94

In spite of their limitations, each of these theories provides insights into aspects of the neuropsychological data, suggesting that an account which integrates the properties of each will be most successful. In our conceptual structure model, we make the assumption that conceptual information is randomly distributed without any category/domain organization. This is likely to be an oversimplification. Given the overlap and structure in the properties of concepts (Box 1), it is clear that even a completely distributed connectionist network will develop its own organization such that semantic space will be 'lumpy'. That is, regions of semantic space will develop where similar concepts are represented close together by virtue of the fact that they share many features which are highly correlated. This will not generate discrete all-or-none categories of concepts, as overlap of features and similarity of structure are probabilistic variables, with individual concepts varying in their similarity to other members of the category and therefore in their distance from the centre of the semantic cluster (see Tranel, Damasio and colleagues for a related account[41,42,52,53]). Moreover, those categories with fewer and less densely inter-correlated properties (such as tools or weapons) will develop less well-defined regions in semantic space than those for which there is a larger pool of shared, correlated information (such as animals). In some cases, a property might be activated by both an artifact and a living thing (e.g. '*used for racing*' is true of racing cars and greyhounds). Moreover, it is possible that some categories might deviate from the typical structure for their domain and so exhibit non-typical characteristics (e.g. a specific artifact category might have few correlated distinctive features). Thus, although the system might organize itself in ways that reflect category and domain structure, there will be no discrete independent stores corresponding to different categories of

Box 2. Conceptual structure: evidence from patients with category-specific deficits

The results of some recent studies of herpes simplex encephalitis patients with category-specific deficits reveal the detailed pattern of loss and preservation of semantic properties across categories and domains[a–d] (Figs I–III). These data are consistent with the predictions of the conceptual structure account but are not readily accommodated by other accounts.

The sensory–functional account predicts that patients with living-things deficits will show poorer performance on sensory (perceptual) properties than functional properties, especially for living things, but the patients do not show this effect (Fig. II). The domain-specific account does not make any predictions about the pattern of preservation of properties within domains, unless complemented by an additional account of the structure of concepts. Moreover, it does not readily capture the data from the word–picture judgment task (Fig. III), which indicate that there is no clear cut-off between an impaired and a preserved category or domain, but rather a graded pattern of results that could result from the interaction of the distinctiveness of properties required to do the task and the conceptual structure of the categories probed.

References

a Moss, H.E. *et al.* (1998) 'Two eyes of a see-through': impaired and intact knowledge in a case of selective deficit for living things. *Neurocase* 4, 291–310

b Moss, H.E. *et al.* The emergence of category-specific deficit in a distributed semantic system. In *Category-Specificity in Mind and Brain* (Forde, E. and Humphreys, G., eds), Psychology Press (in press)

c Bunn. E. *et al.* (1998) Category-specific deficits: the role of familiarity re-examined. *Neuropsychology* 12, 367–379

d Tyler, L.K. and Moss, H.E. (1997) Functional properties of word meaning: studies of normal and brain-damaged patients. *Cognit. Neuropsychol.* 14, 511–545
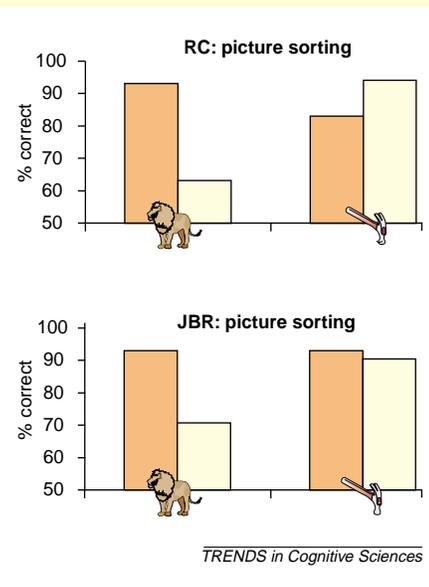

Fig. I. Picture sorting. Two HSE patients (R.C. and J.B.R.) with a well-documented deficit for living-things sorted colour pictures of living and non-living things according to two criteria: (1) category level – animal/bird or tool/vehicle (orange bars); this level of sorting depends on shared properties. (2) property level (pale yellow bars); sorting according to specific properties, which requires knowledge of distinctive information. At the category level there was no deficit for living things, which according to our theory is because of the better preservation of shared properties. The living-things deficit emerged at the property level, because of the loss of distinctive properties of living things relative to artifacts. Adapted from Ref. a.
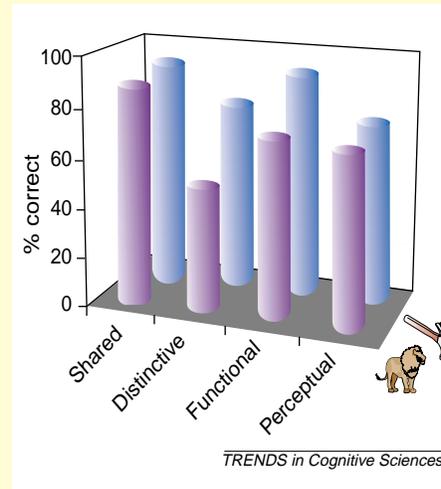

Fig. II. Property verification task. Patients indicated 'true' or 'false' to a set of property statements such as '*Tigers have stripes*' or '*Spades are used for digging*'. Statements were equally divided between true and false, perceptual and functional properties, and shared versus distinctive properties (from property generation norms). Results are shown for matched sets of animals and tools. As predicted by the conceptual structure account, patients were more accurate on shared than distinctive properties, with the poorest performance for the distinctive properties of animals. However, no patient showed the disadvantage for perceptual relative to functional properties of living things, which would be predicted by the sensory–functional account. Adapted from data in Ref. b.
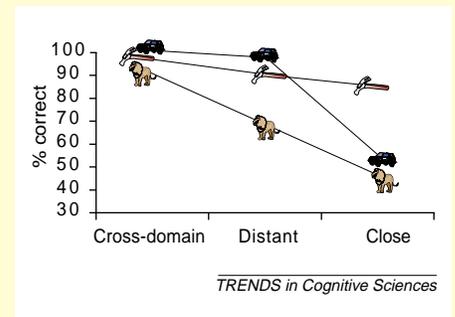

Fig. III. Word–picture judgment. Patients heard a spoken word (e.g. *tiger*), which was immediately followed by a colour picture on a computer screen, and were asked to indicate whether the word and picture referred to the same or to different objects. On the 'different' trials the picture was either a close category co-ordinate (e.g. *lion*), a distant category co-ordinate (e.g. *mouse*) or an object from another domain (e.g. *lorry*). Close and distant pairs were selected after extensive pre-testing with control subjects to determine semantic and visual similarity. Results for a group of four HSE patients, including R.C. and J.B.R., are plotted. Patients made few errors in the cross-domain condition where they can decide on the basis of shared properties. They make more errors as the word and picture refer to increasingly similar concepts, and therefore, increasingly distinctive properties are required to differentiate the concepts. There is a clear dissociation between living things and tools, especially in the close condition where the most distinctive properties are involved. Note, however, that the pattern for vehicles mirrors that of living things rather than tools in the close condition. This is consistent with the results of our property generation norms, which indicated that in some respects vehicles are like living things, in that they have more numerous and highly correlated properties than do tools, and these properties are on average, less distinctive (see Box 1).

knowledge, but rather graded, overlapping regions in semantic space (Fig. 3).

This type of functional organization could be represented in the neural system as the greater involvement of specific cortical regions for certain categories/domains, such that differential impairments will arise as a result of local effects, over and above the general patterns of robustness/vulnerability that have been identified by the conceptual structure account. This differs from Caramazza and Shelton's domain-specific account[2,3] in that these clusters are not independent neural systems that have developed in response to evolutionary pressures. Therefore, even focal brain damage will rarely produce all or none deficits for individual categories. If we assume that the organization of conceptual space at the functional level is reflected at the neural level, we would expect
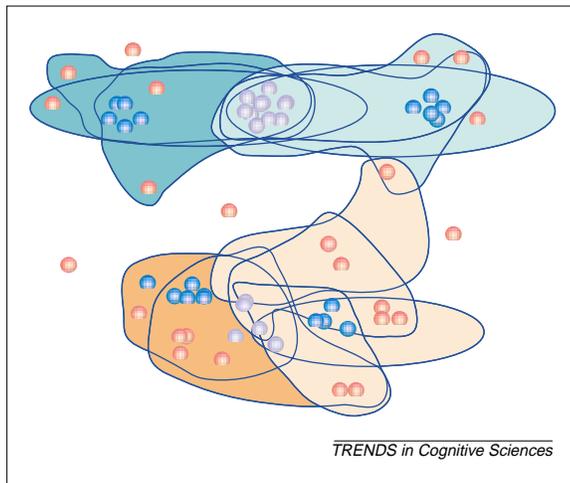
Fig. 3. Proposed clustering of correlated features, and the differences in structure for concepts in the living and non-living domains, as predicted by the conceptual structure model. Each concept is represented as a pattern of activation over a set of features. Living-things concepts (green) have many highly intercorrelated features (represented by coloured circles) shared by all members of the domain and many intercorrelated properties shared by all members of a category, such as birds or mammals (represented by light green versus dark green concepts). Concepts also have some distinctive features, but these do not tend to be highly correlated with each other. Artifacts (orange) have fewer, less densely intercorrelated properties at either the domain or category level. Therefore categories within the domain (e.g. tools, weapons, vehicles) form less well-defined clusters. However, distinctive properties tend to occur in small highly intercorrelated groups; that is, the presence of one property predicts the presence of another within the concept. In this way, domains and categories form 'lumps' within semantic space, but there is no clear cut-off between them.

brain damage to produce graded effects; there may be a certain category or domain of knowledge that is affected more severely than others, but the impact of the damage will stretch beyond category boundaries to affect other 'neighbouring' concepts to various degrees. Moreover, the pattern of impairment will vary as a function of the processing requirements of the specific task at hand (e.g. whether distinctive or shared features are emphasized[33]) and perhaps also by the goal-relevant processing biases associated with different categories[54].

<div style="background:#fffde7">

Questions for future research

- Can existing theories of conceptual knowledge go beyond the representation of concrete nouns (*carrots*, *cats*, *corkscrews*) to other word categories (e.g. abstract nouns, verbs, adjectives)? Are concrete nouns a special case?
- Connectionist accounts of concept representation have incorporated a relatively small number of factors: feature overlap, sparsity, correlation and distinctiveness. Can we identify other important determinants of conceptual structure and model them successfully in a connectionist framework?
- Can feature-based theories alone account for conceptual knowledge? Is it necessary to embed representations of individual concepts within a framework of theories about the world? If so, what is the relationship between conceptual and theoretical knowledge?
- Because most patients with 'living-things' deficits have antero-medial temporal lobe damage, why are these areas not reliably activated in neuroimaging studies of category specificity in normal subjects?

</div>

Conclusions

Understanding the functional and neural architecture of the conceptual system remains a huge, but important, challenge. However, distributed, feature-based accounts of conceptual knowledge provide a promising way forward in terms of explaining the complex patterns of behavioural deficits that arise following brain damage. Moreover, this general approach provides a theoretical context within which integrated accounts of the functional and neural properties of conceptual knowledge can be developed. It remains for future research to determine whether these promises bear fruit.

References

1 Warrington, E.K. and McCarthy, R. (1987) Categories of knowledge: further fractionations and an attempted integration. *Brain* 110, 1273–1296
2 Caramazza, A. and Shelton, J.R. (1998) Domain-specific knowledge systems in the brain: the animate–inanimate distinction. *J. Cogn. Neurosci.* 10, 1–35
3 Shelton, J.R and Caramazza, A. (1999) Deficits in lexical and semantic processing: implications for models of normal language. *Psychonomic Bull. Rev.* 6, 5–27
4 Gonnerman, L. *et al.* (1997) Double dissociation of semantic categories in Alzheimer's Disease. *Brain Lang.* 57, 254–279
5 Tyler, L.K. *et al.* (2000) Conceptual structure and the structure of concepts. *Brain Lang.* 75, 195–231
6 Breedin, S.D. *et al.* (1994) Reversal of the concreteness effect in a patient with semantic dementia. *Cognit. Neuropsychol.* 11, 617–660

7 Warrington, E.K. (1981) Concrete word dyslexia. *Br. J. Psychol.* 72, 175–196
8 Tyler, L.K. *et al.* (1995) Abstract word deficits in aphasia: evidence from semantic priming. *Neuropsychology* 9, 354–363
9 Dennis, M. (1976) Dissociated naming and locating of body parts after left anterior temporal lobe resection: an experimental case study. *Brain Lang.* 3, 147–163
10 Suzuki, K. *et al.* (1997) Category-specific comprehension deficit restricted to body parts. *Neurocase* 3, 193–200
11 DeVreese, L. (1988) Category-specific versus modality-specific aphasia for colours: a review of the pioneer case studies. *Int. J. Neurosci.* 43, 195–206
12 Basso, A. *et al.* (1988) Progressive language impairment without dementia: a case with isolated category specific semantic impairment. *J. Neurol. Neurosurg. Psychiatry* 51, 1201–1207

13 Bunn. E. *et al.* (1998) Category-specific deficits: the role of familiarity re-examined. *Neuropsychology* 12, 3, 367–379
14 De Renzi, E. and Lucchelli, F. (1994) Are semantic systems separately represented in the brain? The case of living category impairment. *Cortex* 30, 3–25
15 Hart, J. and Gordon, B. (1992) Neural subsystems for knowledge. *Nature* 359, 60–64
16 Hillis, A.E. and Caramazza, A. (1991) Category-specific naming and comprehension impairment: a double dissociation. *Brain Lang.* 114, 2081–2094
17 Laiacona, M. *et al.* (1997) Semantic category dissociations: a longitudinal study of two cases. *Cortex* 33, 441–461
18 Moss, H.E. *et al.* (1997) When leopards lose their spots: knowledge of visual properties in category-specific deficits for living things. *Cognit. Neuropsychol.* 14, 901–950

19 Sartori, G. and Job, R. (1988) The oyster with four legs: a neuropsychological study on the interaction of visual and semantic information. *Cognit. Neuropsychol.* 5, 105–132

20 Warrington, E.K. and Shallice, T. (1984) Category specific semantic impairments. *Brain* 107, 829–854

21 Hart, J. *et al.* (1985) Category-specific naming deficit following cerebral infarct. *Nature* 316, 439–440

22 Sheridan, J. and Humphreys, J.W. (1993) A verbal semantic category-specific recognition deficit. *Cognit. Neuropsychol.* 10, 143–184

23 Moss, H.E. and Tyler, L.K. (2000) A progressive category-specific deficit for non-living things. *Neuropsychologia* 38, 60–82

24 Sacchett, C. and Humphreys, G. (1992) Calling a squirrel a squirrel but a canoe a wigwam: a category-specific deficit for artifactual objects and body parts. *Cognit. Neuropsychol.* 73–86

25 Warrington, E.K. and McCarthy, R. (1983) Category specific access dysphasia. *Brain* 106, 859–878

26 Gainotti, G. *et al.* (1995) Neuroanatomical correlates of category-specific impairments: a critical survey. *Memory* 3/4, 247–264

27 Pietrini, V. *et al.* (1988) Recovery from herpes simplex encephalitis: Selective impairment of specific semantic categories with neuroradiological correlation. *J. Neurol. Neurosurg. Psychiatry* 51, 1284–1293

28 Silveri, M.C. and Gainotti, G. (1988) Interaction between vision and language in category-specific semantic impairment. *Cognit. Neuropsychol,* 5, 677–709

29 Caramazza, A. (2000) The organization of conceptual knowledge in the brain. In *The New Cognitive Neurosciences* (2nd edn) (Gazzaniga, M.S., ed.) pp. 1037–1046, MIT Press

30 Martin, A. *et al.* (1996) Neural correlates of category-specific knowledge. *Nature* 379, 649–652

31 Funnell, E. and de Mornay Davies, P. (1996) JBR: a reassessment of concept familiarity and a category-specific disorder for living things. *Neurocase* 2, 461–474

32 Lambon Ralph, M.A. *et al.* (1998) Are living and non-living category-specific deficits causally linked to impaired perceptual or associative knowledge? Evidence from a category-specific double dissociation. *Neurocase* 4, 311–338

33 Moss, H.E. *et al.* (1998) 'Two eyes of a see-through': impaired and intact knowledge in a case of selective deficit for living things. *Neurocase* 4, 291–310

34 Tyler, L.K. and Moss, H.E. (1997) Functional properties of word meaning: studies of normal and brain-damaged patients. *Cognit. Neuropsychol.* 14, 511–545

35 Silveri, M.C. *et al.* (1991) Dissociation between knowledge of living and nonliving things in dementia of the Alzheimer type. *Neurology* 41, 545–546

36 Gitelman, D. *et al.* Voxel-based morphometry of Herpes Simplex Encephalitis. *NeuroImage* (in press)

37 Moore, C.J. and Price, C.J. (1999) A functional neuroimaging study of the variables that generate category-specific object processing differences. *Brain* 122, 943–962

38 Perani, D. *et al.* (1995) Different neural systems for the recognition of animals and man-made tools. *NeuroReport* 6, 1637–1641

39 Devlin, J.T. *et al.* (2000) Susceptibility induced loss of signal: comparing PET and fMRI on a semantic task. *NeuroImage* 11, 589–600

40 Funnell, E. and Sheridan, J. (1992) Categories of knowledge: unfamiliar aspects of living and non-living things. *Cognit. Neuropsychol.* 9, 135–153

41 Tranel, D. *et al.* (1997) Explaining category related effects in the retrieval of conceptual and lexical knowledge for concrete entities: operationalization and analysis of factors. *Neuropsychologia* 35, 1329–1339

42 Tranel, D. *et al.* (1997) A neural basis for the retrieval of conceptual knowledge. *Neuropsychologia* 35, 1319–1327

43 Devlin, J.T. *et al.* (1998) Category-specific semantic deficits in focal and widespread brain damage: a computational account. *J. Cogn. Neurosci.* 10, 77–94

44 Durrant-Peatfield, M. *et al.* (1997) The distinctiveness of form and function in category structure: a connectionist model. In *Proceedings of the Nineteenth Annual Conference of the Cognitive Science Society* (Shafto, M.G. and Langley, P., eds), pp. 193–198, Erlbaum

45 Caramazza A. *et al.* (1990) The multiple semantics hypothesis: multiple confusions? *Cognit. Neuropsychol.* 7, 161–189

46 Hinton, G.E and Shallice, T. (1991) Lesioning an attractor network: investigations of acquired dyslexia. *Psychol. Rev.* 98, 74–95

47 Masson, M.E.J. (1995) A distributed memory model of semantic priming. *J. Exp. Psychol. Learn.* 21, 3–23

48 McRae, K. *et al.* (1997) On the nature and scope of featural representations of word meaning. *J. Exp. Psychol. Gen.* 126, 99–130

49 Small, S. *et al.* (1995) Distributed representations of semantic knowledge in the brain. *Brain* 118, 441–453

50 Zorzi, M. *et al.* (1999) Category-specific deficits in a self-organizing model of the lexical-semantic system. In *Connectionist Models in Cognitive Neuroscience* (Heincke, D. *et al.*, eds), pp. 137–148, Springer-Verlag

51 Moss, H.E. *et al.* The emergence of category specific deficits in a distributed semantic system. In *Category-Specificity in Brain and Mind* (Forde, E. and Humphreys, G.W., eds), Psychology Press (in press)

52 Damasio, A.R. *et al.* (1990) Category-related recognition deficits as a clue to the neural substrates of knowledge. *Trends Neurosci.* 13, 95–98

53 Damasio, A.R. and Damsio, H. (1993) Cortical systems underlying knowledge retrieval: evidence from human lesion studies. In *Exploring Brain Functions: Models in Neuroscience* (Poggio, T.A. and Glaser, G.A., eds), pp. 233–248, John Wiley & Sons

54 Gauthier, I. (2000) What constrains the organization of the ventral temporal cortex. *Trends Cognit. Sci.* 4, 1–2

55 Mummery, C.J. *et al.* (1996) Generating 'tiger' as an animal name or a word beginning with T: differences in brain activation. *Proc. R. Soc. London Ser. B* 263, 989–995

56 Damasio, H. *et al.* (1996) A neural basis for lexical retrieval. *Nature* 380, 499–505

57 Tyler, L.K. and Tavares, P. Is there neural specialization for pictures of living and non-living things? *NeuroImage* (in press)

Book reviews

Books and other types of publication – including software, CD-ROMs, films, art exhibitions, etc. – can be considered for review in *Trends in Cognitive Sciences*. Please send any ideas or proposals for specific items to the Editor.

Publishers: send copies of books or software you would like to be reviewed to the Editor at the address below.

Readers: would you like to review a particular book or other publication? Contact the Editor via e-mail with your proposal.

Trends in Cognitive Sciences
Elsevier Science London, 84 Theobald's Road, London WC1X 8RR.
e-mail: tics@current-trends.com