

# Semantic Memory

Eiling Yee<sup>1</sup>, Michael N. Jones<sup>2</sup>, Ken McRae<sup>3</sup>

<sup>1</sup>Department of Psychological Sciences, University of Connecticut

<sup>2</sup>Department of Psychological and Brain Sciences, Indiana University

<sup>3</sup>Department of Psychology, Social Science Centre, University of Western Ontario

## Abstract

How is it that we know what a *dog* and a *tree* are, or, for that matter, what *knowledge* is? Our semantic memory consists of knowledge about the world, including concepts, facts and beliefs. This knowledge is essential for recognizing entities and objects, and for making inferences and predictions about the world. In essence, our semantic knowledge determines how we understand and interact with the world around us. In this chapter, we examine semantic memory from cognitive, sensorimotor, cognitive neuroscientific, and computational perspectives. We consider the cognitive and neural processes (and biases) that allow people to learn and represent concepts, and discuss how and where in the brain sensory and motor information may be integrated to allow for the perception of a coherent “concept”. We suggest that our understanding of semantic memory can be enriched by considering how semantic knowledge develops across the lifespan within individuals.

**Keywords:** *Semantic memory, Concepts, Distributional Semantic Models, Categories, Abstraction, Generalization, Knowledge, Grounded Cognition, Sensory-motor, Embodiment*

To appear in: In J. T. Wixted & S. Thompson-Schill (Eds), *The Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience* (4th Edition, Volume 3: Language and Thought). New York: Wiley, 2017.

## 1. What is semantic memory?

What do psychologists mean when they use the term *semantic memory*? Almost half a century ago, in 1972, Endel Tulving suggested partitioning the human long-term memory system into two distinct stores: an *episodic* store that contains memories that are linked to a particular time and place (e.g., where you parked your bicycle this morning), and a *semantic* store that contains more general knowledge about the world (e.g., what a bicycle looks like, or is used for). Tulving's proposal was widely adopted, and now many psychologists and cognitive neuroscientists consider episodic and semantic memory to be components of the declarative (or *explicit*) branch of the long-term memory system. Motor knowledge about how you actually ride a bicycle, in contrast, is generally described as a *procedural* skill that is part of another branch of long-term memory—the non-declarative, or *implicit* memory system. This system encompasses knowledge to which we do not have conscious access, but that nevertheless affects our behavior (Squire, 1987).

Early neuropsychological evidence supported the view that episodic memory is distinct from implicit memory, and is at least partially distinct from semantic memory. Amnesic individuals (i.e., people with episodic memory deficits) for example, are able to learn new procedural skills (e.g., maze solving and mirror reading, Cohen & Squire, 1980; Milner et al., 1968) and acquire some new semantic knowledge (e.g., the names of new famous people; O'Kane et al., 2004), despite having no episodic memory of having learned these things.

Although such neuropsychological evidence does suggest that episodic, semantic, and implicit memory systems are at least *partially* distinct, there is also evidence that the semantic system is not completely independent of either episodic or implicit knowledge. In fact, detailed behavioral, computational, and neuroimaging investigations suggest that semantic memory is part of an integrated memory system – a system that is grounded in the sensory, perceptual, and motor systems, and that is distributed across multiple brain regions. In other words, there is now good evidence (some of which we describe later in this chapter) that there is in fact no strict division between semantic memory and the sensory, motor, and episodic information from which semantic information is acquired originally.

In what follows, we discuss semantic memory from cognitive, sensorimotor, cognitive neuroscientific, and computational perspectives. Although much of the presentation focuses on semantic memory for concrete objects, along the way we will see that many of the same issues that concern concrete objects are also relevant for abstract concepts, actions, and events, as well as for the features of each of these.

## 2. What is semantic memory for?

Imagine seeing something, say a dog, for the very first time. It has various attributes, such as a particular shape, motion, smell, sound, color, feel, fur length, and collar. When you see the same animal again a week later, its fur length and smell might be different (perhaps it has recently been bathed and groomed), and it might be wearing a different collar. But its shape, motion, color and bark are the same. Thus, across your two experiences with the animal, some attributes were more constant (shape, sound, color) than others (fur length, smell and collar). By being sensitive to the frequency of the animal's various attributes across episodes of experiencing it, you might derive (or *abstract*) a representation of the concept *dog* (we use the label “dog” for convenience, but the same process would work for an animal we do not have a name for) in which the aspects that are most commonly shared across those episodes are most heavily weighted, and the aspects that those episodes share less frequently have less weight.

The regularities gleaned via the process of abstraction allow *generalization* based on this concept. Therefore, when we see another thing for the first time, to the extent that the new thing's attributes are similar to those that are included in our concept (and taking into account the weights), we can link that new thing with the *dog* concept we formed earlier. Thus, another furry, four legged barking animal would be more strongly associated with the concept *dog* than would a cat, which would in turn be more strongly associated with the concept than would a *table* or a *lemon*. Furthermore, once we have associated the new animal with our *dog* concept, properties of the dog concept that are not, at a particular moment, evident in the new animal can nevertheless be generalized (i.e., mapped on) to the new animal, allowing us to make predictions about aspects of the new animal that we have not directly perceived (e.g., how a new instance of a dog will behave when someone scratches its ears).

Thus, it is via abstraction that we can, as infants (or perhaps later in life), discover that there is a *type* of thing that is a *dog* or a *lemon*, and it is via generalization that we can apply this knowledge to new instances. Even as adults who possess more-or-less fully developed conceptual knowledge, abstraction and generalization continue to be necessary for learning about new things and for applying existing knowledge (imagine moving to the USA from Europe, where poison ivy doesn't grow, and having to abstract, from examples, which plants to avoid because they are poison ivy, and then applying that knowledge on a hike). The example above is intended merely to give a flavor of how the abstraction process might work. In fact, there may not be a single “abstraction process”, rather there may be several distinct properties of the memory system which each contribute to the aspects of

memory we refer to as “abstract”<sup>1</sup>. In the final sections of this chapter, we describe a number of computational models of semantic memory that use computational mechanisms that can support abstraction, not all of which, in fact, require forming semantic representations *per se*.

Now that we have considered what semantic memory is for, we turn to considering how it is structured and organized.

### 3. What is the structure and organization of semantic memory?

What is your semantic representation of a *dog* or a *tree* made up of, and how are these representations organized with respect to one another? In the semantic memory literature, researchers have attempted to answer these questions from both cognitive and cognitive neuroscientific perspectives. Cognitive approaches typically focus on how concepts may be internally structured, acquired, and organized with respect to one another. In contrast, cognitive neuroscientific approaches have focused primarily on how they are organized, using what is known about the structure and functional architecture of the human brain to constrain models. Although cognitive neuroscientific approaches often have little to say about how concepts are acquired and internally structured, the currently dominant class of cognitive neuroscientific models, sensorimotor-based models, are an important exception. We begin this section by briefly summarizing the most prominent cognitive models of how concepts are internally structured. We then turn to theories of the organization of semantic knowledge, with a focus on cognitive neuroscientific approaches.

#### 3.1 What is a concept made of?

Historically, the three most prominent cognitively-oriented theories of the internal structure of concepts have been classical definitional approaches, prototype theory, and exemplar theory. In what follows, we briefly introduce each. For a detailed consideration of these issues, refer to the chapter in this volume by Goldstone, Kersten, and Carvalho (2017).

---

<sup>1</sup> In this chapter we use the term *abstraction* in two related ways: 1) as explained above, to refer to the *process* of deriving information *across* instances that may not be available from any individual instance (in other work, this is sometimes called *induction* or *generalization*), and 2) as discussed in subsequent sections, to refer to different *levels of abstraction*, i.e., differences in degree of detail or precision, where the less detail there is, the greater the level of abstraction, e.g., *poodle* vs. *dog* vs. *animal*. It is easy to imagine that the way that the abstraction *process* (in the first sense) is implemented produces different *levels of abstraction* (in the second sense; see 8.1 for discussion). We reserve the term “generalization” for the related process of *applying* that abstracted knowledge to novel situations (cf. Altmann, 2017), that is, *doing something* with that knowledge. We consider these issues in further detail in Sections 6-8 (see also Barsalou, 2016, for further discussion).

According to the classical theory of concepts, which dates back at least as far as the ancient Greeks, concepts can be thought of as definitions that are built from simpler concepts (e.g., *bachelor* = *unmarried* + *man*). Thus, deciding if someone is a bachelor is a matter of checking whether they are both unmarried and a man. The definitional approach has fallen out of favor, in part because for most concepts, agreeing on a precise definition is difficult or impossible (e.g., what defines a *game*? Wittgenstein, 1953).

Prototype theory (Rosch, 1978) avoids the problem of deciding how to define concepts by positing that concepts are probabilistic, rather than having a definitional structure. That is, for each concept (e.g., *dog*), we encode a list of features (e.g., *has four legs*, *has fur*, *barks*) that are weighted by how frequently that feature has occurred in the dogs we have encountered. When we encounter a new dog, we attempt to match it to these weighted features, with the degree of match determining the likelihood that we will generalize, or correctly classify the new thing as a dog. Prototype theory has an interesting consequence. If what we think of as the most typical dog (e.g., *size*, *shape*, *color*, *fur length*, etc.) is a weighted average of all the dogs we have encountered, this average may be something we have never actually experienced (for review, see Murphy, 2002). Although in their most basic form, prototypes are unorganized lists of features (e.g., the features *has wings* and *flies* are not predictive of one another), there is evidence that we remember concepts better if the features “make sense” according to our prior knowledge (e.g., if the feature *wings* goes with *flies*, rather than with *swims*; Murphy & Allopenna, 1994). This suggests that we do attempt to organize features with respect to one another. In addition, a number of recent versions of prototype-style theories incorporate statistical aspects of featural representations, such as feature distinctiveness (Cree, McNorgan, & McRae, 2006; Randall, Moss, Rodd, Greer, & Tyler, 2004), and correlations among features (McRae, de Sa, & Seidenberg, 1997; Tyler & Moss, 2001).

But do we really need a stored list of weighted features to decide that a new thing is a member of a category? Exemplar models (Medin & Schaffer, 1978; Smith & Medin, 1981) posit that we do not. Instead, according to exemplar models, to decide if something that we encounter is, for example, a dog, we compare it to each of our previous experiences with dogs (all of which we have stored). The more similar it is to those dog exemplars, the more likely we are to decide that it is a dog. Like prototype theory, the exemplar model also predicts that classifying a never-before experienced thing as a dog may be easier than classifying an atypical dog that you have previously seen; something we have never seen can be more similar to our stored exemplars for a category than one

of those stored exemplars is to the rest (consider a novel spotted Labrador, compared to a Chihuahua)<sup>2</sup>.

There is currently no clear consensus on whether exemplar or prototype models best fit human behavior. However, in an intriguing recent article, Murphy (2016) argued that there never has existed a coherent overall exemplar theory of concepts. Murphy's arguments notwithstanding, if we consider that the summary descriptions described in the prototype model must originally develop from examples, it seems likely that remembered exemplars play an important role, especially in the development of conceptual knowledge. It may be the case, however, that as more and more examples continue to be encountered, we also store representations that reflect what is most common across all of those exemplars. That is, we may store abstractions. We consider abstraction further in subsequent sections (see also the chapter in this Handbook by Gureckis).

### 3.2 How are concepts organized?

We now turn to considering models of how conceptual knowledge is organized. We pay special attention to sensorimotor theories because of their influence on the field and because, as we will see, their architecture allows them to address not only the organization of conceptual knowledge, but also its acquisition and internal structure.

#### Hierarchical models

We learn in biology that organisms can be classified using a hierarchical system (e.g., a *tree* is a *plant*, and a *plant* is a *living thing*). Classical cognitive psychological theories describe the organization of knowledge in semantic memory in the same way—as a hierarchy that is structured according to the relations between concepts. There is some psychological evidence for hierarchical structure. For example, individuals are slower to identify that a *tree* is a *living thing* than that a *plant* is a *living thing*, a finding which has been interpreted as reflecting the greater distance to be travelled in the hierarchy to verify the statement (Collins & Quillian, 1969, but cf. Smith 1978). However, after controlling for potentially confounding factors that could have slowed verification responses for more distant properties, such as association between the terms, it

---

<sup>2</sup> Although categorization is performed differently in exemplar and prototype models, the process of conceptualization (e.g., imagining a typical dog when someone mentions dogs in general) results in the two sets of models making similar predictions. According to the exemplar model, each time we imagine a typical dog, we produce something similar to a weighted average of all of our dog exemplars, which would then provide another exemplar for our set of dogs (Hintzman, 1986). This exemplar would be experienced frequently (via all of the times we think about dogs when they are not there), resulting in a stored "best example" which would be similar to a prototype/stored summary description. It would be different from a prototype, however, in that this "best" exemplar would not have any special status. That is, like any other instance of experience, it would be a reflection of, rather than a description of, the category.

has been difficult to obtain empirical evidence that people do routinely use hierarchies to store and retrieve conceptual knowledge (Murphy et al., 2012). For these and other reasons, strict hierarchical models of semantic organization are unlikely candidates at the moment.

#### Neurally-inspired models

In early work, Warrington & McCarthy (1983) described an individual who had more difficulty identifying non-living things than living things. Soon thereafter, Warrington and Shallice (1984) described several patients exhibiting the opposite pattern: more difficulty identifying living than nonliving things. These individuals with brain injuries inspired two influential neural models of the organization of semantic knowledge: 1) domain-specific category-based models; and 2) sensory-functional theory, as well as its successor, sensorimotor-based models. Below we describe each in turn.

#### Domain-specific/category-based models

Patterns of deficits like the ones described above suggest that brain damage can differentially affect categories (or *domains*) of objects, such as living versus nonliving things. One way to account for this pattern is to posit that different categories of objects are processed in distinct brain regions. The domain-specific category-based model (Caramazza & Shelton, 1998) is based on this idea. In this model, due to evolutionary pressures, human brains developed dedicated fast and accurate neural mechanisms for recognizing a few categories that are particularly relevant for survival or reproduction (animals, plant life, conspecifics, and tools). Moreover, because efficiently recognizing each of these categories requires different mechanisms, the adaptations produced a neural organization in which the categories have distinct, non-overlapping representations. This is not to say, however, that these representations must be localized to circumscribed brain regions. A more recent extension of this model, the distributed domain-specific hypothesis, suggests that categories are distributed across the brain, albeit in distinct, non-overlapping neural regions (see Mahon & Caramazza, 2009, for a review).

#### Sensory-functional theory

Rather than interpreting their early observations of individuals with category-specific deficits as evidence that conceptual knowledge in the brain adheres to a category-based organization, Warrington and Shallice (1984) hypothesized that sensory and functional information contribute to create conceptual knowledge, and that different categories of knowledge tend to rely on these two types of information to different extents. For instance, according to sensory-functional theory, representations of living things tend to depend more on visual information than do artifacts, which depend more on information about their function. This means that even without semantic memory being categorically

organized *per se*, if access to one type of information is compromised, deficits could appear to be category-specific (Farah & McClelland, 1991; Warrington & McCarthy, 1987). Although sensory-functional theory was highly influential, the existence of patients whose deficits are not consistent with a sensory-functional division indicates that semantic knowledge cannot be captured by a simple two-way divide (see Mahon & Caramazza, 2009, for a review).

### Sensorimotor-based models

Sensorimotor-based models (also known as “grounded”, or “embodied” models) are an extension of sensory-functional theory. However, in these models, sensory information is divided into as many attributes as there are types of sensory input. According to Alan Allport, who revived such models in the late 20<sup>th</sup> century, the sensorimotor systems through which we experience the world are also used to represent meaning. “The essential idea is that the *same* neural elements that are involved in coding the sensory attributes of a (possibly unknown) object presented to eye or hand or ear also make up the elements of the auto-associated activity-patterns that represent familiar object-concepts in ‘semantic memory’” (Allport, 1985, p. 53)<sup>3</sup>. In recent decades, several models have made similar claims from perspectives that are linguistically oriented (Lakoff & Johnson, 1999), cognitively oriented (Barsalou, 1999) and cognitive neuroscientifically oriented (Damasio, 1989; Pulvermuller, 1999). These models are appealing because unlike many others, they make specific predictions about how semantic information is obtained in the first place (via sensory and motor experience), as well as how and where in the brain this information is processed (in the sensory and motor systems in which it was acquired). For a detailed discussion of this perspective, we refer the reader to the chapter in this volume by Matheson and Barsalou.

A large body of research now supports sensorimotor-based models. Numerous functional neuroimaging, neuropsychological, and neuromodulatory (e.g., TMS: transcranial magnetic stimulation) studies converge to suggest that semantic knowledge for a given object concept is built around its sensory and motor attributes (e.g., its color, shape, smell, and how people use it). These attributes are distributed across brain regions that underlie sensory and motor processing (for reviews see Gainotti, 2000; Kiefer & Pulvermuller, 2012; Noppeney, 2009; Yee et al., 2013). Consequently, at present, most semantic memory researchers would agree that conceptual representations are, at least in part, represented in sensorimotor systems. Active areas of research explore (1) How, and/or where, in the brain this sensorimotor-based information may be combined (or *bound*) into a coherent concept from a jumble of features, and

perhaps relatedly, (2) How sensorimotor-based information might be integrated to produce a representation that involves the sensory and motor systems to differing degrees depending upon the situational demands.

One method for binding features is believed to involve synchronous neural activity. That is, semantically coherent representations may be created by integrating patterns of synchronized neural activity representing different aspects of sensory and motor information (e.g., Singer & Gray, 1995; Schneider et al., 2008). Another (compatible) possibility is that there are regions of cortex that function as *hubs* in that they receive and combine input from multiple modality-specific areas. These hubs are hypothesized to not only bind together features, but to also transform their input such that they represent similarity among concepts that cannot be captured based on individual sensory or motor modalities (for review, see Patterson et al., 2007).

Although most hub models are silent regarding how they determine which inputs are integrated, an early sensorimotor-based model proposed solving this problem by integrating the concept of hubs (which, in this model, are referred to as *convergence zones*) with that of temporal synchrony (Damasio, 1989). According to the convergence zone framework, not only do convergence zones combine input from multiple areas, they also reflect synchronous activity within the areas from which they receive input. Moreover, the convergence zone framework posits that multiple levels of convergence zones (i.e., a hierarchical structure) are necessary to build up semantic representations (see Simmons & Barsalou, 2003, and McNorgan et al., 2011 for evidence consistent with multiple levels of convergence zones from brain damaged and healthy participants, respectively).

There currently is considerable evidence that there are brain regions (in particular, in the anterior temporal lobe and angular gyrus) that function as hubs or convergence zones (for reviews of neuropsychological and neuroimaging evidence for hubs, see Binder, 2016, and Patterson et al., 2007, respectively). Through integrating input from multiple areas, these regions may support abstraction and generalization. We return to the neural regions supporting these processes in Section 6.

### 3.3 Summary of the structure and organization of semantic memory

Where does this consideration of cognitive and neurally oriented models of semantic memory leave us? First, it is worth noting that it has been argued that the hub/convergence zone architecture described above could not only bind features into objects, but, in an analogous fashion, higher order convergence zones could bind objects into generalized events (e.g., *making toast*; Damasio, 1989). It also has been suggested that

<sup>3</sup> The ideas, however, go back at least 300 years to the British empiricist philosophers, John Locke and David Hume.

mechanisms that receive and integrate input about sensory and motor features are likely to extract statistical regularities across these features that correlate with taxonomic categories. In other words, a hub/convergence zone architecture could allow a category-like topography to emerge from sensorimotor input without the need to posit pre-specified divisions among categories, or a strictly hierarchical taxonomic organization (Simmons & Barsalou, 2003).

Second, both of the cognitive models that we have considered, that is, prototype and exemplar models, are, in principle, compatible with the two current models of the organization of semantic memory (domain-specific and sensorimotor models). Although domain-specific category-based models are committed to the existence of neural regions dedicated to a few, evolutionarily relevant categories, these models are silent about how, within these categories, knowledge is represented, as well as about how all other categories of knowledge might be represented. Likewise, although sensorimotor-based models posit that concepts are acquired via sensory and motor experience, and are at least partially represented in sensory and motor cortices, they are also silent about whether the categories that we form are based on a single, probabilistic representation (i.e., a prototype) or stored exemplars.

Although our current understanding of categorization behavior and the organization of the neural systems supporting semantic memory does not allow determining whether prototype or exemplar models are more plausible, it seems likely that examples play an important role in the development of conceptual knowledge. In the next section, we discuss evidence for the importance of such examples. We also consider evidence that as development proceeds, conceptual knowledge may gain the flexibility to rely more or less heavily on the sensorimotor systems (see Gureckis, 2017).

#### **4. How is semantic memory acquired?**

Clearly, to know things about the world, such as what strawberries look and smell like, that bicycles have two wheels, or that elephants have DNA, we must have learned that information during some episode(s). These experiences may be direct, such as actually encountering strawberries, or through language, such as learning about a type of fruit that we have never seen. However, does this mean that all information in semantic memory begins as information in episodic memory (i.e., as information that is linked to a specific time and place)? Or can information be incorporated into our semantic memory even if we were never able to explicitly recall the episode(s) in which we were exposed to it? Opinions on these questions differ. Some authors have argued that episodic memory is the

“gateway” to semantic memory—that is, that semantic memory must be acquired via episodic memory (see Squire & Zola, 1998 for review), although eventually, semantic information may exist independently. In contrast, Tulving (1991) argues that, “new information could be stored in semantic memory in the absence of a functioning episodic system, as presumably happens in very young children and in lower animals without episodic memory” (p. 20).

Evidence from individuals with severe episodic memory deficits can help address the question of whether a functioning episodic memory system is necessary to acquire new semantic knowledge. For instance, research with the famous amnesic patient H.M. revealed that after the surgery that led to his amnesia, he acquired some new semantic knowledge (e.g., for words that came into common use after his surgery; Gabrieli et al., 1988; cf. O’Kane et al., 2004). Two other individuals who became amnesic as adults have also exhibited some, albeit extremely limited, new semantic knowledge after the illnesses that led to their amnesia (Bayley & Squire, 2005; Bayley et al., 2008). Furthermore, individuals who have had amnesia since early childhood appear to have relatively intact semantic knowledge, despite the fact that they had little time to acquire semantic knowledge when their episodic system was intact (Bindschaedler et al., 2011; Gardiner et al., 2008; Vargha-Khadem et al., 1997). Although such evidence seems to suggest that semantic knowledge can be acquired without an intact episodic memory system, it is worth noting that semantic knowledge acquisition in these amnesic patients is not normal (e.g., it is acquired very slowly and laboriously). It is therefore possible that these patients may possess sufficient remaining episodic memory to allow for the acquisition of semantic knowledge (Squire & Zola, 1998). Another (compatible) possibility is that the acquisition of semantic memory normally makes use of the episodic system, but that other (less efficient) points of entry can be more heavily relied upon when the episodic system is damaged (or has not yet developed fully).

What might be some of these “points of entry”? As one might imagine given the prominence of sensorimotor-based models, sensory and motor information that may be acquired implicitly is an obvious candidate. And in fact, there is clear evidence from studies on the development of conceptual knowledge in young children that sensory and motor information are important for developing semantic knowledge about object concepts. Some of this evidence comes from studies examining how young children categorize and make inferences about novel objects. For example, when two-year old children were presented with a novel object and instructed to move it horizontally, they were more likely to consider it to be similar to (i.e., to categorize it with) another novel object whose long axis extended horizontally than one whose long axis

extended vertically, whereas the preference reversed for children who moved the original novel object vertically (Smith, 2005). Thus, motor information has an implicit influence on their categorization, which presumably affects their conceptual representations.

Like 2-year olds, at ages 4-5, young children also show a tendency to prioritize sensory/perceptual information when categorizing objects, as indicated by their inferences about the objects. After learning to label novel creatures as either “ziblets” or “flurps” based on a rule that did not correlate with overall perceptual appearance (e.g., ziblets have more fingers than buttons, but are not otherwise perceptually similar to one another), children were shown a new target creature, told that it had a hidden property (in this case, thick blood) and asked which of two other new creatures also had thick blood. Children consistently selected the creature that had more overall perceptual similarity to the target creature, rather than the creature that, based on the fingers-to-buttons-ratio rule, had the same label (Sloutsky et al., 2007). Thus, at 4-5 years of age, perceptual information that is derived implicitly appears to play an important role in developing semantic knowledge about object concepts, and perhaps a larger role than the explicitly learned label. Such demonstrations with young children suggest that statistical regularities in their sensory and motor input may allow them to form new object concepts, without the need to depend on explicitly learned information.

There is little controversy about the claim that for infants, sensorimotor information is essential to the acquisition of semantic knowledge about object concepts<sup>4</sup>. However, a central question about semantic memory is whether there is a developmental shift with respect to the role of sensorimotor information in object knowledge. Common sense suggests that there should be. Adults have many ways to develop semantic representations for new object concepts (e.g., through verbally described definitions, “jicama is a root vegetable”, or through analogy, “jicama has a texture similar to an apple, but it tastes kind of like a potato”). Infants, however, must rely entirely on the sensory and

---

<sup>4</sup> Debate continues, however, between nativist and empiricist accounts of conceptual development. In nativist accounts, infants’ interpretation of the sensorimotor input is constrained by innate biases or “theories”, for example, that there are core domains of knowledge like objects, animates, and number (Carey, 2009; Spelke et al., 1992). Empiricist accounts claim that there is sufficient structure in the input that concepts can be learned from perceptual experience with the aid of innate *general-purpose* cognitive biases. That is, there is no need for innate biases that pertain specifically to conceptual knowledge (for review, see Sloutsky 2010; Rakison & Lawson, 2013). Both views agree that for older children and adults, hearing common labels for objects promotes categorizing those objects as instances of the same concept. According to the nativist view, humans have an innate predisposition to assume that words refer to concepts (Markman, 1990), whereas according to the empiricist view, words initially have no more weight than other sensory or motor information, but that over development, words gain a more prominent role in categorization as children learn that labels are predictive of category membership (Sloutsky et al., 2001; Deng & Sloutsky, 2012).

motor information that is available to them when they experience the object (e.g., their own experience of the jicama’s taste, smell, texture, and feel).

Empirical evidence is also consistent with a shift. Although young children rely more on perceptual similarity than on labels when making inferences (or predictions) about unseen object properties (e.g., Sloutsky et al., 2007, described above; for a review, see Rakison & Lawson, 2013), older children and adults are more strongly influenced by category labels and by higher order regularities – that is, by regularities that may be detectable across instances of experiencing an object concept. For example, older children and adults are more inclined than younger children to group a car together with a bicycle, presumably in part because of the knowledge that both are used for transport. To explain this shift, Sloutsky (2010) proposed that a “selection-based system may subserv[e] [certain types of] category learning by shifting attention to category-relevant dimension(s) and away from irrelevant dimension(s). Such selectivity may require the involvement of brain structures associated with executive function...[and] there is evidence that many components of executive function critical for learning sparse categories exhibit late developmental onset ...” (p. 1249; cf. Rogers & McClelland, 2004 for a connectionist account of developmental changes in conceptual organization).

Hence, the evidence supports the (common sense) idea that as children develop, their semantic representations increasingly rely on abstracted information, that is, information that is grounded less in direct sensorimotor experience and that depends more on higher order statistical regularities and category labels. We return to the process of abstraction in Sections 6 and 7. However, first we consider whether semantic knowledge is ever “fully developed”. That is, we consider whether changes to, or long-term differences in, experience can result in discernable differences in semantic knowledge across adults.

## 5. Individual differences in semantic memory

Given that semantic memory is often referred to as “general world knowledge”, it seems obvious that individual differences in semantic memory should exist. After all, depending on your experiences, you may or may not know the capital of Vermont or what an ocelot looks like. It follows that cultural factors can be important sources of differences in conceptual processing (see the chapter in this volume by Yap, Ji, & Hong). Perhaps more surprisingly, there are also subtle differences in what different individuals know (or believe) about more common things, such as a trumpet, hammer, or bird, as well as differences in how we categorize them or view their similarity to one another. Moreover, if sensorimotor-based models are correct in

positing that experience determines conceptual representations, then to the extent that experiences differ, representations must differ as well. In this section, we consider individual differences in semantic knowledge.

## 5.1 Differences in categorization

On the surface, categories may appear to be fixed. Once you learn what *hammers* or *birds* are, intuition suggests that these are stable categories that neither vary across time within an individual, nor across individuals. However, behavioral evidence suggests that even the broad categories that we use to organize the world differ across individuals, as well as within an individual (across the lifespan). When asked to indicate which is more closely related, a *robin* and a *nest*, or a *robin* and a *duck*, East Asians are more likely to select the *robin* and *nest* than are Westerners (Ji et al., 2004). This selection reflects sensitivity to *thematic*, or *event-based*, relations. Thematic, in contrast with taxonomic, knowledge concerns what or how things typically “go together”, or play complementary roles in the same action or event. Individual differences can also be found within a culture. Among Western participants, some adults exhibit a consistent preference to categorize thematically, while others prefer to classify taxonomically (e.g., categorizing *robin* with *duck*; Lin & Murphy, 2001), and these preferences remain consistent across implicit and explicit measures (Mirman & Graziano, 2012). Moreover, at as young as age three, children show stable preferences for either thematic or taxonomic categorization, and these preferences can be predicted by aspects of their play and language behavior at 13 months and at 24 months, respectively (Dunham & Dunham, 1995)<sup>5</sup>. Interestingly, these preferences may change over development—there appears to be a shift during young childhood whereby younger children strongly prefer to categorize thematically, but this bias weakens as children get older (for review, see Markman, 1990). Although the reason for this shift is unknown, it has been suggested that it may be related to developing language abilities (e.g., knowing that *robins* and *ducks* are both called *birds* might encourage infants to search for commonalities between robins and ducks; Waxman & Markow, 1995).

In addition to individual, cultural, and developmental differences in the broad categories that we use to organize the world, there are also more subtle differences across individuals in how semantic knowledge is organized and represented. We next turn to some examples of such differences, starting with how

expertise with actions or objects influences an individual’s semantic knowledge.

## 5.2 Expertise-induced differences

Differences in expertise can produce behaviorally and neurally measureable differences in conceptual activation. For instance, professional musicians can identify pictures of musical instruments more quickly and accurately than can novices (non-musicians), while no such difference is found for common objects. Moreover, this pattern is reflected neurally. For professional musicians, pictures of musical instruments activate auditory association cortex and adjacent areas more so than do pictures of common objects, with no such difference being found for novices (Hoenig et al., 2011).

Analogous findings have been reported for ice hockey experts. Simply listening to sentences describing hockey actions (but not everyday actions) activates premotor regions more in hockey experts than in novices (Beilock, Lyons, Mattarella-Micke, Nusbaum & Small, 2008). Thus, the brain regions that are differentially activated for experts (in these two studies, auditory for musicians, and premotor for hockey experts) appear to be due to their specific types of experience. Moreover, the fact that these regions are active when the experts perform tasks that do not require accessing musical or action knowledge suggests that these aspects of experience have become part of their semantic representations of instruments and actions, respectively.

Within non-experts, there is also evidence of experience-based differences in conceptual knowledge. Yee, Chrysikou, Hoffman, and Thompson-Schill (2013) found that performing a sequence of hand motions that are incompatible with those used to interact with frequently manipulated objects such as pencils disrupts thinking about such objects more than it disrupts thinking about less frequently manipulated objects, such as tigers. Critically for the current point, participants’ ratings of their degree of manual experience with an object predicted the degree of interference that the manual task produced. Moreover, for non-experts, like for the experts described above, experience with an object affects its neural representation. When listening to an object’s name, activation in a brain region involved in performing object-related actions (left parietal cortex) correlates with participants’ ratings of their amount of lifetime tactile experience with an object (Oliver, Geiger, Lewandowski & Thompson-Schill, 2009).

Even just a few weeks of sensory or motor experience can produce differences in neural activity during conceptual activation. For instance, Kiefer, Sim, Liebich, Hauk, and Tanaka (2007) trained participants to classify novel objects while either pointing at them or pantomiming actions toward them. After three weeks,

---

<sup>5</sup> Specifically, children who at age 3 prefer to categorize taxonomically exhibited more pointing behavior and labeling of individual objects at 13 and 24 months, respectively. In contrast, children who prefer to categorize thematically at age 3 exhibited more relational play and use of relational terms at 13 and 24 months (Dunham & Dunham, 1995).

when participants simply made judgments on the objects' written names, EEG source analysis revealed that participants who had been trained to pantomime at the objects, but not those who had been trained to point at them, showed early activity in motor regions and later activity in occipito-parietal visual-motor regions. Analogous findings have been reported for novel objects for which olfactory experience was either present or absent (Ghio et al., 2016). These and related findings (Cross et al., 2012; Oliver, Parsons, & Thompson-Schill, 2008; Ruther et al., 2014a; Weisberg, Turenout, & Martin, 2007), suggest that years of experience are not required to impact the neural representations of concepts—shorter-term differences in sensorimotor experience can have a measurable impact as well.

### 5.3 Bodily-induced differences

Differences in sensory or motor experience can be larger and more pervasive than those described above. And as might be expected (given what has been observed for differences in expertise and in relatively short-term experience), larger, bodily-induced differences in experience, such as the absence of input in a sensory modality, can also affect semantic representation. For example, for sighted, but not for blind individuals, implicit similarity judgments about fruits and vegetables are influenced by color similarity (Connolly et al., 2007). Importantly, this is true even for blind participants who have good explicit color knowledge<sup>6</sup>.

However, the dramatic differences in sensory experience between early blind and sighted individuals do not always produce detectable differences in behavioral measures of semantic knowledge, or in its neural organization. For instance, when making judgments that in sighted individuals are thought to rely, in part, on processing visual motion (e.g., is *chiseling* a hand action with a tool?), both blind and sighted participants selectively activate a brain region that in sighted individuals receives input from an area associated with processing visual motion (left posterior middle temporal gyrus, which receives input from area V5/MT; Noppeney et al., 2003). Moreover, blind and sighted participants show category-specific (e.g., animals vs. artifacts) activation in ventral stream brain regions that, in sighted participants, are associated with visual processing (ventral temporal and ventral occipital regions; Mahon et al., 2009). Such findings may appear to be incompatible with sensorimotor-based theories, which posit that differences in visual *experience* explain

why ventral “visual” brain regions respond differently to different categories of stimuli (e.g., between animals and tools). However, given that it is well-established that neural plasticity results in reorganization of the cortex if visual input is absent (for a review, see Amedi, Merabet, Bempohl & Pascual-Leone, 2005), it may be that in blind individuals these ventral regions are sensitive to non-visual factors that correlate with hand action and with the living/non-living distinction (e.g., shape information that is acquired tactilely). Thus, the knowledge represented in these regions may still be based on *experience*, but not on *visual* experience.

A similar argument can help reconcile a sensorimotor-based account with another result that may seem to challenge it. It has been proposed that understanding an observed action (e.g., shooting a basketball) is based, at least in part, on the observer's ability to perform a “motor simulation” of the observed action (e.g., Rizzolatti et al., 2001). However, there is evidence that the absence of motor experience does not necessarily impair knowledge of actions. When presented with videos of actions, individuals born without/with severely shortened upper limbs can recognize, remember, and anticipate actions as accurately as typically developed participants (Vannuscorps & Caramazza, 2016). This finding indicates that knowledge about actions is not necessarily based on the ability to perform those actions (even though such knowledge might ordinarily be represented in that way). Nevertheless, it is perfectly compatible with the idea that action knowledge (and indeed, conceptual knowledge more generally) is experience-based. That is, for individuals *with* experience performing actions, action knowledge will be supported by that experience, but for individuals *without* experience performing actions, action knowledge will be supported by whatever experience they do have of actions, which would include experience visually perceiving them.

Importantly, it appears that the absence of a specific type of sensory or motoric information does not necessarily produce impoverished representations (or if it does, differences are subtle and difficult to detect). Instead, we suggest that there may be a sort of *conceptual compensation* (or adaptation) whereby the information that is available becomes particularly important and informative for representations, at least when the system has had time to adapt or has never developed reliance on a specific modality. Thus, just as experts appear to have different conceptual representations in their areas of expertise compared to novices, individuals with different bodies should also have different (and not necessarily more or less informative) representations. One example of how representations can differ based on long-term bodily experience comes from a highly pervasive individual difference—which hand is typically used to interact with an object. Kan and her colleagues have reported that

---

<sup>6</sup> Interestingly, blind and sighted individuals did not differ in their judgments about household items—a distinction that is consistent with evidence suggesting that color's importance in an object's representation depends upon how useful it is for recognizing the object (e.g., color is important for distinguishing *lemons* from *limes*, but not for distinguishing *toasters* from *microwave ovens*; see Tanaka & Presnell, 1999, for review).

pictures of tools (but not pictures of non-manipulable things like animals) activate left premotor cortex more in right-handers than in left-handers (Kan, Kable, Van Scoyoc, Chatterjee & Thompson-Schill, 2006; see Willems, Hagoort & Casasanto, 2010, for related findings).<sup>7</sup>

Together, the studies described in this section indicate that the organization of semantic knowledge, and its neural representation, can be affected by both large and small individual differences in sensory or motor abilities. More importantly, they show that with differences in long-term experience, representations can develop differently or adapt, such that semantic knowledge that grows out of nonidentical sets of processes can nevertheless be equally informative.

#### 5.4 Summary of individual differences in semantic memory

A fundamental prediction of sensorimotor-based theories is that object concepts include knowledge that is represented in, or processed by, the sensorimotor regions that are routinely active when that object is perceived or interacted with. This means that individuals with different experiences should have different representations and that even within an individual, as experience changes, that individual's representations should change as well. Such individual differences may seem problematic because it would mean that no two individuals' concepts of, for example, *lemon* will be exactly the same, and perhaps more disconcertingly, that your own concept of *lemon* changes, albeit probably subtly, over time. However, the data described above suggest that this is, in fact, what happens.

Much of the research described in Section 5 shows how individual differences in conceptual knowledge manifest as differences in the manner in which different brain regions are recruited, whether through differences in expertise, or through differences in sensory or motor experience. In the next section, we review more broadly the role of different brain regions in conceptual representation.

### 6. Which brain regions support semantic memory, and how?

We suggested in Section 3 that there is ample evidence from the neuroimaging and neuropsychological literatures to support sensorimotor-based models of semantic memory, which claim that the brain regions involved in perception and action also

support meaning. Several of the studies in the previous section provide converging evidence by showing that the sensory and motor areas supporting semantic knowledge appear to change as sensory and motor experience changes. However, there is also general agreement that the neural representation of semantic knowledge extends beyond activation in modality-specific regions. In a comprehensive meta-analytic review of functional neuroimaging studies, Binder et al. (2009) identified several *non* modality-specific regions that consistently have been implicated in tasks requiring semantic knowledge. These include the inferior parietal lobe (including the angular gyrus), large portions of the lateral and ventral temporal lobes, and several parts of the frontal lobe including the inferior frontal gyrus.

A complete discussion of these brain regions is beyond the scope of this chapter (see the chapter by Matheson & Barsalou for an in-depth discussion). Instead, we limit our discussion to the inferior frontal gyrus, the angular gyrus, and the anterior temporal lobe because these regions have often been characterized as “general” semantic areas. That is, they are discussed as being *heteromodal* in supporting multiple modalities of input (e.g., visual and auditory) and multiple categories of semantic knowledge (e.g., animals and artifacts). Evidence from cognitive neuroscience suggests that each of these brain regions may support semantic knowledge in different ways.

The inferior frontal gyrus, for instance, appears to support semantic processing by virtue of promoting the selection of contextually appropriate information (Thompson-Schill et al., 1997). Notably, however, the inferior frontal gyrus is thought to support general-purpose selection, meaning that although this region frequently acts in the service of semantic processing, it also supports selection in non-semantic tasks (for review, see Thompson-Schill, 2003).

The angular gyrus has often been put forth as an important heteromodal association area by virtue of its location between visual, auditory, spatial and somatosensory association areas (for discussion, see Binder, 2009). There is some evidence that this region has a particularly important role in supporting the type of thematic or event-based semantic knowledge that was introduced in Section 5.1 (e.g., that *dog* is related to *leash*; Humphries et al., 2007; Kalenine et al., 2009; Mirman & Granzino, 2012; Schwartz et al., 2011; de Zubicaray et al., 2013). However, because some studies have not linked thematic knowledge in particular to the angular gyrus (see Jackson et al., 2015), more research is needed to understand the specific conditions under which this region may support thematic knowledge or processing (Kalenine & Buxbaum, 2016).

The anterior temporal lobe (ATL) has received perhaps the most attention in the literature on the neural representation of semantic knowledge. One

<sup>7</sup> Perhaps surprisingly, abstract concepts' representations may also be influenced by differences in individual experience. Right-handers' have a tendency to associate “good” with “right” and “bad” with “left”; whereas left handers show the opposite pattern (Casasanto, 2009). Intriguingly, this pattern can be reversed when right hand dominance is compromised due to stroke, or even a brief (12 minute) laboratory-induced handicap (Casasanto & Chrysikou, 2011).

reason for this is the existence of a neurological condition called *semantic dementia*, which, early in the progression of the disease, often is characterized by relatively focal degeneration in the ATL. Individuals with semantic dementia have semantic memory deficits such as problems with naming, recognizing, and classifying objects, regardless of the category of knowledge. In contrast, other cognitive functions are relatively spared. Thus, the ATL appears to support general semantic knowledge (see Hodges & Patterson, 2007, for a review).

Functional neuroimaging studies of unimpaired individuals also implicate the ATL in general semantic processing. For instance, Tyler et al. (2004) had people name pictures using either basic level labels such as *dog* or *hammer* or superordinate level labels such as *animal* or *tool*. They found that the ATL (specifically, the perirhinal cortex, and primarily in the left hemisphere) was more active when people named pictures using basic level labels, as compared to superordinate level labels. In contrast, a region of the posterior temporal lobe (fusiform gyrus, bilaterally) was activated equally during naming at the two levels. These results are consistent with the idea that the ATL supports object categorization, and that the amount of ATL involvement is a function of how much detail is needed to perform the targeted level of categorization (see Rogers et al., 2006 and Rogers & Patterson 2007 for related findings and discussion).<sup>8</sup>

How might the ATL support semantic processing? Recordings of neural activity (via magnetoencephalography) show that during basic level naming of visually presented objects, there are more recurrent interactions between anterior and posterior (fusiform) regions of the left temporal lobe than during superordinate level naming (Clark et al., 2011). Given the posterior fusiform's involvement in processing visual features of objects, it is possible that the ATL functions as the "hub" of a network that facilitates the integration of information across features, and that such integration is necessary for categorization tasks such as determining that an image depicts a *dog* rather than a *cat* (see Martin & Chao, 2001). Furthermore, the ATL's involvement increases as a function of discrimination demands (Clark et al., 2011). Or, to use the vocabulary we used in Section 2, the ATL may support generalization by supporting the process of applying our existing semantic knowledge to the input, in this case to categorize it.

Thus, the ATL's greater involvement in basic relative to superordinate level categorization (Tyler et al., 2004; Rogers et al., 2006) may not be due to *containing* the

perceptual information required for finer-grained discrimination. Rather, the integration that occurs in this region may produce a similarity space that is organized in a way that reflects relationships among multiple features. In this space, higher-order relationships may play a more important role than perceptual information (see Rogers et al., 2004). Thus, when the goal is to match (i.e., categorize) a perceptually presented item with a region of this similarity space, the task is more difficult when attempting to match the input with narrower categories (e.g., *dogs*) than with broader categories (e.g., *living things*) because narrower categories require accessing more of that item's perceptual representation in posterior regions<sup>9</sup>.

More broadly, there is evidence to suggest that integration and abstraction may occur throughout the ventral temporal lobe (e.g., the ventral visual pathway; Tanaka 1996), producing a posterior to anterior gradient of abstraction. Specifically, more posterior regions reflect information that is based more closely on perceptual features, whereas more anterior regions reflect information that, due to integration across features, is somewhat more complex. In particular, anterior regions may reflect higher-order abstract relationships among features. Findings from several studies support this type of gradient (Chan et al., 2011; Grabowski et al., 2001; Kable et al., 2005; Hsu et al., 2011), and many others are consistent more generally with the idea that semantic knowledge is represented at different levels of abstraction (or is integrated to different degrees) in different regions (Fernandino et al., 2015).

Having different levels of abstraction may help alleviate a potential concern regarding sensorimotor-based accounts of conceptual knowledge. If concepts are sensorimotor-based, one might worry that thinking of the concept of *lemon* could cause one to hallucinate a lemon, which people generally do not. Furthermore, thinking of the concept of *kicking* could elicit a kick, which could be embarrassing or perhaps even dangerous. Yet if concepts are indeed also represented at a more abstract level than that underlying direct

<sup>8</sup> Recent work also considers whether different sub regions of the ATL, or the left versus right ATLs, may support different aspects of semantic knowledge (e.g., social knowledge, verbally-mediated semantic knowledge, non-verbal semantic knowledge; for reviews, see Gainotti, 2015; Lambon-Ralph et al., 2017).

<sup>9</sup> This characterization, of basic level categorization requiring more processing than superordinate level categorization, may seem at odds with the well-established phenomena of basic level naming typically being privileged relative to superordinate level naming (e.g., basic level names are learned earlier during development, and, under unspeeded conditions, objects are named faster at the basic, relative to the superordinate level; Rosch, 1976). However, the basic level advantage in naming is likely due to the fact that it is more common to name objects at the basic than at the superordinate level (leading to stronger object-label associations at the basic level), not that recognizing objects at the superordinate level requires more processing. Recent behavioral evidence supports the idea that superordinate level recognition does not require additional processing. Under speeded conditions, categorization at the superordinate level is privileged relative to categorization at the basic level (Rogers & Patterson, 2007; Mack & Palmeri, 2015), and information needed to categorize at the superordinate level is available earlier than information needed to categorize at the basic level (for review, see Clarke, 2015).

sensation and action, then when thinking about, for example, *kicking*, the regions that underlie the motor act of kicking do not need to be sufficiently active to produce a kick.

In the next section, we change gears somewhat to review computational models of conceptual representation that are not based on sensorimotor information. Although much of semantic knowledge comes from direct experience with objects and actions, a great deal of it does not. People learn a huge amount from spoken and written language. We have knowledge of places that we have never been, and of people that we have never met. The computational models described in Section 7 derive semantic knowledge from language input only. Despite being based solely on “abstract” input (language), they successfully capture many aspects of human semantic memory. In a sense, the success of such models can be considered a testament to the power of abstraction.

## 7. Computational models of semantic representation

So far, we have focused on cognitive and neuroscientific models of semantic memory. The former are driven by functional considerations, such as the functions that memory serves, and the functional architecture of distinct memory systems. The latter are driven by how functions relevant to memory appear to be organized across different regions of the brain. In this section, we focus on computational models of semantic representation. Such models have the potential to explain how the nature of the input, or the algorithms operating over that input, constrain the organization of semantic memory, and how the learning process impacts that organization. However, the utility of computational models in understanding semantic memory from a psychological perspective depends on the extent to which the principles governing their operation may be assumed to correspond to psychologically relevant constructs, rather than reflecting implementational choices designed to increase their ability to predict behavior irrespective of psychological plausibility. In what follows, we try to highlight the psychologically relevant constructs.

Recent work has obtained a proxy of mental semantic similarity by employing statistical processes that operate across either large databases of human-generated features (McRae, Cree, Seidenberg, & McNorgan, 2005; Vinson & Vigliocco, 2008), or across behavioral measures such as free association (Abbott, Austerweil, & Griffiths, 2015). Although these approaches have both had considerable success predicting human performance, they are limited in that neither one addresses how representations may be created in the first place (Jones, Hills, & Todd, 2015).

Many current computational models of semantic representation do address how semantic representations may develop in the first place. These *distributional semantic models* use statistical experience (i.e., regularities in the environment) to construct semantic representations. Although for practical purposes, “experience” in these models is usually restricted to a large corpus of text, the models are in principle statistical learners that can be applied to any data that can be represented to them. In the discussion that follows, we use the terms ‘word’ and ‘linguistic context’ because this is typically the input to the models, but in principle, these could be replaced with ‘object’ or ‘event’, and physical or temporal context.

While there are now many distributional semantic models in the literature, they are all based on the *distributional hypothesis*, which is that words that appear in similar linguistic contexts are likely to have related meanings (Firth, 1957; Harris, 1970). For example, *car* may frequently co-occur with *wheel*, *driver*, and *road*. As a result, the model can infer that these words are related in some way. In addition, the model can learn that *car* is similar to *truck* even if the two never directly co-occur. This happens because *car* and *truck* both occur near the same types of words. Similarly, because *car* and *net* rarely appear in the same or similar contexts, according to the distributional hypothesis, they are not likely to be related.

Due in part to the models’ successes at explaining human data, as well as the greater availability of text corpora, the field of distributional semantic models has grown enormously over the past two decades. There are now dozens of models in the literature, each with its own strengths and weaknesses. Rather than focusing on the technical differences among the various models, we instead focus on the psychological mechanisms that are used to build semantic representations. This allows us to organize distributional semantic models into five families: 1) Passive Co-occurrence, 2) Latent Abstraction, 3) Predictive, 4) Bayesian, and 5) Retrieval-Based. For technical differences among the models, we refer the reader to Riordan and Jones (2011) or Turney and Pantel (2010).

While all distributional semantic models are based on the distributional hypothesis, the families of models differ in theoretically important ways in terms of the mechanisms they posit for how distributional information is learned, and the type of information that is stored in the semantic representation. They thus differ with respect to how they model *abstraction*. These mechanisms range from simple Hebbian learning to Bayesian inference and reinforcement learning. In addition to differences in learning mechanisms, these families also differ in *when* they hypothesize that semantic abstraction takes place from episodic experience. Some models suffer from psychological implausibility in that they abstract meaning via a

learning mechanism that is applied only after all episodes have been learned, whereas others accumulate semantic information continuously as experience unfolds. In contrast, there are other models, like the exemplar models described earlier, that posit that only episodic traces are stored, and semantic abstraction is an artifact of retrieval from episodic memory. A challenge for such models is to account for evidence suggesting that episodic memories for one's entire life can be lost without losing access to semantic knowledge (see Tulving, 2002 for review). One possibility is that in such cases, the ability to explicitly retrieve episodic memories is lost, but the episodic traces themselves remain in some form. If true, these traces could allow for a type of implicit retrieval that supports semantic abstraction. For an in-depth review of the theoretical claims of different models, see Jones, Willits, and Dennis (2015). Below, we highlight only the main properties of the five families of models.

### 7.1. Passive co-occurrence models

Members of the family of passive co-occurrence models all share the characteristic that they simply count the co-occurrence frequency among words within an “attentional context”. The size of the attentional context is usually based loosely on working memory capacity, such that co-occurrences might be counted within, for example, a several-word moving window or a sentence, but rarely a full paragraph. In this sense, this family theoretically implements ideas of Hebbian learning from repeated stimulus pairings. Hence, models of this type are *unsupervised* learning models. Theoretically, they posit that sophisticated semantic representations are the product of a relatively simple associative count mechanism operating on statistical regularities in the linguistic environment. These models are therefore quite direct instantiations of Firth's (1957) famous claim that “you shall know a word by the company it keeps”.

One major benefit to these models is that they are incremental learners. In other words, they develop their semantic representations continuously as experience unfolds. This allows them to make predictions about how representations change and “develop” as a function of the statistics that the model experiences. To relate passive co-occurrence models to the cognitively-oriented theories introduced in Section 3, passive co-occurrence models could be construed as modern instantiations of prototype models because statistical tendencies are, in effect, prototypes.

Perhaps the earliest passive co-occurrence model was the Hyperspace Analog to Language (HAL; Lund & Burgess, 1996). HAL simply slides a fixed width window (typically 5 to 10 words) across a text corpus, counting distance-weighted co-occurrences between words and storing these in a “long-term memory” matrix. Each word's semantic representation in the matrix is a function of the distance between it and every other word

within the window across learning. In effect, HAL produces a similarity space, with words that are similar distances from the same other words being in the same region of space (i.e., having similar meanings). Considering HAL's simplicity, it has been able to explain a large amount of human data, ranging from semantic priming to categorization (Burgess & Lund, 2000). A model similar to HAL was used to predict human fMRI brain activations when thinking about features of concepts given the concept's name and a picture (Mitchell et al., 2008). The original model has spawned several new versions with features that more carefully approximate human cognitive abilities (Shaoul & Westbury, 2010).

Another approach to building semantic representations using passive co-occurrence is commonly referred to as Random Vector Accumulation models (e.g., BEAGLE; Jones & Mewhort, 2007). These models treat each word as a random vector (of arbitrary dimensionality) such that initially, *dog* and *cat* are represented as being no more similar to one another than are *dog* and *bike*. Then, each time a word is experienced, its representation is summed together with all of the other vectors in its context (e.g., with the vectors of all of the other words in the sentence). As a consequence, across learning, a word's vector becomes a composite pattern of distributed activity reflecting its history of co-occurrence with other words. This means that, because *dog* and *cat* occur in similar contexts, across learning they move closer together in semantic space, whereas *bike*, due to not occurring in such similar contexts, moves further away. Thus, these models calculate “co-occurrence” differently than other models in that rather than calculating frequencies (i.e., rather than performing counts) to form semantic representations, they form representations by summing together all the vectors that constitute a word's context.

Although the passive models we have described learn from first-order co-occurrence, they are not limited to representing first-order statistical relationships. For example, the distributed representations of Random Vector Accumulation models such as BEAGLE end up having rich higher-order statistical relationships—even without direct co-occurrence, the model develops similar representations between words that appear in similar contexts. Thus, passive co-occurrence models show that sophisticated semantic representations can emerge via applying a simple Hebbian-based count to episodic co-occurrences.

### 7.2. Latent abstraction models

Latent Abstraction models also record the frequency of co-occurrence of words across linguistic contexts. However, rather than simply counting co-occurrences, they go a step further by reducing the dimensionality of the “episodic” matrix into an abstracted, lower dimensional “semantic” matrix. Hence, all models of this type posit a cognitive process that operates on episodic

experiences, reducing those experiences in such a way that it induces higher-order “latent” semantic relationships. Typically, latent abstraction models implement data reduction via mechanisms such as principal components analysis or factor analysis. Because latent abstraction models borrow dimensional reduction methods from statistics, they can only learn in batch—all episodic traces must first be represented before the learning mechanism can be applied (but see Olney, 2011). Thus, how they *acquire* semantic representations seems psychologically implausible. However, they are worth discussing because once the representations are acquired, the resulting representations appear to be quite plausible.

The best-known latent abstraction model is Latent Semantic Analysis (LSA; Landauer & Dumais, 1997). LSA begins with a word-by-document frequency matrix from a large text corpus. This is reduced into a semantic space in which two words can be very similar if they occurred in sufficiently similar contexts, even if they never directly co-occur (e.g., synonyms). The learning process in LSA is obviously not meant to simulate what humans do—we are unlikely to represent our lifetime of episodic experiences all at once and then factor that matrix. Landauer and Dumais were clear that humans do not use the same kind of dimensional reduction technique to learn semantics but, rather, that they use “some cognitive mechanism that achieves the same goal” (p. 218). While what that mechanism might be remains elusive, LSA has inspired many similar latent abstraction models (e.g. COALS; Rhode, Gonnerman, & Plaut, 2009) that have been among the most successful in the literature at explaining human data in a variety of semantic tasks.

### 7.3. Predictive models

Members of the family of *predictive* distributional semantic models also produce abstract distributed representations of word meanings, but their learning mechanisms are based on predictive coding and error-driven learning, two principles that are core to theories of reinforcement learning. Most of these models are connectionist in architecture. For example, early recurrent connectionist models such as those studied by St. John and McClelland (1990) and Elman (1990) learn a distributed pattern of elements across their hidden layers that reflects each word’s co-occurrence with other words (as in HAL, described earlier) in a continuous (technically, infinite) window. Rogers and McClelland (2004) likewise studied the ability of classic feedforward Rumelhart networks to learn distributed representations from simple languages, and found that very sophisticated (even hierarchical) higher-order relations naturally emerged after sufficient experience.

Both recurrent (e.g., Elman, 1990) and feed-forward (e.g., Rogers & McClelland, 2004) supervised networks learn by feeding activation forward through the network to generate its predicted output, and then derive an

error signal—the difference between the predicted output and the actual value observed in the training corpus. This signal is used to backpropagate (Rumelhart, Hinton, & Williams, 1986) the error through the layers of the network to increase the likelihood that the correct output will be predicted given the input in the future. While early predictive models produced interesting generalization behavior, they required many passes through the data to learn stable representations, and did not scale well beyond small artificial languages as effectively as did latent abstraction models such as LSA. Nonetheless, such models are interesting because by constraining the size of the hidden layers, they develop representations that are more akin to latent semantic variables (c.f. LSA) than to simple passive co-occurrence.

There has been a recent resurgence of interest in predictive models of distributional semantics. Howard et al. (2011) trained a predictive version of the Temporal Context Model, a recurrent model of error-driven hippocampal learning, on a large text corpus and demonstrated impressive performance on word association tasks. The predictive Temporal Context Model associates local item representations to a gradually drifting representation of time (temporal context) to learn distributed semantic representations from a text corpus. Although it differs considerably from classic connectionist models in architecture, it shares the relation to error-driven reinforcement learning theories.

A new type of model that has gained immense popularity recently due to its performance in the field of computational linguistics is the “neural embedding” model. Perhaps the best known is Google’s word2vec model (Mikolov et al., 2013), partially due to its remarkable performance on semantic tasks. But a key difference between this model and other systems led by industry (that are engineered to perform a *single* task well) is that word2vec is based on the same theoretical reinforcement-learning architecture that original recurrent language models were based on (Elman, 1990). Word2vec is a predictive model with hidden and recurrent layers, very much like the predictive Temporal Context Model, but operating on discrete time steps. It uses several training “tricks” to scale up to extremely large word corpora that Elman’s (1990) networks were never able to. Word2vec has had a major impact on the machine learning literature because it outperforms every other semantic model on a large battery of semantic tasks while using an architecture on which the field was no longer focused (see Baroni et al., 2014 for a careful comparison of state-of-the-art co-occurrence models and word2vec). Hollis and Westbury (2016) recently explored the organizational principles that word2vec is discovering to construct its semantic representations, concluding that the model primarily converges on affective dimensions of language when constructing semantic representations, very much in the

spirit of Osgood's early theories of meaning (Osgood, Suci, & Tannenbaum, 1957). Despite its excellent scaling properties and performance when trained on massive amounts of text, it remains to be seen if word2vec is an appropriate cognitive model of learning at human-like scales. For example, Asr, Willits, and Jones (2016) found that word2vec had much more difficulty learning simple semantic categories from child-directed speech using the CHILDES corpus than did a simple principal components analysis-based latent abstraction model.

An important benefit of predictive models is that they embody principles that seem to be ubiquitous within the cognitive system (cf. the recent interest on prediction during language comprehension, Altmann & Mirkovic, 2009). Moreover, the simplicity of Elman's (1990) recurrent network is inherently attractive because of its transparency regarding how, and what, it learns. However, the issue of scaling, both scaling up to realistically large adult-sized input, and scaling down to realistically small, child-sized input, is an important one. Being concerned with scaling is no different (for a cognitive psychologist) than being concerned with the hypothesis that links principles of learning and semantic organization to the architecture of the brain that embodies such principles.

#### 7.4. Bayesian models

Bayesian models of cognition have seen immense growth over the past decade, both due to their success at explaining human data across many cognitive levels, and because the approach offers a unifying theoretical framework to understand human cognition as rational probabilistic inference. The Bayesian approach assumes that the pattern of word co-occurrences across documents is generated by mixtures of semantic topics, where a topic is a probability distribution over words, and a word is a probability distribution over topics. The topics themselves are estimated from the observed data using Bayesian inference.

It is helpful to think of topics as folders on a shelf, where each folder contains words that are most associated with that topic, and each possible topic has one folder. For example, reaching into the *finance* folder you are more likely to pull out words such as *market*, *bank*, or *investment* than *giraffe*, *zebra*, or *hippo* (this same principle underlies LSA, described earlier). Whereas all folders contain all words, their probabilities differ. *Giraffe* is much more likely to be pulled out of the *zoo* folder than is *investment*. In the model, it is assumed that documents were generated by weighted mixtures of these topic folders. A news article describing a new fertility drug may have been generated by a recipe that calls for a mix of topics including 1 part *health*, 2 parts *pharmaceutical*, a dash of *reproduction*, and a pinch of *finance*. This "recipe" example is a slight oversimplification of a graded

process in that all topics may be sampled from, but their contributions are probabilistic.

Hence, a topic model estimates the most likely set and mixture of topics that would have generated the text. When applied to a large general text corpus, the model presumably uncovers general semantic topics that represent the shared world knowledge across individuals. When applied to a more specific text corpus, a topic model can uncover the author-specific topics that are most likely to have generated the observed data; this type of analysis can be very useful to explore knowledge expertise. For example, Murdock, Allen, and DeDeo (2015) recently used topic models to explore Darwin's reading notebooks, providing new insights about how his theory of natural selection emerged from the semantic path of his readings and writings.

Topic models differ from traditional distributional semantic models in theoretically important ways that deserve some highlighting here. First, topic models are generative. They specify a process by which words in documents are generated as a pattern of topic mixtures. Whereas abstraction models such as LSA do uncover variance components representing word co-occurrences in the text corpus, the process is more one of data mining and cannot easily be reversed to explain how future documents would be generated from the model.

Second, whereas most distributional semantic models are geometric models that represent a word as a point in high-dimensional space, topic models represent words as probability distributions. As a result, topic models naturally represent asymmetric semantic relations that are seen in human data, but that are difficult to account for with geometric models. For example, in free association norms, asymmetries are common: *baby* is a much more likely response to *stork* than *stork* is to *baby*. In a spatial representation, however, the distance between *baby* and *stork* must be the same regardless of which one serves as the cue. This symmetry need not be the case in topic models, where  $p(\textit{baby}|\textit{stork}) > p(\textit{stork}|\textit{baby})$ . Griffiths Steyvers, and Tenenbaum (2007) demonstrate several examples of how topic models explain patterns of human semantic data that violate symmetry and cannot be explained by purely geometric models (but see Jones, Gruenenfelder, & Recchia, 2011).

#### 7.5. Retrieval-based models

A final family of distributional semantic model has seen considerably less attention in the literature, but nonetheless deserves mention because members of this family make a radically different theoretical claim about semantic representation. *Retrieval-based* models essentially posit that there is no such thing as semantic memory, only episodic memory. Thus, they have considerable overlap with the exemplar-based models

introduced in Section 3.1 and in the chapter in this Handbook by Ashby and Valentin.

In retrieval-based models, semantic representations are an artifact (i.e., an emergent property) of retrieval from episodic memory. Most models of this family posit that each experience with a word lays down a unique trace in episodic memory. When a word is encountered, all of its previous experienced contexts are reactivated, and the process of retrieval from episodic memory incidentally recruits episodes of contexts that are quite similar. The aggregate of all episodes in the retrieved representation (implemented as a multi-dimensional vector) contains higher-order relationships of semantic similarity. Even though two words (e.g., synonyms) may never have co-occurred in the same context, their retrieved vectors will be similar to each other.

The constructed semantics model of Kwantes (2005) is an example of a retrieval-based version of LSA. This model is based heavily on Minerva 2 (Hintzman, 1986), a multiple-trace memory model that originally was designed as an existence proof that “schema abstraction”, that is, the process by which a prototype appears to be abstracted from exemplars (Posner & Keele, 1968), could be an emergent phenomenon from an episodic memory model. This was a significant demonstration because schema abstraction had been used as strong evidence of multiple memory systems—both episodic and semantic systems, consistent with Tulving’s classic taxonomy. Kwantes’ constructed semantics model is similar to LSA, but the data reduction mechanism is a natural byproduct of episodic retrieval. In addition, the semantic representation itself is never stored; it simply is an emergent artifact of episodic retrieval.

Retrieval-based models therefore have two major distinctions from all other distributional semantic models. First, semantic memory in these models does not exist as a stored structure; the model stores only episodes. This essentially is equivalent to storing the word-by-context matrix that is used as input to other models, such as LSA or Topics. Second, there is no abstraction mechanism when learning. In contrast to models such as LSA, Topics, BEAGLE, and word2vec that all apply a data reduction mechanism to construct an abstracted semantic representation for storage during input (singular value decomposition, Bayesian inference, backpropagation), retrieval-based models essentially do this at output. In other words, all other models place the abstraction at encoding, whereas retrieval models place abstraction at retrieval. In addition, the abstraction is not a purposeful mechanism *per se*. Instead, abstraction incidentally occurs because our memory retrieval mechanism is reconstructive. Hence, semantic memory in retrieval-based models is essentially an accident due to our imperfect memory retrieval process.

## 7.6 Challenges for distributional semantic models

Despite their impressive success at accounting for human data, distributional semantic models face some challenges as psychological models of human semantic representation. First and foremost, they routinely are criticized because they learn exclusively from text, that is, they construct semantic representations by applying a processing mechanism to statistical regularities in how words are used in a text corpus. In contrast, human semantic representations are constructed from a lifetime of linguistic and perceptual experience. Perceptual information is an inherent part of the organization of human semantic memory, but much of this information cannot be learned from statistics in a text corpus—it must be learned from multisensory experience (see the chapter by Matheson & Barsalou).

The challenge resulting from distributional semantic models being limited to text input however, is not a limitation of them as learning models *per se*, but is rather a limitation in the type of data we are currently able to feed to them. Hence, it is possible that a distributional semantic learning model could be applied to perceptual and motoric information as well as linguistic information to construct a more refined semantic representation, if we could get the model to “see”, etc. Indeed, such a suggestion was made in the discussion section of the original LSA article (Landauer & Dumais, 1997). Recent distributional semantic models have taken up this challenge, constructing joint probabilistic or composite semantic spaces that integrate complementary information sources from both linguistic and perceptual streams. “Perception” in these models is either represented using semantic feature norms (McRae et al., 2005; Vinson & Vigliocco, 2008), geon-based representations (Kievit-Kylar & Jones, 2011), or computer vision algorithms applied to natural images (Bruni, Tran, & Baroni, 2014). For an extended discussion, see the chapter in this Volume by Cai and Vigliocco. Reasonable success has been observed in several theoretical frameworks, including Bayesian models (Andrews, Vigliocco, & Vinson, 2009; Steyvers, 2009), random vector models (Jones & Recchia, 2010), and retrieval-based models (Johns & Jones, 2012). It remains to be seen whether, by providing distributional semantic models with not only linguistic and perceptual information, but also the ability to interact with the environment (cf. autonomous robots), such models could mimic human representations not only with respect to processing (e.g., the predictions that they make) and organization (e.g., similarity space), but also with respect to format (e.g., with parts supporting action also representing action knowledge).

A second challenge faced by distributional semantic models is that most do not reflect the evidence that the episodic and semantic systems are at least partially distinct (see Section 4). Some neurobiologically inspired

computational models of semantic memory, however, do posit some distinction between these systems, taking inspiration from the different computational properties of the neocortical and hippocampal structures that support memory (Howard, Shankar, & Jagadisan, 2011; Kumaran, Hassabis, & McClelland, 2016; Kumaran & McClelland, 2012; McClelland, McNaughton, & O'Reilly, 1995; Norman & O'Reilly, 2003; Schapiro, Turk-Browne, Botvinick, & Norman, 2017). Most of these models are based on Complementary Learning Systems theory (McClelland, McNaughton, & O'Reilly, 1995). According to this theory, episodic memory is supported by hippocampal structures that rapidly encode distinct episodes by means of large changes in connectivity within the hippocampus and between the hippocampus and neocortex. In contrast, semantic memory is supported by slower and smaller changes in neocortical connectivity that encode regularities encountered across multiple episodes. Although few of the models based on Complementary Learning Systems have been scaled up to cope with realistically-sized conceptual spaces (but see Howard et al., 2011), their role in potentially constraining psychological theories of semantic memory is undeniably important and highlights the value of using neurobiological considerations to shape computational models.

## 7.7 Summary of computational models of semantic representation

Because we have discussed some of the limitations of distributional semantic models, it might be tempting to think that these limitations, coupled with the diversity of the models, compromise their contribution to psychological theories of semantic memory. However, even the simplest computational models of learning, abstraction, and semantic representation (e.g. Elman's recurrent network model; Elman, 1990) can strongly influence psychological theory. Concepts such as prediction, abstraction, generalization, latent variables, and semantic space can be better understood by having an implemented computational perspective on how they might arise in *any* computational system. And even if it seems implausible that people wait to process the input until all of it has been gathered (c.f. the Latent Semantic models described in Section 7.2), the idea that there exists latent structure is important when considering what is encoded in semantic memory, and how it might be derived from experience. Similarly, although the idea that there is no such thing as semantic memory *per se* (c.f. the Retrieval-Based models of Section 7.3) challenges classical interpretations of the distinction between experience and abstraction, the retrieval-based models that make this claim are conceptually similar to the exemplar-based psychological models (Section 3.1) that posit that the process of *retrieval*, rather than *encoding*, may be responsible for how it is that people can generalize from

one experience to another. Thus, even principles that, on the surface, seem psychologically unlikely may ultimately prove to inform our understanding of cognitive processing. Finally, the models inspired by Complementary Learning Systems theory (7.6) make clear that fully understanding semantic memory will require better knowledge of the relationship between it and episodic memory, as well as an understanding of the neurobiology.

## 8. Semantic memory: a common thread, and a look ahead

The idea of abstraction recurs throughout this chapter. This is appropriate because abstraction is arguably at the root of semantic knowledge. In the remainder of this chapter, we revisit some essential roles of abstraction and raise questions for future research.

### 8.1 The roles of abstraction

We have suggested that the *process* of abstraction is one that, supported by sensitivity to statistical regularities, allows people to form and store (or, on exemplar/retrieval-based models, allows us to compute at retrieval) “abstracted” semantic information from our experiences (Section 2). This abstracted information would reflect what is most common across our experiences with a given object or event, such that idiosyncratic properties, like a leaf sticking to your dog's fur, would be minimized. It is through this kind of abstraction that we are able to discern what various objects and events have in common, and group them together into concepts.

We also described abstraction in terms of levels of abstraction (Section 6). The idea here is that information may be represented with different degrees of precision or detail, with representations containing more detail being referred to as “less abstract” or “at lower levels of abstraction” and representations containing less detail being referred to as “more abstract” or “at higher levels of abstraction”. For example, knowing that something is a *canary* requires more detailed semantic knowledge than knowing that something is a *bird*. In hierarchical descriptions of semantic knowledge (Section 3.1), the levels in the hierarchy can also be thought of as levels of abstraction, with superordinate (i.e., more inclusive) levels in the hierarchy corresponding to higher levels of abstraction (Rosch, 1978). Evidence suggests that depending upon the level at which we classify things (e.g., classifying a *labrador* as either a *dog* or a *living thing*), we rely on different brain regions to varying degrees (Section 6; cf. Hsu et al., 2011).

How are these different kinds of abstraction related? It is easy to imagine a model in which the way that the process of abstraction is implemented produces different levels of abstraction. For instance, it is possible

that sensitivity to statistical regularities not only allows us to derive information from across instances that is not apparent in any individual instance (as described in Section 2), but also that sensitivity to regularities across features from different modalities produces more complex relationships among features than could be apparent in any individual modality (Damasio 1989). Such sensitivities may result in higher level representations that have less perceptual detail and are further removed from the perceptual input in any single modality. Subsequently combining these higher level representations could produce yet higher level representations that even less directly reflect the unimodal perceptual input (for discussion, see Binder, 2016). Thus, sensitivity to cross-modal regularities could produce multiple levels of abstraction.

An important open question is how this cross-modal integration process might happen. One possibility, discussed in Sections 3 and 6, is that due to their connectivity and/or proximity to multiple sensory and motor areas, certain brain regions such as the anterior temporal lobe and angular gyrus, function as “hubs” in which information from multiple modalities converges and can therefore be integrated. Another possibility, however, is that rather than (or in addition to) integration being made possible by the *spatial* convergence of information, integration may occur as a result of the *temporal* convergence of information. In particular, researchers have asked questions like, to what extent does “the integration of multiple aspects of reality... depend on the time-locked co-activation of geographically separate sites of neural activity within sensory and motor cortices...” (Damasio, 1989, p. 39)? Interesting areas of research are the possible roles of various frequency bands when considering synchronous neural activity, and the possible roles of coherent oscillatory activity in integrating information from distributed brain regions (Singer & Gray, 1995). Several researchers have been exploring this question with respect to auditory and visual information (Widmann et al., 2007; Schneider et al., 2008; Bastiaansen et al., 2008; Van Ackeren & Rueschemeyer 2014).

## **8.2 Can abstraction help explain the representation of abstract concepts and features?**

We would be remiss to leave the topic of abstraction without considering an important type of semantic knowledge that we have not yet considered—knowledge about abstract concepts. So far, we have focused on what might be termed “simple concepts”, that is, the kind of object and action concepts that we learn early on during development, and that we use in our basic interactions with the world. However, concrete concepts can contain features (e.g., the functions of many objects, such as *used to tell time*) that do not have obvious sensorimotor correlates. Similarly, much of our

language is about concepts that have no physical manifestation, such as intellectual discourse or presidential debates. What sensory or motor attributes, or abstraction across those attributes, might constitute our representations of *purpose* or *hope*, concepts that seem to be devoid of physical and perceptual instantiation?

The predominant theory of how abstract concepts are represented emphasizes the importance of emotional or affective information in their representation (Vigliocco et al., 2009). Neural investigations have provided support for this idea in that brain regions associated with emotion processing are more active when processing abstract than concrete words (Vigliocco et al., 2014). However, while this account may help explain the representation of abstract concepts that do indeed involve emotion (e.g., *love*), many abstract concepts are not as clearly associated with emotion (e.g., *purpose*). It has therefore been suggested that abstract concepts for which emotional and/or sensory and motor attributes are lacking are more dependent on linguistic (Paivio 1991), and contextual/situational information (Barsalou & Wiemer-Hastings, 2005). That is, their mention in different contexts (i.e., episodes) may gradually lead us to an understanding of their meaning in the absence of sensorimotor content (Pulvermuller, 2013). Neural investigations have supported at least the linguistic portion of this proposal. Brain regions known to support language show greater involvement during the processing of abstract relative to concrete concepts (see Wang et al., 2010, for review).

If contextual/situational information and linguistic information (including verbal labels) are indeed important components of abstract knowledge, these should serve as important inputs into the process of abstraction that we have already described. Although we focused on sensorimotor inputs in our prior descriptions of this process, in principle it should function over any input from which statistical regularities can be derived. In fact, the computational models described in Section 7 are a testament to the ability of abstraction processes to function over non-sensorimotor input.

## **8.3 Is there differential reliance on different types of semantic information across development and across levels of abstraction?**

The question of whether abstract knowledge relies more on linguistic or situational information compared to concrete concepts raises the question of whether other aspects of semantic knowledge also rely more on some types of information more than others. For instance, as discussed in Section 5.1, labels may serve as an “invitation to form categories” in that they encourage attending to commonalities between things that share

the same label (Waxman & Markow, 1995). It is therefore possible that for categories for which the commonalities among instances are particularly hard to identify (such as superordinate categories, like “animal” or “tool”), labels may play a particularly important role in that without them, the commonalities would go unnoticed. Relatedly, it is possible that one reason that superordinate categories are later to develop is that they are particularly dependent on the ability to focus on specific aspects of a representation (e.g., the label), and the ability to focus in this way depends on prefrontal cortex, the brain region that takes longest to develop (see Sloutsky, 2010 for discussion). It will be interesting to discover whether the development of prefrontal cortex indeed influences the extent to which children rely on labels versus sensorimotor information when learning new semantic knowledge (see Section 4).

## 9. An integrated memory system

The data and theories that we have discussed in this chapter show that semantic memory should not be considered as fully independent from either episodic or implicit memory. In this section, we first review the evidence that semantic memory is part of an integrated memory system, and we then consider an implication of this integration.

### 9.1 Episodic, implicit, and semantic memories

There are several reasons to believe that episodic memory is an important component of semantic memory. First, without a functioning episodic memory, acquiring new semantic knowledge is slow and laborious, suggesting that episodic memory typically is important for the development of semantic knowledge (Section 4). Second, for most concepts, the acquisition process begins with a specific episode, although that episode may not later be consciously linked to the concept. This means that early in the process of learning about a new object, our knowledge may be heavily (or at least more heavily than later) influenced by episodic information. Conversely, there is evidence that our semantic knowledge supports our ability to acquire new episodic information (Kan et al., 2009). Our everyday experience is thus due to an interplay between episodic and semantic memory (see Altmann, 2017, for an account of this interplay in the context of language and event comprehension). Exemplar/retrieval-based models of semantic memory blur the distinction further, essentially eliminating semantic representations and instead positing that semantic knowledge is an emergent property of the process of retrieving episodes (see Sections 3.1 and 7.5).

Unlike episodic memory, implicit memory is typically defined as knowledge that we never had conscious

access to, but nevertheless affects our responses (e.g., procedural knowledge, such as how to ride a bike, or perceptual priming whereby our response to a stimulus is facilitated by prior exposure to a related stimulus, despite that fact that we are not aware of the relationship). By positing that the sensory and motor regions that are active when we perceive or interact with an object also have a role in representing it, sensorimotor-based models of semantic memory make clear predictions that implicit knowledge has a role in semantic knowledge. For instance, according to sensorimotor-based models, knowledge that is not easy to verbalize, such as how to ride a bike, can be part of your representation of *bike*, and how you position your hand and fingers while using a pencil can be part of your representation of *pencil*. Similarly, sensorimotor-based models posit that similarity based on implicit knowledge, such as how objects are manipulated, should shape how semantic knowledge is organized, an assertion for which ample evidence exists (e.g., Myung et al., 2006).

### 9.2 Do concepts really change across time and differ between individuals?

Sensorimotor-based models of semantic memory are committed to representations being experience-based. And experiences necessarily differ across time and across individuals. This may seem to create a problem for sensorimotor-based models because most people have the intuition that their semantic representations are static, and that when they talk to others about them, they are talking about the same things. We suggest that both of these intuitions are false (albeit not entirely false). First, given that semantic representations necessarily change across development (see Section 4), we must at least accept changes then. This raises the question of when, if ever, development should be considered to end, and, relatedly, how to consider the changes that result from the acquisition of new, or more detailed semantic knowledge in adulthood. Second, as described in Section 5, there are clear individual differences in the representation of semantic knowledge that result from individual differences in the experiences that make up both our episodic memories and our implicit memories.

Despite these differences, we still (usually) seem to be referring to approximately the same things when we refer to an *apple* (whether the same goes for more abstract concepts like *justice* is open to debate). We suggest that the commonalities in human experience, due in part to the commonalities in the structure of the world in which we live, and in part to our common sensory and motor systems, allow our semantic representations, and the labels that refer to them, to be similar enough for successful communication (for further consideration of this point, see Casasanto & Lupyan, 2015; Connell & Lynott, 2014; Taylor & Zwaan, 2009; Yee & Thompson-Schill, 2016). Moreover, if, as

suggested by Complementary Learning Systems theory, semantic memory changes slowly via small changes in neocortical connectivity, we need not worry that sudden changes in our experiences will radically alter our conceptual knowledge. Rather, we would expect changes due to experience to build gradually as our experiences accumulate. Thus, an integrated model of the semantic memory system must take account of the episodic and the implicit knowledge that influence semantic memory across the lifespan.

## 10. A final note

The study of semantic memory is more than just the study of what we know and how we know it. To the extent that our knowledge shapes who we are, it is the study of who we are and why. A lesson to be learned from the insights that underpin the theories, data, and models that we have described is that our understanding of human memory relies on complementary investigations into behavior, neurobiology, and computation. Moreover, each of these perspectives is strengthened by considering the transition from the immature state to its mature counterpart. The challenge is to develop a model of human memory that reflects, across the lifespan, both the computations afforded by its neurobiology as well as the behaviors that these computations produce.

## Acknowledgements

We thank Katherine White, Gerry Altmann, and Charles Davis for helpful comments and discussion.

## References

- Abbott, J. T., Austerweil, J. L., & Griffiths, T. L. (2015). Random walks on semantic networks can resemble optimal foraging. *Psychological Review*, 122(3), 558-569.
- Allport, D. A. (1985). Distributed memory, modular subsystems and dysphasia. In S. K. Newman & R. Epstein (Eds.), *Current Perspectives in Dysphasia* (pp. 207-244). Edinburgh: Churchill Livingstone.
- Altmann, G.T.M. (2017). Abstraction and Generalization in Statistical Learning: Implications for the relationship between semantic types and episodic tokens. *Philosophical Transactions of the Royal Society B*.
- Altmann, G.T.M., & Mirkovic, J. (2009). Incrementality and prediction in human sentence processing. *Cognitive Science*, 33, 583-609.
- Amedi, A., Merabet, L. B., Bermpohl, F., & Pascual-Leone, A. (2005). The occipital cortex in the blind lessons about plasticity and vision. *Current Directions in Psychological Science*, 14(6), 306-311.
- Andrews, M., & Vigliocco, G. (2010). The hidden-Markov topic model: A probabilistic model of semantic representation. *Topics in Cognitive Science*, 2, 101-113.
- Andrews, M., Vigliocco, G. & Vinson, D. P. (2009). Integrating experiential and distributional data to learn semantic representations. *Psychological Review*, 116, 463-498.
- Arndt, J., & Hirshman, E. (1998). True and false recognition in MINERVA2: Explanations from a global matching perspective. *Journal of Memory and Language*, 39(3), 371-391.
- Ashby, F. G. & Valentin, V. V. (2017). The categorization experiment: Experimental Design and Data Analysis. In J. T. Wixted, E. J. Wagenmakers (Eds), *The Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience* (4th Edition, Volume 5: Methodology). New York: Wiley.
- Asr, F. T., Willits, J. A., & Jones, M. N. (2016). Comparing predictive and co-occurrence based models of lexical semantics trained on child-directed speech. *Proceedings of the 37th Meeting of the Cognitive Science Society*.
- Bayley PJ, Squire LR. 2005. Failure to acquire new semantic knowledge in patients with large medial temporal lobe lesions. *Hippocampus* 15, 273-280.
- Bayley, P. J., O'Reilly, R. C., Curran, T., & Squire, L. R. (2008). New semantic learning in patients with large medial temporal lobe lesions, *Hippocampus*, 18, 575-583.
- Baroni, M., Dinu, G., & Kruszewski, G. (2014, June). Don't count, predict! A systematic comparison of context-counting vs. context-predicting semantic vectors. In *ACL (1)* (pp. 238-247).
- Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral Brain Science*, 22, 577660.
- Barsalou, L. W. (2016). On staying grounded and avoiding quixotic dead ends. *Psychonomic bulletin & review*, 1-21.
- Barsalou, L. W., & Wiemer-Hastings, K. (2005). Situating abstract concepts. *Grounding cognition: The role of perception and action in memory, language, and thought*, 129-163.
- Bastiaansen, M. C., Oostenveld, R., Jensen, O., & Hagoort, P. (2008). I see what you mean: theta power increases are involved in the retrieval of lexical semantic information. *Brain and language*, 106(1), 15-28.
- Beilock, S. L., Lyons, I. M., Mattarella-Micke, A., Nusbaum, H. C., & Small, S. L. (2008). Sports experience changes the neural processing of action language. *Proceedings of the National Academy of Sciences*, 105, 13269-13273.
- Binder, J. R. (2016). In defense of abstract conceptual representations. *Psychonomic Bulletin & Review*, 23(4), 1096-1108.
- Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*, 19(12), 2767-2796.
- Bindschaedler, C., Peter-Favre, C., Maeder, P., Hirsbrunner, T., & Clarke, S. (2011). Growing up with bilateral hippocampal atrophy: from childhood to teenage. *Cortex*, 47(8), 931-944.
- Bruni, E., Tran, N. K., & Baroni, M. (2014). Multimodal Distributional Semantics. *J. Artif. Intell. Res. (JAIR)*, 49(1-47).
- Bullinaria, J. A., & Levy, J., P. (2007). Extracting semantic representations from word co-occurrence statistics: A computational study. *Behavior Research Methods*, 39, 510-526.
- Burgess, C. & Lund, K. (2000) The dynamics of meaning in memory. In E. Dietrich & A. B. Markman (Eds.), *Cognitive dynamics: Conceptual and representational change in humans and machines* (pp. 117-156). Mahwah, NJ: Erlbaum.
- Cai, Z.G. & Vigliocco, G. (2017). Word processing. In J. T. Wixted & S. Thompson-Schill (Eds), *The Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience* (4th Edition, Volume 3: Language and Thought). New York: Wiley.
- Caramazza, A., Shelton, J.R., (1998). Domain-specific knowledge systems in the brain the animate-inanimate distinction. *Journal of Cognitive Neuroscience* 10, 1-34
- Carey, S. (2009). The origin of concepts. New York: Oxford University Press.
- Casasanto, D. (2009). Embodiment of abstract concepts: good and bad in right-and left-handers. *Journal of Experimental Psychology: General*, 138(3), 351.

- Casasanto, D., & Chrysikou, E. G. (2011). When left is “right” motor fluency shapes abstract concepts. *Psychological Science*, 22(4), 419-422.
- Casasanto, D., & Lupyan, G. (2015). All concepts are ad hoc concepts. *The conceptual mind: New directions in the study of concepts*, 543-566.
- Chan, A. M., Baker, J. M., Eskandar, E., Schomer, D., Ulbert, I., Marinkovic, K., Cash, S. S., & Halgren, E. (2011). First-pass selectivity for semantic categories in human anteroventral temporal lobe. *The Journal of Neuroscience*, 31(49), 18119-18129.
- Clarke, A., Pell, P. J., Ranganath, C., & Tyler, L. K. (2016). Learning warps object representations in the ventral temporal cortex. *Journal of Cognitive Neuroscience*, 28(7), 1010-1023.
- Clarke, A., Taylor, K. I., & Tyler, L. K. (2011). The evolution of meaning: Spatiotemporal dynamics of visual object recognition. *Journal of Cognitive Neuroscience*, 23, 1887-1899.
- Cohen, N. J., & Squire, L. R. (1980). Preserved learning and retention of pattern analyzing skill in amnesia: Dissociation of knowing how and knowing that. *Science*, 210, 207-209.
- Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, 82, 407-428.
- Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, 8, 240-247.
- Connell, L., & Lynott, D. (2014). Principles of representation: Why you can't represent the same concept twice. *Topics in Cognitive Science*, 6, 390-406.
- Connolly, A. C., Gleitman, L. R., & Thompson-Schill, S. L. (2007). The effect of congenital blindness on the semantic representation of some everyday concepts. *Proceedings of the National Academy of Sciences*, 104(20), 8241-8246. PMID: 17483447.
- Cox, G. E., Kachergis, G., Recchia, G., & Jones, M. N. (2011). Toward a scalable holographic word-form representation. *Behavior research methods*, 43(3), 602-615.
- 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, & Cognition*, 32, 643-658.
- Cross, E. S., Cohen, N. R., Hamilton, A. F. D. C., Ramsey, R., Wolford, G., & Grafton, S. T. (2012). Physical experience leads to enhanced object perception in parietal cortex: insights from knot tying. *Neuropsychologia*, 50(14), 3207-3217.
- Damasio, A. R. (1989). Time-locked multiregional retroactivation: A systems-level proposal for the neural substrates of recall and recognition. *Cognition*, 33(1), 25-62.
- Deng, W., & Sloutsky, V. M. (2012) Carrot-eaters and moving heads: Salient features provide greater support for inductive inference than category labels. *Psychological Science*, 23, 178-186
- Dennis, S. (2005). A memory-based theory of verbal cognition. *Cognitive Science*, 29, 145-193.
- Dunham, P., & Dunham, F. (1995). Developmental antecedents of taxonomic and thematic strategies at 3 years of age. *Developmental Psychology*, 31(3), 483.
- Elman, J. (1990). Finding structure in time. *Cognitive Science*, 14, 179-211.
- 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.
- Firth, J. R. (1957). A synopsis of linguistic theory, 1930-1955. In *Philological Society (Great Britain) (Ed.)*, *Studies in linguistic analysis*. Oxford: Blackwell.
- Gabrieli, J. D., Cohen, N. J., & Corkin, S. (1988). The impaired learning of semantic knowledge following bilateral medial temporal-lobe resection. *Brain and cognition*, 7(2), 157-177.
- Gainotti G. 2000. What the locus of brain lesion tells us about the nature of the cognitive defect underlying category-specific disorders: a review. *Cortex*. 36, 539-559.
- Gainotti, G. (2015). Is the difference between right and left ATLs due to the distinction between general and social cognition or between verbal and non-verbal representations? *Neuroscience & Biobehavioral Reviews*, 51, 296-312.
- Gardiner, J.M., Brandt, K.R., Baddeley, A.D., Vargha-Khadem, F., Mishkin, M. (2008). Gardiner, J. M., Brandt, K. R., Baddeley, A. D., Vargha-Khadem, F., & Mishkin, M. (2008). Charting the acquisition of semantic knowledge in a case of developmental amnesia. *Neuropsychologia*, 46(11), 2865-2868.
- Ghio, M., Schulze, P., Suchan, B., & Bellebaum, C. (2016). Neural representations of novel objects associated with olfactory experience. *Behavioural brain research*, 308, 143-151.
- Goldstone, R. L., Kersten, A. & Carvalho, P. (2017). Categorization and concepts. In J. T. Wixted & S. Thompson-Schill (Eds), *The Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience* (4th Edition, Volume 3: Language and Thought). New York: Wiley.
- Grabowski, T. J., Damasio, H., Tranel, D., Boles Ponto, L. L., Hichwa, R. D., & Damasio, A. R. (2001). A role for left temporal pole in the retrieval of words for unique entities. *Human Brain Mapping*, 13, 199-212.
- Griffiths, T. L., Steyvers, M., & Tenenbaum, J. B. (2007). Topics in semantic representation. *Psychological Review*, 114, 211-244.
- Gureckis, T. (2017). Category learning. In J. T. Wixted, E. Phelps & L. Davachi (Eds), *The Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience* (4th Edition, Volume 1: Learning & Memory). New York: Wiley.
- Harris, Z. (1970). Distributional structure. In *Papers in Structural and Transformational Linguistics* (pp. 775-794). Dordrecht, Holland: D. Reidel Publishing Company.
- Hintzman, D. L. (1986). “Schema abstraction” in a multiple-trace memory model. *Psychological Review*, 93, 411-428.
- Hodges, J. R., & Patterson, K. (2007). Semantic dementia: a unique clinicopathological syndrome. *The Lancet Neurology*, 6(11), 1004-1014.
- Hoening K, Muller C, Herrnberger B, Spitzer M, Ehret G, and Kiefer M. (2011). Neuroplasticity of semantic maps for musical instruments in professional musicians. *NeuroImage*, 56(3), 1714-1725.
- Hollis, G., & Westbury, C. (2016). The principals of meaning: Extracting semantic dimensions from co-occurrence models of semantics. *Psychonomic Bulletin & Review*, 1-13.
- Howard MW, Shankar KH, Jagadisan UK. 2011 Constructing semantic representations from a gradually changing representation of temporal context. *Topics in Cognitive Science*. 3, 48 – 73. (doi:10.1111/j. 1756-8765.2010.01112.x)
- Howard, M. W., Shankar, K. H., & Jagadisan, U. K. (2011). Constructing semantic representations from a gradually changing representation of temporal context. *Topics in Cognitive Science*, 3(1), 48-73.
- Hsu, N.S., Kraemer, D.J.M., Oliver, R.T., Schlichting, M.L., & Thompson-Schill, S.L. (2011). Color, Context, and Cognitive Style: Variations in Color Knowledge Retrieval as a Function of Task and Subject Variables. *Journal of Cognitive Neuroscience*, 29(3), 2544-2557.
- Hummel, J. E., & Holyoak, K. J. (2003). A symbolic-connectionist theory of relational inference and generalization. *Psychological Review*, 110, 220-264.

- Humphries, C., Binder, J. R., Medler, D. A., & Liebenthal, E. (2007). Time course of semantic processes during sentence comprehension: an fMRI study. *Neuroimage*, 36(3), 924-932.
- Jackson, R. L., Hoffman, P., Pobric, G., & Ralph, M. A. L. (2015). The nature and neural correlates of semantic association versus conceptual similarity. *Cerebral Cortex*, 25(11), 4319-4333.
- Ji, L. J., Zhang, Z., & Nisbett, R. E. (2004). Is it culture or is it language? Examination of language effects in cross-cultural research on categorization. *Journal of personality and social psychology*, 87(1), 57.
- John, M. F. S., & McClelland, J. L. (1990). Learning and applying contextual constraints in sentence comprehension. *Artificial Intelligence*, 46(1-2), 217-257.
- Johns, B. T., & Jones, M. N. (2012). Perceptual inference from global lexical similarity. *Topics in Cognitive Science*, 4:1, 103-120.
- Jones, M. N., & Mewhort, D. J. K. (2007). Representing word meaning and order information in a composite holographic lexicon. *Psychological Review*, 114(1), 1.
- Jones, M. N., & Recchia, G. L. (2010). You can't wear a coat rack: A binding framework to avoid illusory feature migrations in perceptually grounded semantic models. In S. Ohissou & R. Catrambone (Eds.), *Proceedings of the 32nd Annual Cognitive Science Society*. Austin TX: CSS.
- Jones, M. N., Gruenenfelder, T. M., & Recchia, G. (2011). In defense of spatial models of lexical semantics. In L. Carlson, C. Hölscher, & T. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 3444-3449). Austin, TX: Cognitive Science Society.
- Jones, M. N., Hills, T. T., & Todd, P. M. (2015). Hidden processes in structural representations: A reply to Abbott, Austerweil, and Griffiths). *Psychological Review*, 122, 570-574.
- Jones, M. N., Willits, J. A., Dennis, S., & Jones, M. (2015). Models of semantic memory. *Oxford Handbook of Mathematical and Computational Psychology*, 232-254.
- Kable, J. W., Kan, I. P., Wilson, A., Thompson-Schill, S. L., and Chatterjee, A. (2005). Conceptual representations of action in lateral temporal cortex. *Journal of Cognitive Neuroscience*, 17, 855-870.
- Kalénine, S., Peyrin, C., Pichat, C., Segebarth, C., Bonthoux, F., & Baciou, M. (2009). The sensory-motor specificity of taxonomic and thematic conceptual relations: A behavioral and fMRI study. *Neuroimage*, 44(3), 1152-1162.
- Kalénine, Solène, and Laurel J. Buxbaum. "Thematic knowledge, artifact concepts, and the left posterior temporal lobe: Where action and object semantics converge." *Cortex* 82 (2016): 164-178.
- Kan, I. P., Alexander, M. P., & Verfaellie, M. (2009). Contribution of prior semantic knowledge to new episodic learning in amnesia. *Journal of Cognitive Neuroscience*, 21, 938-944.
- Kan, I. P., Kable, J.W., Van Scoyoc, A., Chatterjee, A., & Thompson-Schill, S. L. (2006). Fractionating the left frontal response to tools: Dissociable effects of motor experience and lexical competition. *Journal of Cognitive Neuroscience*, 18, 267-277.
- Kanerva, P. (2009). Hyperdimensional computing: An introduction to computing in distributed representations with high-dimensional random vectors. *Cognitive Computation*, 1, 139-159.
- Kiefer, M., Sim, E.-J., Liebich, S., Hauk, O. & Tanaka, J. (2007). Experience-dependent plasticity of conceptual representations in human sensory-motor areas. *Journal of Cognitive Neuroscience*, 19, 525-542.
- Kievit-Kylar, B., & Jones, M. N. (2011). The semantic Pictionary project. In L. Carlson, C. Hölscher, & T. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 2229-2234). Austin, TX: Cognitive Science Society.
- Kumaran D, Hassabis D, McClelland JL. 2016 What learning systems do intelligent agents need? complementary learning systems theory updated. *Trends in Cognitive Science*. 20, 512 – 534. (doi:10.1016/j.tics. 2016.05.004)
- Kumaran D, McClelland JL. 2012 Generalization through the recurrent interaction of episodic memories: a model of the hippocampal system. *Psychological Review*, 119, 573 – 616. (doi:10.1037/a0028681)
- Kwantes, P. J. (2005). Using context to build semantics. *Psy Bull & Review*, 12, 703-710.
- Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to western thought*. Basic books.
- Landauer, T., & Dumais, S. (1997). A solution to Plato's problem: The Latent Semantic Analysis theory of the acquisition, induction, and representation of knowledge. *Psychological Review*, 104, 211-240.
- Lin, E. L., & Murphy, G. L. (2001). Thematic relations in adults' concepts. *Journal of Experimental Psychology: General*, 130(1), 3.
- Lund, K., & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavioral Research Methods, Instrumentation, and Computers*, 28, 203-208.
- Mack, M. L., & Palmeri, T. J. (2015). The dynamics of categorization: Unraveling rapid categorization. *Journal of Experimental Psychology: General*, 144(3), 551.
- Mahon, B. Z., & Caramazza, A. (2009). Concepts and categories: A cognitive neuropsychological perspective. *Annual Review of Psychology*, 60, 27.
- Mahon, B.Z., Anzellotti, S., Schwarzbach, J., Zampini, M., and Caramazza, A. (2009). Category-specific organization in the human brain does not require visual experience. *Neuron*, 63, 397-405.
- Martin, A., & Chao, L. L. (2001). Semantic memory and the brain: Structure and processes. *Current Opinion in Neurobiology*, 11, 194-201.
- Matheson, H. E. & Barsalou, L. (2017). Embodiment and grounding in cognitive neuroscience. In J. T. Wixted & S. Thompson-Schill (Eds), *The Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience* (4th Edition, Volume 3: Language and Thought). New York: Wiley.
- McClelland JL, McNaughton BL, O'Reilly RC. 1995 Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, 102, 419. (doi:10.1037/0033-295X.102.3.419)
- McKoon, G., & Ratcliff, R. (1992). Spreading activation versus compound cue accounts of priming: Mediated priming revisited. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18, 1155-1172.
- McNorgan, C., Reid, J., & McRae, K. (2011). Integrating conceptual knowledge within and across representational modalities. *Cognition*, 118(2), 211-233. Chicago.
- McRae, K., Cree, G., Seidenberg, M. S., & McNorgan, C. (2005). Semantic feature production norms for a large set of living and nonliving things. *Behavior Research Methods*, 37, 547-59.
- 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, 99-130.
- Medin, D. L. & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review*, 85, 207-238.
- Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review*, 85, 207-238. doi:10.1037/0033-295X.85.3.207
- Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. In *Advances in neural information processing systems* (pp. 3111-3119).

- Milner, B., Corkin, S., & Teuber, H. L. (1968). Further analysis of the hippocampal amnesic syndrome: 14 year follow-up study of H.M. *Neuropsychologia*, 6, 215–234.
- Mirman, D., & Graziano, K. M. (2012). Individual differences in the strength of taxonomic versus thematic relations. *Journal of experimental psychology: General*, 141(4), 601.
- Mitchell, T. M., Shinkareva, S. V., Carlson, A., Chang, K. M., Malave, V. L., Mason, R. A., & Just, M. A. (2008). Predicting human brain activity associated with the meanings of nouns. *Science*, 320(5880), 1191-1195.
- Murdock, J., Allen, C., & DeDeo, S. (2015). Exploration and Exploitation of Victorian Science in Darwin's Reading Notebooks. *arXiv preprint arXiv:1509.07175*.
- Murphy, G. L. (2002). *The Big Book of Concepts*. Cambridge, MA: MIT Press.
- Murphy, G. L. (2016). Is there an exemplar theory of concepts? *Psychonomic Bulletin & Review*, 23(4), 1035-1042.
- Murphy, G. L., & Allopenna, P. D. (1994). The locus of knowledge effects in concept learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 904–919. doi:10.1037/0278-7393.20.4.904
- Murphy, G. L., Hampton, J. A., & Milovanovic, G. S. (2012). Semantic memory redux: an experimental test of hierarchical category representation. *Journal of memory and language*, 67(4), 521-539.
- Myung, J. Y., Blumstein, S. E., & Sedivy, J. C. (2006). Playing on the typewriter, typing on the piano: manipulation knowledge of objects. *Cognition*, 98(3), 223-243.
- Noppeney, U. (2009). The sensory-motor theory of semantics: Evidence from functional imaging. *Language and Cognition*, 1(2), 249-276.
- Noppeney, U., Friston, K. J., & Price, C. J. (2003). Effects of visual deprivation on the organization of the semantic system. *Brain*, 126(7), 1620-1627.
- Norman KA, O'Reilly RC. 2003 Modeling hippocampal and neocortical contributions to recognition memory: A complementary-learning- systems approach. *Psychological Review*, 110, 611 – 646. (doi:10.1037/0033-295X.110.4.611)
- O'Kane G., Kensinger E. A., Corkin S. 2004. Evidence for semantic learning in profound amnesia: an investigation with patient H.M. *Hippocampus* 14, 417–425 (doi:10.1002/hipo.20005)
- Oliver, R. T., Geiger, E. J., Lewandowski, B. C., & Thompson-Schill, S. L. (2009). Remembrance of things touched: How sensorimotor experience affects the neural instantiation of object form. *Neuropsychologia*, 47, 239-247.
- Oliver, R. T., Parsons, M. A., & Thompson-Schill, S. L. (2008). Hands on learning: Variations in sensorimotor experience alter the cortical response to newly learned objects. *Cognitive Neuroscience Society*, San Francisco, CA.
- Olney, A. M. (2011). Large-scale latent semantic analysis. *Behavior research methods*, 43(2), 414-423.
- Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning*. Urbana, IL: University of Illinois Press.
- Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology*, 45, 255–287.
- 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. *Nature Reviews Neuroscience*, 8, 976-987.
- Perfetti, C. (1998). The limits of co-occurrence: Tools and theories in language research. *Discourse Processes*, 25, 363-377.
- Posner, M. I., & Keele, S. W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology*, 77(3p1), 353-363.
- Pulvermüller, F. (1999). Words in the brain's language. *Behavioral and Brain Sciences*, 22(02), 253-279.
- Pulvermüller, F. (2013). How neurons make meaning: brain mechanisms for embodied and abstract-symbolic semantics. *Trends in Cognitive Sciences*, 17(9), 458-470.
- Rakison, D. H., & Lawson, C. A. (2013). Categorization. *The Oxford Handbook of Developmental Psychology*, Vol. 1: Body and Mind. (pp. 591-627). Oxford University Press.
- Randall, B., Moss, H. E., Rodd, J., Greer, M., & Tyler, L. K. (2004). Distinctiveness and correlation in conceptual structure: Behavioral and computational studies. *Journal of Experimental Psychology: Language, Memory, and Cognition*, 30, 393-406.
- Recchia, G. L., & Jones, M. N. (2009). More data trumps smarter algorithms: Training computational models of semantics on very large corpora. *Behavior Research Methods*, 41, 657-663.
- Riordan, B., & Jones, M. N. (2011). Redundancy in linguistic and perceptual experience: Comparing distributional and feature-based models of semantic representation. *Topics in Cognitive Science*, 3, 1-43.
- Rizzolatti, G., Fogassi, L., & Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews Neuroscience*, 2(9), 661-670.
- Rogers, T. T. and McClelland, J. L. (2004). *Semantic cognition: A parallel distributed processing approach*. Cambridge, MA: MIT Press
- Rogers, T. T., & Patterson, K. (2007). Object categorization: Reversals and explanations of the basic-level advantage. *Journal of Experimental Psychology: General*, 136, 451-469.
- Rogers, T. T., Hocking, J., Noppeney, U., Mechelli, A., Gorno-Tempini, M. L., Patterson, K., & Price, C. J. (2006). Anterior temporal cortex and semantic memory: reconciling findings from neuropsychology and functional imaging. *Cognitive, Affective, & Behavioral Neuroscience*, 6(3), 201-213.
- Rohde, D. L. T., Gonnerman, L., & Plaut, D. C. (2009). An improved model of semantic similarity based on lexical co-occurrence. *Cognitive Science*.
- Rosch, E. (1978). Principles of categorization. In E. Rosch and B. B. Lloyd (Eds.), *Cognition and Categorization* (pp. 27–48). Hillsdale, NJ: Erlbaum.
- Rosch, E., Mervis, C. B., Gray, W., Johnson, D., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, 8, 382–439.
- Rüther, N. N., Brown, E. C., Klepp, A., & Bellebaum, C. (2014). Observed manipulation of novel tools leads to mu rhythm suppression over sensory-motor cortices. *Behavioural brain research*, 261, 328-335.
- Schapiro AC, Turk-Browne NB, Botvinick MM, Norman KA. 2017 Complementary learning systems within the hippocampus: a neural network modeling approach to reconciling episodic memory with statistical learning. *Phil. Trans. R. Soc. B* 372,
- Schneider, T. R., Debener, S., Oostenveld, R., & Engel, A. K. (2008). Enhanced EEG gamma-band activity reflects multisensory semantic matching in visual-to-auditory object priming. *Neuroimage*, 42(3), 1244-1254.
- Schneider, T. R., Debener, S., Oostenveld, R., & Engel, A. K. (2008). Enhanced EEG gamma-band activity reflects multisensory semantic matching in visual-to-auditory object priming. *Neuroimage*, 42(3), 1244-1254.
- Schwartz, M. F., Kimberg, D. Y., Walker, G. M., Brecher, A., Faseyitan, O. K., Dell, G. S., ... & Coslett, H. B. (2011). Neuroanatomical dissociation for taxonomic and thematic knowledge in the human brain. *Proceedings of the National Academy of Sciences*, 108(20), 8520-8524.
- Shaoul, C., & Westbury, C. (2006). Word frequency effects in highdimensional co-occurrence models: A new approach. *Behavior Research Methods*, 38, 190–195.

- Simmons, W. K., & Barsalou, L. W. (2003). The similarity-in-topography principle: Reconciling theories of conceptual deficits. *Cognitive Neuropsychology*, *20*, 451–486.
- Singer, W., & Gray, C. M. (1995). Visual feature integration and the temporal correlation hypothesis. *Annual review of neuroscience*, *18*(1), 555-586.
- Sloutsky, V. M. (2010). From perceptual categories to concepts: What develops? *Cognitive science*, *34*(7), 1244-1286.
- Sloutsky, V. M., Kloos, H., & Fisher, A. V. (2007). When looks are everything: Appearance similarity versus kind information in early induction. *Psychological Science*, *18*, 179–185.
- Sloutsky, V. M., Lo, Y. F., & Fisher, A. V. (2001). How much does a shared name make things similar? Linguistic labels, similarity, and the development of inductive inference. *Child development*, *72*(6), 1695-1709.
- Smith, E. E. (1978). Theories of semantic memory. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes* (Vol. 6, pp. 1–56). Potomac, MD: Erlbaum.
- Smith, E. E., & Medin, D. L. (1981). *Categories and concepts*. Cambridge, MA: Harvard University Press.
- Smith, E. E., Shoben, E. J., & Rips, L. J. (1974). Structure and process in semantic memory: A featural model for semantic decisions. *Psychological Review*, *81*, 214-241.
- Smith, L. B. (2005). Cognition as a dynamic system: Principles from embodiment. *Developmental Review*, *25*(3), 278-298.
- Spelke, E. S., Breinlinger, K., Macomber, J., & Jacobson, K. (1992). Origins of knowledge. *Psychological Review*, *99*(4), 605.
- Squire, 1987 L.R. Squire, *Memory and brain*, Oxford University Press, New York (1987).
- Squire, L. R., & Zola, S. M. (1998). Episodic memory, semantic memory, and amnesia. *Hippocampus*, *8*, 205–211.
- Steyvers, M. (2009). Combining feature norms and text data with topic models. *Acta Psychologica*, *133*, 234-243.
- Tanaka, J. M., & Presnell, L.M. (1999). *Perception & Psychophysics*, *61*, 1140-1153.
- Tanaka, K. (1996). Inferotemporal cortex and object vision. *Annual review of neuroscience*, *19*(1), 109-139.
- Taylor, L. J. & Zwaan, R. A. (2009). Action in cognition: The case of language. *Language and Cognition*, *1*, 45-58.
- Thompson-Schill, S. (2003). Neuroimaging studies of semantic memory: Inferring how from where. *Neuropsychologia*, *41*, 280-292.
- Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left prefrontal cortex in retrieval of semantic knowledge: A re-evaluation. *Proceedings of the National Academy of Sciences*, *94*, 14792-14797.
- Tulving (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of Memory* (pp. 381-403). New York: Academic Press.
- Tulving, E. (1991). Concepts of human memory. In L. Squire, G. Lynch, N.M. Weinberger, & J.L. McGaugh (Eds.), *Memory: Organization and Locus of Change* (pp. 3-32). New York: Oxford University Press.
- Tulving, E. (2002). Episodic memory: from mind to brain. *Annual review of psychology*, *53*(1), 1-25.
- Turney, P. D. and Pantel, P. (2010). From frequency to meaning: Vector space models of semantics. *Journal of Artificial Intelligence Research*, *37*, 141-188.
- Tyler, L. K., & Moss, H.E. (2001). Towards a distributed account of conceptual knowledge. *Trends in Cognitive Science*, *5*, 244–252.
- Tyler, L. K., Stamatakis, E. A., Bright, P., Acres, K., Abdalah, S., Rodd, J. M., & Moss, H. E. (2004). Processing objects at different levels of specificity. *Journal of Cognitive Neuroscience*, *16*, 351-362.
- van Ackeren, M. J., & Rueschemeyer, S. A. (2014). Cross-modal integration of lexical-semantic features during word processing: evidence from oscillatory dynamics during EEG. *PLoS one*, *9*(7), e101042.
- Vannuscorps, G., & Caramazza, A. (2016). Typical action perception and interpretation without motor simulation. *Proceedings of the National Academy of Sciences*, *113*(1), 86-91.
- Vargha-Khadem, F., Gadian, D. G., Watkins, K. E., Connelly, A., Van Paesschen, W., Mishkin, M. (1997). Differential effects of early hippocampal pathology on episodic and semantic memory. *Science* *277*, 376–380.
- Vigliocco, G., Kousta, S. T., Della Rosa, P. A., Vinson, D. P., Tettamanti, M., Devlin, J. T., & Cappa, S. F. (2014). The neural representation of abstract words: the role of emotion. *Cerebral Cortex*, *24*(7), 1767-1777.
- Vigliocco, G., Meteyard, L., Andrews, M., & Kousta, S. (2009). Toward a theory of semantic representation. *Language and Cognition*, *1*(2), 219-247.
- Vinson, D. P., & Vigliocco, G. (2008). Semantic feature production norms for a large set of objects and events. *Behavior Research Methods*, *40*(1), 183-190.
- Wang, J., Conder, J. A., Blitzer, D. N., & Shinkareva, S. V. (2010). Neural representation of abstract and concrete concepts: A meta-analysis of neuroimaging studies. *Human brain mapping*, *31*(10), 1459-1468.
- Warrington, E. K., & McCarthy, R. A. (1983). Category specific access dysphasia. *Brain*, *106*, 869-878.
- Warrington, E. K., & McCarthy, R. A. (1987). Categories of knowledge. Further fractionation and an attempted integration. *Brain*, *110*, 1273-1296.
- Warrington, E. K., & Shallice, T. (1984). Category specific semantic impairments. *Brain*, *107*, 829-854.
- Waxman, S. R., & Markow, D. B. (1995). Words as invitations to form categories: Evidence from 12- to 13-month-old infants. *Cognitive psychology*, *29*(3), 257-302.
- Weisberg J, Turenout M, Martin A. (2007). A neural system for learning about object function. *Cerebral Cortex*. *17*, 513–521. doi: 10.1093/cercor/bhj176.
- Widmann, A., Gruber, T., Kujala, T., Tervaniemi, M., & Schröger, E. (2007). Binding symbols and sounds: evidence from event-related oscillatory gamma-band activity. *Cerebral Cortex*, *17*(11), 2696-2702.
- Willems, R.M., Hagoort, P., & Casasanto, D. (2010). Body-specific representations of action verbs: Neural evidence from right- and left-handers. *Psychological Science*, *21*, 67-74.
- Wittgenstein, L. (1953/2010). *Philosophical investigations*. John Wiley & Sons.
- Yap, S., Ji, L. & Hong, E. (2017). Culture, cognition, & Language. In J. T. Wixted & S. Thompson-Schill (Eds), *The Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience* (4th Edition, Volume 3: Language and Thought). New York: Wiley.
- Yee, E., & Thompson-Schill, S. L. (2016). Putting concepts into context. *Psychonomic Bulletin & Review*, *23*, 1015-1027.
- Yee, E., Chrysikou, E., Hoffman, E., & Thompson-Schill, S.L. (2013). Manual Experience Shapes Object Representation. *Psychological Science*, *24* (6), 909-919.
- Yee, E., Chrysikou, E.G., Thompson-Schill, S.L. (2013). Semantic Memory. In Kevin Ochsner and Stephen Kosslyn (Eds.), *The Oxford Handbook of Cognitive Neuroscience, Volume 1: Core Topics* (pp. 353-374). Oxford University Press.