Word Learning as the Confluence of Memory Mechanisms: Computational and Neural Evidence

Prahlad Gupta

In this article, I review evidence for the proposal that the processing and learning of words can usefully be understood as lying at the intersection of a variety of memory mechanisms. I begin with consideration of the temporally dynamic and serially ordered nature of human spoken language, focusing particularly on spoken word forms, and discuss the computational consequences of these properties and how they constrain the manner in which certain critical aspects of language are likely to be processed in the brain. In the second part, I discuss another fundamental property of language - its arbitrariness - and discuss how this once again is a functional characteristic that has important implications for how language must be processed in the mind/brain. Relevant neuroscientific evidence is briefly reviewed along with each of these discussions. The third part of the article brings together these ideas in a framework that integrates functional and computational considerations in word processing and word learning. I also discuss how this functionally and computationally derived proposal is quite consistent with other recent architectural suggestions derived from less computational and more neurophysiological points of view (further discussion on natural vocabulary acquisition is provided in de Groot, Volume 1, Chapter 23, and Ellis, Volume 2, Chapter 31). The present article thus provides an integration of various sources of evidence that bear on word-level processing and word learning.

Serial Ordering in Spoken Language

Spoken language is processed over time. Unlike written language, in which a unit such as a word is present throughout the process of reading, and is present in its

The Handbook of the Neuropsychology of Language, First Edition. Edited by Miriam Faust. © 2012 Blackwell Publishing Ltd. Published 2012 by Blackwell Publishing Ltd. entirety at the conclusion of a writing event, a spoken word is never present in its entirety during listening, nor is there any point during production when the spoken form is present in its entirety. At every point during spoken language processing (both listening and producing), all that is present as a stimulus is the currently spoken/heard piece of the speech stream.

This essential characteristic of spoken language is so obvious that it may appear uninteresting and/or inconsequential – emphasizing it might seem much like saying that cars run on wheels, which does not provide much insight into how cars work. In fact, however, the evanescence of spoken language has profound implications for how it must be processed in the mind/brain. A useful starting point in thinking about these implications is the fact that a novel word form, on first exposure, is a novel sequence of sounds. The task of repeating such a stimulus immediately after exposure to it requires the listener to encode the serial order of this sequence during its presentation, and then replicate this serial order when the stimulus is no longer present. That is, immediate repetition of a novel word form requires the encoding and retrieval of a novel serial ordering of constituent sounds. What are the mechanistic underpinnings of such a serial ordering task?

Twenty years ago, Jordan (1986) pointed out that, despite the fundamental importance of serially ordered action to human behavior, no general theory of serial ordering had emerged. Nor is any such generally accepted theory available today. However, many mathematically and/or computationally specified accounts have been developed of various aspects of serially ordered behavior (many of them in the context of the immediate serial recall task, e.g., Botvinick & Plaut, 2006; Brown, Preece, & Hulme, 2000; Burgess & Hitch, 1992, 1999; Gupta, 1996; Hartley & Houghton, 1996; Page & Norris, 1998). Although these accounts differ in numerous respects, they do all converge on one notion: in all of these accounts, producing a serially ordered sequence has the computational requirement of maintaining state or context information. For example, in order to replicate the sequence "BACDAB," a system must disambiguate the first and second occurrences of "A," so as to be able to produce "C" following the first "A," but "B" following the second "A." This requires maintenance of state or context information - information about where the system currently is in producing the sequence, and, in particular, information that distinguishes the state of "currently producing A" for the two instances of A. Thus in all computational accounts of serial ordering, the system must maintain some kind of state information.¹ That is, computational analysis indicates that a serial ordering task such as nonword repetition cannot be performed without a serial ordering mechanism, which in turn requires maintenance of state information.²

But maintenance of such state information is nothing if not *memory* for sequential information, as will be discussed in greater detail below. Computational analysis thus indicates that encoding and repeating a novel sequence requires some kind of serial ordering memory. This, then, provides a clear answer to our question of what is required computationally for repetition of a novel word form or nonword: it indicates the computational necessity in nonword repetition of *serial ordering*, and of a serial ordering *memory* mechanism. Importantly, this requirement follows directly from the temporally dynamic and serially ordered nature of spoken language – in this case, of word forms.

Over the last two decades, the relationship between phonological short-term memory (PSTM) and language processing (and especially phonological vocabulary learning - i.e., the learning of novel phonological word forms) has become a major focus of investigation in psychological research, generating extensive bodies of study in the traditional domains both of memory research and of language research (e.g., Dollaghan, 1987; Gathercole & Baddeley, 1989; Gathercole, Service, Hitch, Adams, & Martin, 1999; Gathercole, Willis, Emslie, & Baddeley, 1992; Gupta, MacWhinney, Feldman, & Sacco, 2003; Martin & Saffran, 1997; Martin, Saffran, & Dell, 1996; Montgomery, 2002; Saffran, 1990; for review, see Baddeley, Gathercole, & Papagno, 1998; Gathercole, 2006). Among the results that initiated these bodies of research were the findings that novel word repetition ability (i.e., the ability to immediately repeat possible but nonoccurring word forms, also termed nonwords) is correlated with immediate serial list recall ability on the one hand, and with vocabulary achievement on the other, in normally developing children (Gathercole & Baddeley, 1989) and in children with specific language impairment (SLI; Gathercole & Baddeley, 1990a). Since these initial reports, an overwhelming amount of further evidence has documented the existence of a relationship between vocabulary size and/or new word learning, nonword repetition, and immediate serial recall (e.g., Atkins & Baddeley, 1998; Baddeley, 1993; Baddeley, Papagno, & Vallar, 1988; Gathercole & Baddeley, 1990b; Gathercole et al., 1999; Gathercole et al., 1992; Gathercole, Hitch, Service, & Martin, 1997; Gupta, 2003; Gupta et al., 2003; Michas & Henry, 1994; Papagno, Valentine, & Baddeley, 1991; Papagno & Vallar, 1992; Service, 1992; Service & Kohonen, 1995).

What has remained unexplained, however, is the nature of the observed relationships between PSTM and the processing and learning of novel word forms. The computational analysis outlined above suggests a formal reason why the dynamic and serially ordered nature of spoken language (in this case, of spoken word forms) should necessitate a reliance on short-term memory. In recent work, Gupta and Tisdale (2009) concretized this analytic formulation in the form of a computational model. Gupta and Tisdale (2009) constructed a model that was exposed to word forms represented as input phonological sequences, and that attempted to repeat each word form immediately after presentation. The model incorporated the ability to learn from each such exposure. Over many exposures to many word forms, the model learned about the corpus of word forms to which it was being exposed, and thus acquired a phonological vocabulary. Gupta and Tisdale (2009) were then able to examine various aspects of the model's behavior and functioning, including: factors that affected its phonological vocabulary learning; its ability to repeat unlearned sequences (i.e., its nonword repetition) as well as factors that affected this ability; and, of greatest relevance for the present chapter, how the model instantiated PSTM.

The structure (*architecture*) of the model is shown in Figure 8.1a, and is an adaptation of an architecture introduced by Botvinick & Plaut (2006). The model has



Figure 8.1 (a) Architecture of the model. (b) Processing regimen in the model, illustrated for the word form *flugwish*.

an input layer at which a representation of an entire syllable is presented, and an output layer that uses the same representation scheme, at which the model's output is produced. The representation of a syllable, at both the input and the output layers, is in terms of a CCVCC (i.e., consonant-consonant-vowel-consonant-consonant) template. That is, a syllable is represented at the input layer across a set of units that are divided into five slots. Activation of units in the first slot denotes the first C (if any) of the syllable, activation of units in the second slot denotes the second C (if any) of the syllable, activation of units in the third slot denotes the V of the syllable, and so on. Within each of these slots, the various phonemes that are legal for that slot for English are represented as different patterns of activations across a set of units. For example, for the encoding scheme used, there are 17 different phonemes of English that are legal for the first C slot. These phonemes were represented as different patterns of activation across five units constituting the first C slot. Similarly, the 21 phonemes that are possible in the V slot were represented as different patterns of activation across a set of five units constituting the V slot; and so on for the various slots shown at the input and output layers in Figure 8.1a.

The model also has an intermediate layer of 200 units. Such an intermediate layer, which does not directly receive the model's input or directly produce the model's output, is usually termed a *hidden layer* in such models, and therefore the units it contains are typically termed *hidden units*, as shown in Figure 8.1a. All units in the input layer project to all units in the hidden layer, and all units in the hidden layer project to all units in the output layer, as is common in such connectionist models

of cognitive phenomena. An additional aspect of the architecture is the selfconnections on the hidden layer (indicated by the circular arrow from the hidden layer back to itself), which denote a connection from every unit in the hidden layer to itself and to every other unit in the hidden layer. These connections are termed *recurrent* connections. The model's task is to accept as input a *sequence* of syllables, and to produce as output a *sequence* (for this model, the same sequence) of syllables. It is well established that for connectionist models that perform sequential processing of this kind, the presence of recurrent connections is critical. Thus, the recurrent connections in the present model are crucial for it to be able to perform the task of inputting and repeating phonological word forms presented as sequences of syllables.

Figure 8.1b illustrates the regimen of presentation and desired output in the model, for the example word form *flugwish*. The procedure is the same irrespective of whether or not the word form has been presented to the model previously. Following presentation of the first syllable flug at the input, the model's task is to produce that same syllable at the output. The model's actual output may or may not be correct. Either way, after the model has produced an output, the activation pattern at the hidden layer is transmitted across the recurrent connections, thus transmitting information to *itself* that will arrive at the next time-step, so that when the second syllable wish is presented, the model's hidden layer actually receives input from two sources: the input representing wish, and input from its own previous state. When presented with this second syllable at the input, the model's task, as for the first syllable, is to produce the input syllable at the output. Again, the output may or may not be correct. Again, the hidden layer activation pattern is transmitted across the recurrent connections to be available at the next time-step. The input at this next step is actually an indication of the end of input, denoted by activation of the Recall unit in the input layer. At this point the model's task is to produce at the output layer the entire sequence of syllable representations previously presented at the input layer, i.e., flug followed by wish, and then activate the Stop unit at the output layer, to signify the end of production of the word form. As this repetition must be performed in the absence of any external input representing the word form, the network must necessarily have encoded some internal memory representation of the word form to allow it to now produce it in correct sequence (i.e., to perform immediate serial recall of the word form). At each point during recall, the model's hidden layer receives input from activation of the Recall unit, and from its own state at the previous time-step. (Note that the activation of the Recall unit is only a cue, and carries no information about the specific word form that was presented, because this same unit is activated as a cue for all word forms). Thus overall, the model attempts to match its own production (i.e., repetition) of a syllable sequence constituting a word form with the observed linguistic sequence provided by the environment. At the end of presentation and repetition of one word form, the model's connections weights are adjusted using a learning procedure for neural networks with recurrent connections that is known as back propagation through time, whose details are beyond the scope of this article (for

further discussion, see Botvinick & Plaut, 2006; Gupta & Tisdale, 2009; Rumelhart, Hinton, & Williams, 1986).

The essence of what this model does is to encode the serially ordered sequence of constituents comprising a word form, and then, after input has ended, to reproduce that serially ordered sequence, in the absence of any further informative input. That is, the model performs the task of serially ordered production of word forms. The model's serial ordering capability is critically dependent on the recurrent connections on the hidden layer; as described above, they provide the ability for the model to know where it is in producing the current word form, by providing information about what had already been produced. Gupta and Tisdale (2009) pointed out that this information is information about the past, and thus indubitably constitutes memory information, and that this information is overwritten when a subsequent word form sequence is produced, so that it is short-term memory information. They were also able to show that it can be regarded as phonological short-term memory information. Thus, the mechanisms in the model that provided this information provided PSTM functionality. This in turn indicated that PSTM functionality was crucial to serially ordered word form production in the model. Gupta and Tisdale (2009) also demonstrated that impairment of this functionality is severely disruptive to novel word from repetition as well as learning. The Gupta and Tisdale (2009) model thus provided a concrete demonstration of how the temporally dynamic and serially ordered nature of spoken language has implications for its processing in the mind/brain: it necessitates a reliance on short-term memory.

There is a fair amount of evidence regarding the neural substrates of such PSTM functionality, which appears to map onto an interactive neural system that is importantly dependent on temporoparietal cortex but that also encompasses anterior perisylvian regions. For instance, patients with posterior damage in general appear to suffer from span deficits (Risse, Rubens, & Jordan, 1984). Furthermore, in reviewing the neuropsychological syndrome of "pure STM" deficit, which involves reduced auditory-verbal short-term memory in the absence of other major language and cognitive deficits, Shallice and Vallar (1990) conclude that the condition reflects impairment to a short-term "input phonological store," and, "based on clinicalanatomical correlations, that the anatomical region compromised in this deficit is left inferior parietal cortex (angular gyrus and supramarginal gyrus). In patients with preserved span abilities, these areas seem quite consistently spared (Shallice & Vallar, 1990). This suggests that subsets of left temporoparietal cortex (e.g., left inferior parietal cortex) may be particularly specialized for the temporary storage of phonological information, and thus particularly crucial for verbal short-term memory. Supporting evidence that areas of temporoparietal cortex play a role in the temporary maintenance of information comes from single-cell recordings in primates, which showed memory-related planning activity in posterior parietal cortex (Gnadt & Anderson, 1988); and from a positron emission tomography (PET) study involving a verbal short-term memory task which revealed a supramarginal focus of activation, which the authors interpreted as the locus of phonological storage involved in the verbal short-term memory task (Paulesu, Frith, & Frackowiak,

1993). Furthermore, Baddeley et al. (1988) have described a patient, P. V., who has a pure short-term memory deficit. P. V. was able to learn meaningful paired associates in a familiar language. However, she was unable to learn to associate an unfamiliar word (in an unfamiliar language) with a familiar word in a familiar language, which is akin to learning a new vocabulary item. The fact that P. V. was a pure shortterm memory patient suggests that the critical damage in her case was to left inferior parietal cortex. This in turn indicates that this area of cortex does play a role in vocabulary acquisition. So, one part of the neural substrate commonly underlying PSTM and vocabulary acquisition appears to be left inferior parietal cortex. Beyond this, it appears that the encoding and retrieval of verbal sequences may be subserved by regions of inferior/posterior parietal cortex (especially BA 40, but also BA 39 and BA 7), while active maintenance and rehearsal are subserved by regions of (pre) frontal cortex, especially Broca's area (BA 44/45), premotor and supplementary motor cortex (BA 6), and dorsolateral prefrontal cortex (BA 9/46; Awh et al., 1996; Gupta & MacWhinney, 1997; Paulesu et al., 1993; Risse et al., 1984; Shallice & Vallar, 1990). However, other studies have failed to find inferior parietal activation during verbal short-term memory tasks (e.g., Chein & Fiez, 2001; Fiez et al., 1996; Grasby et al., 1993), and a clear consensus has not as yet emerged regarding the role of various neural regions in sub-aspects of PSTM tasks, although there is a fair consensus on the overall involvement of inferior/posterior parietal cortex.

The Arbitrariness of Language

Virtually every introductory textbook on language points out that language is arbitrary. That is, the mapping of phonological forms onto meanings does not follow any identifiable pattern within a language (much less across languages). For example, in English, the fact that the phonological form hat maps onto the meaning "something to wear on the head," cannot be taken to predict that the similar phonological form rat will map onto a meaning that is similar to that of hat. Morphology provides exceptions to this arbitrariness, so that, for example, the presence of an -s at the end of a noun of English does fairly consistently indicate plurality, and hence arguably something about meaning. There are also submorphological regularities, such as, for instance, similarities in meaning that are signaled by initial segment clusters such as sn-, as in sneer, snigger, and snide. Despite these exceptions, it is quite clear that the form-meaning mapping incorporates a high degree of arbitrariness in human languages; hence the definitional nature of this arbitrariness. Once again, this property of language is so fundamental that emphasizing it might appear to be uninformative. In fact, however, the arbitrariness of language is once again a functional characteristic that has important implications for how language must be processed in the mind/brain.

These implications arise from the integration of a number of ideas, which are worth clarifying here. The first of these ideas pertains to the distinction between systematic mappings and arbitrary mappings. A systematic mapping can be defined

as a function (in the mathematical sense of a transformation of a set of inputs into a corresponding set of outputs) in which inputs that are similar on some specifiable dimension are mapped to outputs that are similar on some specifiable dimension. An example of a systematic mapping is a function whose input is the orthographic representation of a word and whose output is the reversed spelling of the same word. In this mapping, the similar orthographic forms BUTTER and BETTER map onto the also similar orthographic forms RETTUB and RETTEB. As another example, the function mapping the length of a bar of mercury in a thermometer onto temperature is systematic, in that numerically similar lengths map onto numerically similar temperatures. An arbitrary mapping, in contrast, is a function in which inputs that are similar on some specifiable dimension are mapped to outputs that are not necessarily similar on any specifiable dimension. For example, the mapping between the names of countries and the names of their capital cities is arbitrary: phonologically similar country names (e.g., Canada, Panama) do not map onto capital city names that are consistently similar phonologically or on any other identifiable dimension (Ottawa, Panama City). As another example, the mapping between human proper names and the personality characteristics of those bearing them is arbitrary within a particular gender and culture. That is, the phonologically similar names John and Don do not map onto personality types that are more similar on any identifiable dimension than the personality types associated with the phonologically dissimilar names John and Fred. The property of arbitrariness discussed above for human languages is an instantiation of precisely this type of arbitrary mapping between the forms and meaning of words.

The second idea is that connectionist networks are devices that instantiate mappings. When an input is provided, such a network transforms the input stimulus into an output response, thus instantiating a mapping. The distinction between systematic and arbitrary mappings thus becomes relevant to such networks, and, in particular, to the nature of learning that can occur in connectionist networks whose input and output representations allow for measurement of similarity - i.e., which employ distributed representations at the input and output. The defining characteristic of such representations is that a stimulus is represented as a pattern of activation that is distributed across a pool of units, with each unit in the pool representing a feature that comprises the entity; there is no individual unit that represents the whole entity. The most important characteristic of distributed representations is that they enable similar stimuli to have similar representations. If such a connectionist system instantiates a systematic mapping, presentation of a novel input stimulus leads to production of a correct or close-to-correct output response simply by virtue of generalization based on prior knowledge: because the representations are distributed, the network will respond to the novel input in a manner that is similar to the response for previous similar inputs; because the mapping is systematic, this will be approximately the correct response. Little or no learning (adjustment of connection weights) is therefore needed for production of a correct response to a novel stimulus. Thus even though distributed connectionist networks incorporate incremental weight adjustment together with a slow learning rate (because fast learning rates can lead to unstable learning and/or interference with previously established weights – what McCloskey & Cohen, 1989, termed *catastrophic interference*), if the mapping that such a network instantiates is systematic, then learning the correct response to a novel input can be fast, requiring only a few exposures to the novel input–output pairing (because, even on first exposure, the response is close to correct).

The situation is different, however, where a mapping is arbitrary. In a distributed connectionist network that instantiates an arbitrary mapping, presentation of a novel input stimulus is unlikely to lead to production of a near-correct response: previous learning does not help, precisely because the mapping is arbitrary. Learning to produce the correct response will require considerable weight change. Therefore, because weight change is made only incrementally in a distributed connectionist network, learning a new input–output pairing in an arbitrary mapping can only occur gradually, over many exposures, at each of which the weights are adjusted slightly. However, the learning of arbitrary associations of items of information such as those comprising episodes and new facts can occur swiftly in humans, often within a single encounter, and without catastrophic interference. Gradual weight change in distributed connectionist networks thus cannot offer an account of such learning behavior. Such learning would, however, be possible in a connectionist system that employed orthogonal or localist representations (which do not overlap and hence do not interfere with each other) together with a faster learning rate.

These points suggest a functional requirement for two types of networks: one employing distributed representations that incorporates the desirable property of generalizing appropriately for novel inputs, which also enables it to quickly learn new entries in a systematic mapping; and one that employs orthogonal representations and a faster learning rate. McClelland, McNaughton, and O'Reilly (1995) proposed that these two functional requirements are indeed provided by the human brain, in the form of what have respectively been termed the procedural memory system and the declarative memory system. The procedural memory system, which provides for the learning and processing of motor, perceptual, and cognitive skills, is believed to be subserved by learning that occurs in nonhippocampal structures such as neocortex and the basal ganglia (e.g., Cohen & Squire, 1980; McClelland et al., 1995; Mishkin, Malamut, & Bachevalier, 1984; Squire, Knowlton, & Musen, 1993), and can be thought of as operating like distributed connectionist networks (Cohen & Squire, 1980; McClelland et al., 1995). The declarative memory system is believed to be subserved by the hippocampus and related medial temporal lobe structures (we will refer to this loosely as "the hippocampal system"); these structures provide for the initial encoding of memories involving arbitrary conjunctions, and also for their eventual consolidation and storage in neocortex (e.g., Cohen & Squire, 1980; Mishkin et al., 1984; Squire et al., 1993). It can be thought of as a system that converts distributed representations into localist non-overlapping ones, and swiftly establishes associations between such converted representations (Cohen & Eichenbaum, 1993; McClelland et al., 1995). That is, the hippocampal system performs fast learning, based on orthogonalized representations, thus constituting the second necessary type of network and providing a basis for the swift encoding of arbitrary associations of the kind that comprise episodic and factual information. Neocortex and the hippocampal system thus perform complementary learning functions, and these functions constitute the essence of procedural and declarative memory, respectively. McClelland et al. (1995) marshaled a variety of arguments and evidence to support these proposals. Their framework offers a means of reconciling the weaknesses of distributed connectionist networks with the human capacity for fast learning of arbitrary associations as well as with neurophysiological data.

It should be noted that different learning tasks are not viewed as being routed to one or other learning system by some controller based on whether each task is better suited to declarative or procedural learning. Rather, both learning systems are engaged in all learning behavior. However, for any given learning task, components of the task that constitute arbitrary mappings will be ineffectively acquired by the procedural system, and will only be effectively acquired by the declarative system, with later consolidation into the procedural system then being necessary. Any components of the task that constitute systematic mappings may be acquired by the declarative system but can also be effectively acquired directly by the procedural system, so that their declarative learning and later consolidation does not add much benefit. McClelland et al.'s (1995) framework has been widely influential, and constitutes the third idea that Gupta and Dell (1999; Gupta & Cohen, 2002) incorporated.

Gupta and Dell (1999; Gupta & Cohen, 2002) noted that phonology incorporates a systematic mapping, in that similar input phonology representations map onto similar output phonology representations; and that in contrast, the mapping between word forms and meanings is largely arbitrary as discussed above, with similar phonological word form representations not being guaranteed to map onto similar meanings. Furthermore, in humans, learning a new word can in general occur relatively rapidly, which implies that learning can occur relatively rapidly for both the systematic phonology of a novel word, and its links with semantics. Based on these observations and the assumption that the human lexical system employs distributed representations, Gupta and Dell (1999; Gupta & Cohen, 2002) suggested that the fast learning of new distributed representations of phonological word forms in the systematic input-output phonology mapping can be accomplished by a distributed connectionist-like procedural learning system even if it incorporates incremental weight adjustment. However, the swift establishment of the expressive and receptive links (i.e., learning associations between distributed phonological and semantic representations, which are in an arbitrary mapping) cannot be accomplished via incremental weight adjustment alone, and necessitates a computational mechanism employing orthogonal representations and a faster learning rate - i.e., a hippocampus-like system.

This hypothesis is consistent with the kinds of impairments observed in hippocampal amnesics. Such patients are virtually unable to learn new word meanings (e.g., Gabrieli, Cohen, & Corkin, 1988; Grossman, 1987), which is an indication of

their impairment in declarative memory. However, these same patients exhibit intact repetition priming for both known and novel words (e.g., Haist, Musen, & Squire, 1991), which is an indication of their relatively spared procedural memory. More recently, it has been reported that some children who suffered early damage to limited parts of the hippocampal system nevertheless achieve vocabulary levels in the low normal range by early adulthood (Vargha-Khadem et al., 1997). While the broader implications of this finding have been the matter of debate (Mishkin, Vargha-Khadem, & Gadian, 1998; Squire & Zola, 1998), the results are not inconsistent with the present hypothesis. They may indicate that not all parts of the hippocampal system are equally critical for learning of associations between word meanings and word forms, but remain consistent with the larger body of evidence indicating that parts of the hippocampal system are critical for normal learning of such associations (and for semantic memory more generally). Gupta and Dell's (1999; Gupta & Cohen, 2002) proposal regarding the differential engagement of procedural and declarative memory systems in word learning thus appears to be consistent with computational analysis of the requirements of word learning as well as with neuropsychological data. The distinction between the roles of procedural and declarative learning is similar to a view proposed by Ullman (2001, 2004), who also suggests that these two types of learning play specific roles in language learning, but who suggests that they underlie the distinction between syntax and the lexicon, rather than underlying different aspects of word learning.

Thus once again, computational analysis of fundamental properties of language (in this case, arbitrariness and systematicity) leads to constraints on how language must be processed in the mind/brain. Interestingly, we are once again led to a dependence of novel word processing/learning on *memory* systems.

Word Learning as a Confluence of Memory Systems: An Integrated View

The ideas outlined above can be seen as offering the beginnings of an integration of the domains of input phonology and output phonology in the lexical domain. As is clear from Figure 8.1a, input-side processing in the model incorporates a key element of what is required for spoken word recognition, in that a sequence of sublexical elements is transduced into an internal word form representation. The model also, of course, incorporates a key element of word production: an internal representation is transformed into a sequence of output phonological representations. In the model, the input and output phonology processes are tightly integrated, and are subserved by the same internal representation. In this sense, the model offers a tentative integration of the domains of spoken word recognition and spoken word production, or input and output phonology.

Because its learning algorithms constitute procedural learning, the model also incorporates Gupta and Dell's (1999; Gupta & Cohen, 2002) proposal that the learning of phonological forms is accomplished via procedural memory/learning.



Figure 8.2 Word learning as a confluence of memory systems.

Because the model does not simulate semantic processing or the linking of semantic with phonological representations, it does not speak to the second aspect of Gupta's proposal, viz. that the linking of word form representations to semantic representations requires declarative memory/learning. To the extent that the proposal is correct, however, we are led to a view of word learning in which the seemingly simple process of learning a new word is a rich confluence of short-term, procedural, and declarative memory systems. This view is sketched in Figure 8.2, which illustrates the interaction of PSTM, procedural learning, and declarative learning in providing for novel word learning. It should be clear that the implemented model of Gupta and Tisdale (2009) constitutes the phonological processing subset of this integrated architecture.

Neuroscientific Evidence for the Integrated View

We have already surveyed neuroscientific evidence relevant to the individual ideas presented above. It is important to note, however, that recent integrative views about the neural substrates of language processing/learning also fit quite well with the overall functional/computational scheme laid out above. Hickok and Poeppel (2007), for instance, have proposed two streams of speech-related processing, in which a ventral stream processes speech signals for comprehension, and a

dorsal stream maps acoustic speech signals to frontal lobe articulatory networks. In terms of Figure 8.2, the ventral stream can be seen as consisting of the pathway from input phonology to semantic representations, while the dorsal stream can be seen as that from input phonology to output phonology. This implies that phonological word form representations are a point of contact between the two pathways, which is entirely consistent with Hickok and Poeppel's formulation, in which a "phonological network" connects the two streams (Hickok & Poeppel, 2007; Figure 8.1).

Hickok and Poeppel's (2007) ventral stream includes what we have identified as the arbitrary mapping from phonology to semantics. Based on the computational analysis presented above, this would be expected to require involvement of the hippocampal system. Although Hickok and Poeppel (2007) do not explicitly discuss this, the physical location of the proposed ventral stream is entirely consistent with the engagement of subcortical medial temporal structures such as the hippocampal system. And indeed, Rodriguez-Fornells, Cunillera, Mestres-Misse, and de Diego-Balaguer (2009), building on Hickok and Poeppel's (2007) formulation, posit three pathways that are relevant to language learning: (1) a *dorsal audio-motor interface*, corresponding with Hickok and Poeppel's (2007) dorsal stream; (2) a *ventral meaning integration interface*, corresponding to Hickok and Poeppel's (2007) ventral stream; and (3) an *episodic-lexical interface*. Rodriguez-Fornells et al. (2009) explicitly characterize this third pathway as incorporating declarative learning structures such as the hippocampal system.

Rodriguez-Fornells et al. (2009) also posit an integrative role for the basal ganglia, noting their importance to sequential tasks. As we noted earlier, the basal ganglia and neocortex are regarded as important substrates that incorporate procedural learning (e.g., Cohen & Squire, 1980; McClelland et al., 1995; Mishkin et al., 1984; Squire et al., 1993), and can be thought of as operating like distributed connectionist networks (Cohen & Squire, 1980; McClelland et al., 1995). In the conceptualization shown in Figure 8.2, the operation of such procedural learning is pervasive throughout the system. Thus, consistently with the formulation of Rodriguez-Fornells et al. (2009), it does indeed play an integrative role. The two streams that Rodriguez-Fornells et al. (2009) add to Hickok and Poeppel's (2007) formulation can thus be seen as the complementary operation of declarative and procedural memory discussed in this article and depicted in Figure 8.2.

The present chapter's integrated view of word processing and learning as lying at the confluence of several memory systems is thus in considerable consonance with current thinking in integrative neurophysiology. Furthermore, the present formulation offers an analysis of the functional architecture that is derived from computational consideration of fundamental characteristics of language. The combination of such integrative approaches from behavioral, computational, and neurophysiological perspectives appears particularly well suited to furthering our understanding of language, and the present chapter has aimed to review evidence that word processing and word learning are domains of language that are well amenable to such integrative treatment.

Notes

- 1 Maintenance of state information is necessary, but not sufficient, for a system to be able to produce specific serially ordered sequences of output. There must also be a sequential control policy that specifies, among other things: (1) how the current state of the system is determined (it is usually here that the state information is required); (2) how the output is determined from the current state; and (3) how the output maps onto the sequential elements of the desired sequential behavior. For example, in accounts of serial ordering that employ a time-varying context signal (e.g., Burgess & Hitch, 1992, 1999; Brown et al., 2000; Gupta, 1996; Hartley & Houghton, 1996), the procedure that updates the context signal is an aspect of the sequential control policy, as is the specification of how the updated context signal is translated into an output representing a sequence element. In a simple recurrent (SRN; e.g., Elman, 1990), the procedure that merges the "context layer" activation with the current input is part of the sequential control policy, as is the specification of the target sequence element to be produced at each point in time. Thus having and executing a sequential control policy, together with maintenance of state information, are the necessary and sufficient conditions for serially ordered sequential behavior. For the present discussion, the most directly relevant of these requirements is the maintenance of state information, and we therefore do not provide further analysis of the sequential control policy. It may be noted, however, that such analysis would be closely related to the theory of formal languages, automata, and computation (Hopcroft & Ullman, 1979).
- 2 In some accounts, the maintenance of state information is more implicit, but is nevertheless critical to the production of a serially ordered sequence. For instance, in some models, the serial ordering is encoded directly by the structure of the representations such as linguistic syntactic frames or slot-filler representations (Dell, 1986; Dell, Schwartz, Martin, Saffran, & Gagnon, 1997; Kuehne, Gentner, & Forbus, 2000; Levelt, Roelofs, & Meyer, 1999; Warker & Dell, 2006) or storage structures such as arrays (e.g. Nairne, 1990; Nairne & Neumann, 1993) themselves. In such accounts, it is assumed that some process unpacks this structurally encoded serial ordering (e.g., Dell et al., 1997). But this requires maintenance of context: During sequence production, such a process (which instantiates the sequential control policy) would need to keep track of which structural element of the frame or array was currently being accessed, which is functionally identical to maintaining state or context information.

References

- Atkins, P. W. B., & Baddeley, A. D. (1998). Working memory and distributed vocabulary learning. *Applied Psycholinguistics*, 19, 537–552.
- Awh, E., Jonides, J., Smith, E. E., Schumacher, E. H., Koeppe, R. A., & Katz, S. (1996). Dissociation of storage and rehearsal in verbal working memory: Evidence from positron emission tomography. *Psychological Science*, 7, 25–31.
- Baddeley, A. D. (1993). Short-term phonological memory and long-term learning: A single case study. *European Journal of Cognitive Psychology*, 5, 129–148.
- Baddeley, A. D., Gathercole, S. E., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, 105, 158–173.

- Baddeley, A. D., Papagno, C., & Vallar, G. (1988). When long-term learning depends on short-term storage. *Journal of Memory and Language*, 27, 586–595.
- Botvinick, M., & Plaut, D. (2006). Short-term memory for serial order: A recurrent neural network model. *Psychological Review*, 113, 201–233.
- Brown, G. D. A., Preece, T., & Hulme, C. (2000). Oscillator-based memory for serial order. *Psychological Review*, 107, 127–181.
- Burgess, N., & Hitch, G. J. (1992). Toward a network model of the articulatory loop. *Journal of Memory and Language*, 31, 429–460.
- Burgess, N., & Hitch, G. J. (1999). Memory for serial order: A network model of the phonological loop and its timing. *Psychological Review*, 106, 551–581.
- Chein, J. M., & Fiez, J. A. (2001). Dissociation of verbal working memory system components using a delayed serial recall task. *Cerebral Cortex*, 11, 1003–1014.
- Cohen, N. J., & Eichenbaum, H. (1993). Memory, amnesia, and the hippocampal system. Cambridge, MA: MIT Press.
- 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.
- Dell, G. S. (1986). A spreading activation theory of retrieval in sentence production. *Psychological Review*, 93, 283–321.
- Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical access in normal and aphasic speakers. *Psychological Review*, 104, 801–838.
- Dollaghan, C. (1987). Fast mapping in normal and language-impaired children. Journal of Speech and Hearing Disorders, 52, 218-222.
- Elman, J. L. (1990). Finding structure in time. Cognitive Science, 14, 179-211.
- Fiez, J. A., Raife, E. A., Balota, D. A., Scwarz, J. P., Raichle, M. E., & Petersen, S. E. (1996). A positron emission tomography study of the short-term maintenance of verbal information. *Journal of Neuroscience*, *16*, 808–822.
- Gabrieli, J. D. E., Cohen, N. J., & Corkin, S. (1988). The impaired learning of semantic knowledge following bilateral medial temporal-lobe resection. *Brain*, 7, 157–177.
- Gathercole, S. E. (2006). Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics*, *27*, 513–543.
- Gathercole, S. E., & Baddeley, A. D. (1989). Evaluation of the role of phonological STM in the development of vocabulary in children: A longitudinal study. *Journal of Memory and Language*, 28, 200–213.
- Gathercole, S. E., & Baddeley, A. D. (1990a). Phonological memory deficits in languagedisordered children: Is there a causal connection? *Journal of Memory and Language*, 29, 336–360.
- Gathercole, S. E., & Baddeley, A. D. (1990b). The role of phonological memory in vocabulary acquisition: A study of young children learning arbitrary names of toys. *British Journal of Psychology*, *81*, 439–454.
- Gathercole, S. E., Hitch, G. J., Service, E., & Martin, A. J. (1997). Short-term memory and new word learning in children. *Developmental Psychology*, 33, 966–979.
- Gathercole, S. E., Service, E., Hitch, G. J., Adams, A.-M., & Martin, A. J. (1999). Phonological short-term memory and vocabulary development: Further evidence on the nature of the relationship. *Applied Cognitive Psychology*, *13*, 65–77.
- Gathercole, S. E., Willis, C., Emslie, H., & Baddeley, A. D. (1992). Phonological memory and vocabulary development during the early school years: A longitudinal study. *Developmental Psychology*, 28, 887–898.

- Gnadt, J. W., & Anderson, R. A. (1988). Memory related motor planning activity in posterior parietal cortex of macaque. *Experimental Brain Research*, *70*, 216–220.
- Grasby, P. M., Frith, C. D., Friston, K. J., Bench, C., Frackowiak, R. S. J., & Dolan, R. J. (1993). Functional mapping of brain areas implicated in auditory-verbal memory function. *Brain*, 116, 1–20.
- Grossman, M. (1987). Lexical acquisition in Alcoholic Korsakoff psychosis. Cortex, 23, 631-644.
- Gupta, P. (1996). Word learning and verbal short-term memory: A computational account.In G. W. Cottrell (Ed.), *Proceedings of the eighteenth annual meeting of the Cognitive Science Society* (pp. 189–194). Mahwah, NJ: Erlbaum.
- Gupta, P. (2003). Examining the relationship between word learning, nonword repetition, and immediate serial recall in adults. *Quarterly Journal of Experimental Psychology (A)*, 56, 1213–1236.
- Gupta, P., & Cohen, N. J. (2002). Theoretical and computational analysis of skill learning, repetition priming, and procedural memory. *Psychological Review*, *109*, 401–448.
- Gupta, P., & Dell, G. S. (1999). The emergence of language from serial order and procedural memory. In B. MacWhinney (Ed.), *The emergence of language*, 28th Carnegie Mellon Symposium on Cognition. Mahwah, NJ: Erlbaum.
- Gupta, P., & MacWhinney, B. (1997). Vocabulary acquisition and verbal short-term memory: Computational and neural bases. *Brain and Language*, 59, 267–333.
- Gupta, P., MacWhinney, B., Feldman, H., & Sacco, K. (2003). Phonological memory and vocabulary learning in children with focal lesions. *Brain and Language*, *87*, 241–252.
- Gupta, P., & Tisdale, J. (2009). Does phonological short-term memory causally determine vocabulary learning? Toward a computational resolution of the debate. *Journal of Memory and Language*, 61, 481–502.
- Haist, F., Musen, G., & Squire, L. R. (1991). Intact priming of words and nonwords in amnesia. *Psychobiology*, 19, 275–285.
- Hartley, T., & Houghton, G. (1996). A linguistically constrained model of short-term memory for nonwords. *Journal of Memory and Language*, *35*, 1–31.
- Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, *8*, 393–402.
- Hopcroft, J. E., & Ullman, J. D. (1979). *Introduction to automata theory, languages, and computation*. Redwood City, CA: Addison-Wesley.
- Jordan, M. I. (1986). Serial order: A parallel distributed processing approach (Report 8604). La Jolla, CA: Institute for Cognitive Science, University of California, San Diego.
- Kuehne, S. E., Gentner, D., & Forbus, K. D. (2000). Modeling infant learning via symbolic structural alignment. In *Proceedings of the twenty-second annual conference of the Cognitive Science Society* (pp. 286–291). Mahwah, NJ: Erlbaum.
- Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22, 1–75.
- Martin, N., & Saffran, E. (1997). Language and auditory-verbal short-term memory impairments: Evidence for common underlying processes. *Cognitive Neuropsychology*, 14, 641–682.
- Martin, N., Saffran, E. M., & Dell, G. S. (1996). Recovery in deep dysphasia: Evidence for a relation between auditory-verbal STM capacity and lexical errors in repetition. *Brain and Language*, 52, 83–113.
- McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and

failures of connectionist models of learning and memory. *Psychological Review*, 102, 419–457.

- McCloskey, M., & Cohen, N. J. (1989). Catastrophic interference in connectionist networks: The sequential learning problem. In G. H. Bower (Ed.), *The psychology of learning and motivation*, (vol. 24, pp. 109–165). New York: Academic Press.
- Michas, I. C., & Henry, L. A. (1994). The link between phonological memory and vocabulary acquisition. *British Journal of Developmental Psychology*, *12*, 147–164.
- Mishkin, M., Malamut, B., & Bachevalier, J. (1984). Memories and habits: Two neural systems. In G. Lynch, J. McGaugh, & N. Weinberger (Eds.), *Neurobiology of learning and memory* (pp. 65–77). New York: Guilford.
- Mishkin, M., Vargha-Khadem, F., & Gadian, D. G. (1998). Amnesia and the organization of the hippocampal system. *Hippocampus*, *8*, 212–216.
- Montgomery, J. W. (2002). Understanding the language difficulties of children with specific language impairments: Does verbal working memory matter? *American Journal of Speech-Language Pathology*, 11, 77–91.
- Nairne, J. S. (1990). A feature model of immediate memory. *Memory and Cognition*, 18, 251–269.
- Nairne, J. S., & Neumann, C. (1993). Enhancing effects of similarity on long-term memory for order. Journal of Experimental Psychology: Learning, Memory, and Cognition, 19, 329–337.
- Page, M. P. A., & Norris, D. (1998). The primacy model: A new model of immediate serial recall. *Psychological Review*, 105, 761–781.
- Papagno, C., Valentine, T., & Baddeley, A. D. (1991). Phonological short-term memory and foreign-language learning. *Journal of Memory and Language*, 30, 331–347.
- Papagno, C., & Vallar, G. (1992). Phonological short-term memory and the learning of novel words: The effects of phonological similarity and item length. *Quarterly Journal of Experimental Psychology*, 44A, 47–67.
- Paulesu, E., Frith, C. D., & Frackowiak, R. S. J. (1993). The neural correlates of the verbal component of working memory. *Nature*, *362*, 342–345.
- Risse, G. L., Rubens, A. B., & Jordan, L. S. (1984). Disturbance of long-term memory in aphasic patients: A comparison of anterior and posterior lesions. *Brain*, 107, 605–617.
- Rodriguez-Fornells, A., Cunillera, T., Mestres-Misse, A., & de Diego-Balaguer, R. (2009). Neurophysiological mechanisms involved in language learning in adults. *Philosophical Transactions of the Royal Society B*, *364*, 3711–3735.
- Rumelhart, D., Hinton, G., & Williams, R. (1986). Learning internal representations by error propagation. In D. Rumelhart & J. McClelland (Eds.), *Parallel distributed processing* (vol. 1, *Foundations*, pp. 318–362). Cambridge, MA: MIT Press.
- Saffran, E. M. (1990). Short-term memory impairment and language processing. In A. Caramazza (Ed.), *Cognitive neuropsychology and neurolinguistics*. Mahwah, NJ: Erlbaum.
- Service, E. (1992). Phonology, working memory, and foreign-language learning. Quarterly Journal of Experimental Psychology, 45A, 21–50.
- Service, E., & Kohonen, V. (1995). Is the relation between phonological memory and foreign language learning accounted for by vocabulary acquisition? *Applied Psycholinguistics*, *16*, 155–172.
- Shallice, T., & Vallar, G. (1990). The impairment of auditory-verbal short-term storage. In G. Vallar & T. Shallice (Eds.), *Neuropsychological impairments of short-term memory*. Cambridge: Cambridge University Press.

- Squire, L. R., Knowlton, B., & Musen, G. (1993). The structure and organization of memory. Annual Review of Psychology, 44, 453–495.
- Squire, L. R., & Zola, S. M. (1998). Episodic memory, semantic memory, and amnesia. *Hippocampus*, 8, 205–211.
- Ullman, M. T. (2001). The declarative/procedural model of lexicon and grammar. *Journal of Psycholinguistic Research*, 30, 37–69.
- Ullman, M. T. (2004). Contributions of memory circuits to language: The declarative/ procedural model. *Cognition*, 92, 231–270.
- Vargha-Khadem, F., Gadian, D. G., Watkins, K. E., Connely, A., van Paesschen, W., & Mishkin, M. (1997). Differential effects of early hippocampal pathology on episodic and semantic memory. *Science*, 277, 376–380.
- Warker, J. A., & Dell, G. S. (2006). Speech errors reflect newly learned phonotactic constraints. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 387–398.

1 th