# Word learning, phonological short-term memory, phonotactic probability and long-term memory: towards an integrated framework

## Prahlad Gupta* and Jamie Tisdale

*Department of Psychology, University of Iowa, Iowa City, IA 52242, USA*

Word learning is studied in a multitude of ways, and it is often not clear what the relationship is between different phenomena. In this article, we begin by outlining a very simple functional framework that despite its simplicity can serve as a useful organizing scheme for thinking about various types of studies of word learning. We then review a number of themes that in recent years have emerged as important topics in the study of word learning, and relate them to the functional framework, noting nevertheless that these topics have tended to be somewhat separate areas of study. In the third part of the article, we describe a recent computational model and discuss how it offers a framework that can integrate and relate these various topics in word learning to each other. We conclude that issues that have typically been studied as separate topics can perhaps more fruitfully be thought of as closely integrated, with the present framework offering several suggestions about the nature of such integration.

**Keywords:** word learning; integrative framework; phonological short-term memory; phonological learning; short-term and long-term memory interactions; computational modelling

Although human children and adults learn new words effortlessly, understanding how they do so is a complex endeavour. Although it seems clear that word learning or vocabulary acquisition entails learning a word form, a meaning and the link between them (e.g. Saussure 1916; Desrochers & Begg 1987), this simple formulation encompasses a multitude of different abilities and subcomponents. As a result, the terms 'vocabulary acquisition' and 'word learning' have been used to mean a wide variety of different things, as revealed by consideration of the many different facets of what has been studied as word learning (for review, see Gupta 2005b).

In this article, we begin by outlining a very simple functional framework that despite its simplicity can serve as a useful organizing scheme for thinking about various types of studies of word learning. In §2, we review a number of themes that in recent years have emerged as important topics in the study of word learning, and relate them to the functional framework, noting nevertheless that these topics have tended to be somewhat separate areas of study. In the third part of the article, we describe a recent computational model developed by Gupta & Tisdale (2009), and discuss how it offers a framework that can integrate and relate these various topics to each other. We conclude that issues that have typically been studied as separate topics can perhaps more fruitfully be thought of as closely integrated, with the

present framework offering several suggestions about the nature of such integration.

## 1. FUNCTIONAL ASPECTS OF WORD LEARNING

In considering the functional aspects of word learning, it is useful to begin with some terminology. First of all, we will reserve the term vocabulary acquisition to refer to the overall phenomenon of acquiring a vocabulary of words over an extended period of time. We will reserve the term word learning to refer to the specific phenomenon of learning one or a small number of new words in a relatively short period of time. Thus, vocabulary acquisition is the cumulative outcome of multiple instances of word learning.

We next need to specify what we mean by *word form*, *meaning* and *link*. We will operationalize word form as an auditorily experienced human speech stimulus; we are thus discussing spoken, not written language. We use the term word form rather than *word* to emphasize the fact that such a stimulus may be a known word or a novel word (to a particular language learner/user). We assume the existence (in the cognitive system of such a user) of an internal phonological representation for any word form that is processed by the system. Although the nature of such an internal representation is likely to differ for known versus novel word forms, there must nevertheless be some internal representation that is evoked even by a novel word form; we therefore adopt the neutral term *word-form representation*.

What is the meaning of a word form? We will sidestep the immense amount of debate that has centred on this question, and simply operationalize the

* Author for correspondence (prahlad-gupta@uiowa.edu).

notion of meaning as an internal mental representation of an object(s), action(s), event(s) or abstract entity/entities. We term this internal representation a *semantic representation*. Such a representation can exist independently of a name. The meaning of a word form to a particular individual (rather than to a community of individuals) is then simply the semantic representation that is evoked via activation of the internal word-form representation of that word form.

Finally, what is the link between a word form and a meaning? We will operationalize the notion of link as a connection between representations, whose existence allows the one representation to activate the other. There are actually two links that must be considered: one from the word-form representation to the semantic representation (the *receptive link*); and one from the semantic representation to the word-form representation (the *expressive link*).

The view just laid out is depicted in figure 1 and should be relatively uncontroversial. It follows from this view that achieving full mastery of a new word requires creation of a word-form representation (presumably richer than that evoked by a novel word form), creation of a semantic representation, creation of an expressive link and creation of a receptive link. Although this, too, is presumably uncontroversial, framing the discussion in these terms serves an important purpose: it emphasizes that the question of how words are learned requires investigation of each of these aspects of mastery. Learning a word-form representation, learning a semantic representation, learning an expressive link and learning a receptive link can thus be thought of as core functional components of word learning.

As this framework provides a functional description of necessary aspects of word learning, it can serve as a simple way to classify a great deal of research on word learning. For instance, one major thrust of research in word learning can be characterized as investigating how the meanings of new words are inferred—that is, how the semantic representations are created. When exposed to a new word form in conjunction with a number of possible referents, how does a language learner infer the correct referent, if it is not explicitly indicated? Even if the referent can be determined, what is it? If it is an object, is it an animal, a bird or a machine? If it is an action, is it like running, or is it something else? If it is more abstract: what exactly is being denoted when the language learner's parents talk about the 'mortgage'? Is the entity the sole possessor of the name, or is it an example of the kinds of things so named? What aspect of the entity is being denoted? (e.g. Quine 1960)—the entity itself, all of it, some part of it and if so, which part, the colour, the shape or the substance? And, once a word has been learned in the sense of the language user being reliably able to pick out the correct referent when provided with the word form, how is the word's reference extended?—that is, how does the learner determine to which *other* objects that name applies, thereby inferring a *category*? A great deal of research can be seen as having addressed such questions. For instance, numerous studies have investigated the process whereby children extend the novel words to untrained
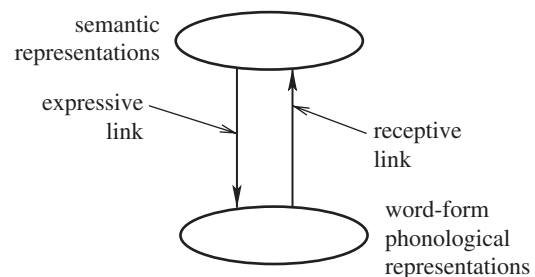


Figure 1. Functional aspects of word learning.

referents (Nelson & Bonvillian 1973; Oviatt 1982; Samuelson & Smith 1999). Others have examined the rate of acquisition as a function of the child's referential or non-referential orientation (e.g. Leonard *et al.* 1981). Some studies have addressed the complex issue of how meaning is inferred from context of usage (social, linguistic non-syntactic or syntactic) rather than from perceptual information, as it necessarily must be for abstract referents (see Sternberg 1987 for review). The extensive literature on constraints or biases or tendencies in children's word learning can also be seen as having addressed the issue of how the learner figures out what the denoted entity is (e.g. Mervis & Pani 1980; Markman 1984, 1989; Markman & Hutchinson 1984; Mervis 1984; Baldwin & Markman 1989; Merriman & Bowman 1989; Au & Glusman 1990; Merriman & Schuster 1991; Mervis *et al.* 1994; Hollich *et al.* 2000*a*). Thus, one major focus of research on word learning can be seen as investigating creation of the semantic representation component of the overall word-learning task.

A second major thrust of research in word learning can be characterized as investigating the learning of the word forms themselves. This research typically studies situations where word forms are presented in isolation to a learner who already knows the phonological structure of the language, thereby side-stepping issues of the development of phonology, segmentation, attention and identification. A good deal of early research of this type focused on the learning of lists of nonsense syllables in the tradition of Ebbinghaus. This research has shown that such learning is strongly influenced by the associational value of the word forms to be learned (e.g. Underwood & Schulz 1960). A few studies have required subjects to learn auditorily presented word forms as pre-training for a subsequent process in which they were mapped to semantics (Horowitz & Gordon 1972; Pressley & Levin 1981). There has been some examination of the effect of the phonological composition of the novel word on its learning (e.g. Leonard *et al.* 1981). Some recent studies have examined whether newly learned word forms exhibit the same kinds of cohort effects and neighbourhood density effects as well-known words, and have found evidence that they do (e.g. Magnuson Tanenhaus Aslin & Dahan 2003).

A third line of experimental research in word learning has involved provision of both semantics and word form to the learner, requiring that both be learned. In such investigation, creation of the expressive and receptive links becomes important to performance of the task. These studies have examined both childrens' and

adults' learning of novel word forms. These novel forms have included low-frequency real words, phonologically permissible non-words in the subject's native language and items drawn from foreign languages. An important dimension along which these studies can be categorized is according to the *direction of mapping* they investigated. In studies of *receptive word learning*, the emphasis has been on the ability to produce some indication of the meaning of the target word (such as a definition of its meaning or identification of its referent), when presented with the word form as a cue (e.g. Paivio & Desrochers 1981; Desrochers & Begg 1987; Pressley *et al.* 1987; Mandel *et al.* 1995; Tincoff & Jusczyk 1999); in these studies, creation of the receptive link is likely to have been more important to the task than creation of the expressive link. Conversely, in studies of *expressive word learning*, the emphasis has been on the ability to produce the target word form, when cued with some aspect of its meaning, or with its referent (e.g. Pressley *et al.* 1980; Ellis & Beaton 1993; Gallimore *et al.* 1977); here, creation of the expressive link is likely to have been more important than the receptive link. A number of studies have examined both directions of mapping (e.g. Horowitz & Gordon 1972; Ellis & Beaton 1993; Storkel 2001).

## 2. RECENT FOCI OF INVESTIGATION

While the functional framework outlined above is useful as a means of thinking about broad categories of research in word learning, in recent years, considerable attention has been given to a number of issues that cut across these functional aspects.

### (a) *Phonotactic probability and neighbourhood density*

One such issue relates to *phonotactic probability*, a construct that refers to the frequency with which various phonological segments and segment sequences occur in a language (e.g. Jusczyk *et al.* 1994). For example, the sequence /tr/ occurs more frequently than the sequence /fr/ at the beginning of words of English, and thus has a higher word-initial phonotactic probability in English. The notion of phonotactic probability can also be applied to word forms as a whole, as a measure of how closely the segments and segment sequences within a word form conform to the distribution of those segments and segment sequences in the language. For example, the word form *cat* incorporates the segments /k/, /æ/ and /t/, and also incorporates the biphones /kæ/ and /æt/. The higher the frequency of occurrence of these segments in first, second and third position, respectively, in words of English, and the higher the frequency of the biphones in words of English, the higher the phonotactic probability of cat. Thus, phonotactic probability can be seen as one means of gauging the similarity of a target word form to words of the language in general. The construct of *neighbourhood density* is a different gauge of the similarity of a word form to words of the language. The neighbourhood of a target word form is the set of words of language that are within some specified criterion of similarity to it. The criterion of similarity that has

most often been used is a one-phoneme distance, so that the neighbourhood of the word form cat is the set of all words of the language that differ from it by one phoneme (e.g. *mat*, *cot*, *cab*, . . .). Thus a word form that has many neighbours (i.e. whose neighbourhood has *high density*) is closely similar to many words of the language, whereas a word form that has relatively few neighbours (i.e. whose neighbourhood has *low density*) is closely similar to fewer words of the language.

Although the constructs of phonotactic probability and neighbourhood density originate in the study of spoken word recognition, there has been increasing interest in what role they may play in word learning. For instance, Storkel & Rogers (2000) and Storkel (2001) found that both expressive and receptive word learning is better for word forms of higher phonotactic probability, in children aged 3 through 13 years. In these studies, the referents of the words were objects, so that these were noun-learning tasks; in a subsequent study, Storkel (2003) obtained similar effects of phonotactic probability in a verb-learning task. In a study that examined learning of novel word forms alone (word-form learning) as well as learning of novel word forms paired with referents (word learning), Heisler (2005) found facilitative effects of both higher phonotactic probability and higher neighbourhood density. Retrospective analyses indicate that higher neighbourhood density is associated with an earlier age of acquisition, once again suggesting that neighbourhood density facilitates word learning, at least for certain types of words (Coady & Aslin 2003; Storkel 2004). Hollich *et al.* (2000b, 2002) and Swingley & Aslin (2007) experimentally examined the effect of neighbourhood density on word learning in toddlers, finding that this variable did affect word learning, although in these studies the effect was found to be reversed, so that learning appeared to be better for word forms from low- rather than high-density neighbourhoods. Samuelson *et al.* (in preparation) found, however, that higher neighbourhood density was facilitatory in toddlers' word learning, and offered an explanation for the seemingly contradictory findings across studies, showing computationally that when the high-density neighbourhood has a single dominant neighbour, the low-density word form is learned better, as in Swingley & Aslin (2007), but when the high-density neighbourhood has several neighbours, the high-density word form is learned better.

Thus, there is now considerable evidence that phonotactic probability and neighbourhood density affect word learning. An obvious question relates to what the locus of these effects is. In terms of the functional framework discussed above (figure 1), do these variables affect learning of the word-form representation, the semantic representation or the links? On the one hand, as these are phonological variables, it appears likely that their effects are related to phonological word-form learning. On the other hand, the effects appear to apply to both expressive and receptive word learning, which might suggest additional loci of influence. More generally, whatever their locus, it has remained unclear what the mechanistic bases

might be of these effects (but see Samuelson *et al.* (in preparation) for a preliminary computational account of both facilitatory and inhibitory neighbourhood density effects in receptive word learning).

## (b) *Phonological short-term memory*

A second issue that has received considerable attention in recent years is the nature of the associations that are obtained between word learning and phonological short-term memory (PSTM; for review, see Baddeley *et al.* 1998; Gathercole 2006). Evidence for such a relationship has come from observation of correlations between immediate serial memory ability, non-word repetition ability and vocabulary size, in a variety of developmental and neurologically impaired populations, as well as from experimental investigation of factors affecting verbal short-term memory and the learning of new words (Baddeley *et al.* 1988; Gathercole & Baddeley 1989, 1990; Papagno *et al.* 1991; Gathercole *et al.* 1992, 1997; Papagno & Vallar 1992; Service 1992; Baddeley 1993; Michas & Henry 1994; Service & Kohonen 1995; Atkins & Baddeley 1998; Gupta, 2003, 2005a; Gupta *et al.* 2003). An emerging view of this relationship is that immediate serial recall and non-word repetition are both tasks that draw on the mechanisms of PSTM fairly directly, and that the learning of new words is also in some way supported by PSTM (e.g. Brown & Hulme 1996; Baddeley *et al.* 1998; Gathercole *et al.* 1999; Gupta 2003). Interpreting this in terms of our functional framework, it is thought that the hypothesized role of PSTM occurs in the learning of word-form representations (e.g. Baddeley *et al.* 1998). However, although there are now approximately 200 published studies that examine the relationship between PSTM and non-word repetition and/or vocabulary learning, the precise nature of this relationship has remained unclear.

## (c) *Vocabulary size*

A third issue relates to the role of vocabulary size in word learning. The rationale for expecting such a role is that a learner with a vocabulary size of, say, 10 000 words can be expected to have considerably greater phonological knowledge than a learner who has a vocabulary of, say, 2000 words, and that this greater knowledge should facilitate learning further new words; thus, greater vocabulary size should lead to better word learning, providing a sort of positive feedback loop. The posited facilitatory role of vocabulary size has been especially prominent in discussion of the repetition (rather than learning) of novel word forms (i.e. in discussion of non-word repetition), where it has contrasted with the view described above, namely, that non-word repetition is critically dependent on PSTM. The non-word repetition task has come to be regarded as highly relevant to studying word learning because, as noted in the previous section, it is robustly predictive of word learning and vocabulary achievement. The mechanisms underlying non-word repetition ability thus become important as candidate mechanisms of word learning. A number of studies have shown that repetition accuracy is better for non-words of high rather than low

phonotactic probability or wordlikeness (e.g. Gathercole 1995; Munson 2001), indicating that long-term phonological knowledge affects non-word repetition performance—which in turn suggests that long-term knowledge (or vocabulary size) may be a determinant of word learning itself (e.g. Munson 2001; Edwards *et al.* 2004). Munson (2001; Munson *et al.* 2005) also found that the phonotactic probability effect in non-word repetition decreased with increasing vocabulary size, which could indicate that non-word repetition is dependent on vocabulary size, rather than on PSTM. The question of what role prior vocabulary knowledge plays in word learning thus makes contact with both of the first two issues identified above. It is related to the first issue in that any effect of the phonotactic probability of a new word is an effect based on the relation of that word form to prior vocabulary knowledge. It is related to the second issue in that it has been proposed as a determinant of non-word repetition and word learning that stands in contrast to the posited role of PSTM. There has been relatively little discussion of the likely locus of the effect of vocabulary size on word learning, although it appears likely that the phonotactic effects noted above occur at the level of word-form representations. In addition, as with the previously discussed themes, it has remained unclear by what process vocabulary size might have an effect on word learning and/or non-word repetition.

## (d) *Long-term memory systems*

A fourth issue of relatively recent focus relates to the role of long-term memory systems in word learning. A number of streams of thinking can be identified here.

## (e) *Procedural and declarative memory*

In one such stream, it has been suggested that different components of learning a new word rely differentially on what have been termed *procedural memory* and *declarative memory* (see Gupta & Dell 1999; Gupta & Cohen 2002 for more detailed discussion). Gupta & Dell's (1999; Gupta & Cohen 2002) proposal is based on 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 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 that 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 a distributed connectionist system instantiates a systematic mapping, presentation of a novel input stimulus leads to the 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 the 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 the 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 behaviour. 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 a slow learning rate and distributed representations that incorporate 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 *et al.* (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 non-hippocampal structures such as neocortex and the basal ganglia (e.g. Cohen & Squire 1980; Mishkin *et al.* 1984; Squire *et al.* 1993; McClelland *et al.* 1995), 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 the 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) marshalled 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 behaviour. 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 & Dell (1999; Gupta & Cohen 2002) incorporated.

Gupta & Dell (1999; Gupta & Cohen 2002) noted that phonology incorporates a systematic mapping, in that similar input phonology representations map onto similar output phonology representations. In contrast, the mapping between phonological word forms and meanings is largely arbitrary (with morphophonology as an exception), in that similar phonological word form representations are not guaranteed to map onto similar meanings, and indeed, this arbitrariness is generally viewed as one of the defining characteristics of human language. Gupta & Dell (1999; Gupta & Cohen 2002) further noted that humans can in general learn a new word 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 & 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 *et al.* 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 *et al.* 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 & 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 neurospychological 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.

### (f) Interaction of long-term and short-term memory in the verbal domain

Another stream relating word learning to long-term memory is concerned with how long-term and short-term memory systems interact in the verbal domain, with a recent edited volume having been devoted to exploring this topic (Thorn & Page 2009). This issue can be seen as a generalization of the second and third issues identified above, in that it encompasses the notion that *both* PSTM and long-term prior knowledge are involved in and may interact in service of word learning, while also extending the notion of such interactions into other verbal domains such as, for instance, list recall and list learning (see also Gathercole (1995, 2006), Gupta (1995, 1996), Gupta & MacWhinney (1997) and Gupta *et al.* (2005) for similar views regarding the interaction of long-term and short-term memory).

### (g) Impact of new learning on prior long-term knowledge

A third stream of research has examined the processes by which newly learned phonological forms become established in the lexicon, as assessed by whether they interfere with the processing of previously known words (e.g. Gaskell & Dumay 2003; Dumay & Gaskell 2005, 2007). In these studies, certain aspects of the learning of the new word forms, such as facilitation in responding to them, emerged relatively quickly, within a single session. However, evidence that the new lexical entries were interfering with earlier ones (i.e. that they had established a *lexical footprint* as Gaskell & Dumay (2003) termed it) emerged only after a period of time, with an intervening night's sleep being critical, thus suggesting a role of a longer-term consolidation process of the kind studied

in the long-term memory literature. Generalizing these different facets of the learning of a new word, Leach & Samuel (2007) distinguished between what they termed *lexical configuration* which refers to knowledge about the word itself, such as its phonology or meaning or grammatical properties, and *lexical engagement*, which refers to how the word interacts with other lexical elements, for example, by interfering with them. Leach & Samuel (2007) noted that studies of word learning have typically focused on the former and emphasized that it is equally important to investigate the latter.

With these streams of research also, it is not entirely clear what the locus is of the hypothesized roles of long-term memory or its interaction with short-term memory. In Gupta's formulation (Gupta & Dell 1999; Gupta & Cohen 2002), there is an explicit correspondence between procedural learning/memory and phonological word-form learning and between declarative learning/memory and learning of the receptive and expressive links. The locus of interactions between short-term and long-term memory is currently unclear. The locus of the memory consolidation effects described by Gaskell and co-workers (Gaskell & Dumay 2003; Dumay & Gaskell 2005, 2007) would appear to be at the level of the phonological word-form representations. In each case, however, the mechanisms have remained unspecified.

## 3. TOWARDS A COMPUTATIONAL ACCOUNT

It is worth noting that a number of the issues discussed in the previous section are the very ones highlighted in a recent article in a handbook of psycholinguistics (Samuel & Sumner 2009). Although primarily concerned with surveying the field of spoken word recognition, Samuel & Sumner (2009) note that models of spoken word recognition must be able to account for changes in the lexicon over time—that is, they must account for word learning. With regard to word learning, issues highlighted by Samuel & Sumner (2009) include the role of phonological short-term memory in word learning (our second issue); the manner in which newly learned words are affected by neighbourhood density (our first issue) and the processes whereby newly learned words interfere with previously known words (an aspect of our fourth issue). The issues we have discussed above thus have considerable currency. However, as noted in our preceding discussion, there is little current understanding of the mechanistic bases of these phenomena. In addition, there is little understanding of whether or how these phenomena relate to each other. Clearly, each has been identified as an important aspect of word learning. But are they all disparate effects? Or is there some way in which they are related to each other?

Recent computational work by Gupta & Tisdale (2009) offers the beginnings of a framework within which several of these issues can be understood and integrated. In what follows, we first summarize relevant aspects of this computational work, and then discuss how it can provide an integration of issues discussed above. Before discussing the model, however,

it will be useful to augment the functional account of word learning introduced in the first section of this article, extending it to a simple model of lexical processing. Figure 2 provides a simple conceptualization of the processing of word forms. To the earlier functional scheme, the present conceptualization adds additional detail about phonological processing. In particular, it emphasizes that spoken word forms are *sequences* of sounds—sequences of sublexical units such as phonemes. This is depicted by the incoming arrow representing phonological input. The internal word-form representation is thus one that is activated by a *sequence* of sounds. It is also the basis for producing an output sequence of phonological patterns that constitute production of a word form, as shown by the outward arrow depicting output phonological sequences. These additions highlight the fact that the internal word-form representation must incorporate sequential information about the entire word form, both because it must distinguish words such as *bat* from *tab* in input processing, and because it must enable production of word forms as sequences, in output processing.

In this conceptualization, when the sequence of input sounds constituting a spoken word form (whether known or novel) is presented to the system, it evokes an internal representation of that word form. The input is thought of as consisting of a sequence of sublexical units, such as phonemes or syllables. The internal representation of the word form, once evoked, can evoke any associated semantic representation, via the receptive link. Conversely, if a semantic representation is activated in the system, it can activate any associated internal word-form representation, via the expressive link. When an internal word-form representation has been activated (whether by phonological input or via semantics), it forms the basis for the system to be able to produce the word form, as a sequence of output sublexical units. Such production constitutes either naming (if the internal word-form representation was activated via semantics) or repetition (if the activation was via an input phonological sequence).

Although the essential structure of this processing model is the same as that of the earlier functional framework, it is distinguished by its emphasis on processing aspects related to the level of word-form representations. It was this processing subsystem that formed the basis of Gupta & Tisdale's (2009) computational investigations, in which they examined the learning of word forms *without* any associated semantics. The primary motivation for their work was to examine the roles of PSTM and vocabulary size in vocabulary learning. As noted previously, these variables are thought to have their hypothesized effects at the phonological word-form level. Gupta & Tisdale (2009) therefore 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
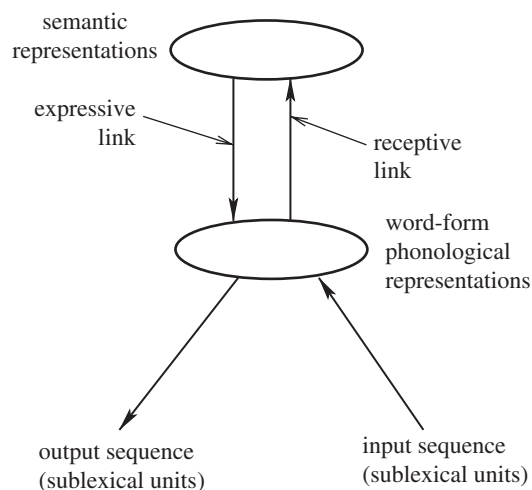
Figure 2. Model of lexical processing.

acquired a *phonological vocabulary*. Gupta & Tisdale (2009) were then able to examine various aspects of the model's behaviour and functioning, including: factors that affected its phonological vocabulary learning; its ability to repeat unlearned sequences (i.e. its non-word repetition) as well as factors that affected this ability; and how the model instantiated PSTM.

## (a) *The model of Gupta & Tisdale*

The model developed by Gupta & Tisdale (2009) was a neural network or 'connectionist' model. The architecture of the model was adapted from recent work by Botvinick & Plaut (2006) and is shown in figure 3. The model had an input layer at which a distributed representation of an entire syllable was presented; this input representation was assumed to have been created by an earlier stage of processing. The model also had an output layer at which the model's response was produced. The representation of a syllable, at both the input and the output layers, was in terms of a CCVCC (i.e. Consonant–Consonant–Vowel–Consonant–Consonant) frame. That is, the representation scheme for a syllable consisted of units divided into five slots. Activation of units in the first slot denoted the first C (if any) of the syllable, activation of units in the second slot denoted the second C (if any) of the syllable, activation of units in the third slot denoted the V of the syllable, and so on. Within each of the slots, the various phonemes that are legal for that slot for English were 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. Five bits (i.e. binary digits) are needed to represent 17 different binary patterns and five binary units were therefore necessary for the first slot. Thus, each of the 17 possible phonemes was represented as a different binary pattern of activation across this set of five units. Similarly, the 21 phonemes that are possible in the V slot were represented as patterns of activation across a set of five units constituting the V slot (five units suffice for up to 32 different patterns) and so

on for the various slots shown at the input and output layers in figure 3. As an example, the syllable bat (in IPA, /bæt/) was represented by presentation of the binary pattern representing /b/ in the first slot, the pattern representing /æ/ in the third slot, and the pattern representing /t/ in the fourth slot. The binary patterns encoding particular phonemes did not incorporate phonological feature information and thus did not encode phonological similarity between phonemes. The representation scheme did, however, encode phonological similarity between syllables—for example, the representations of /bæt/ and /kæt/ differed only in the activation pattern in the first slot.

In addition to the input and output layers, the model also had a hidden layer of 200 units. All units in the input layer projected to all units in the hidden layer, and all units in the hidden layer projected to all units in the output layer. The hidden layer additionally had recurrent self-connections such that every unit in the hidden layer had a projection (i.e. connection) to every unit in the hidden layer, including itself. This is depicted in figure 3 by the circular arrow on the hidden layer. The model also incorporated feedback connections from the output layer back to the hidden layer.

The model took as its input a sequence of one or more syllables constituting a monosyllabic or polysyllabic phonological word form. Syllabification of the input was assumed to occur in an earlier stage of processing. Thus, a word form was presented to the model one syllable at a time. Each phoneme was represented individually within a syllable, using the CCVCC scheme described previously. However, the model did not incorporate phoneme-level sequencing within a syllable, and in addition abstracted away from numerous detailed aspects of speech processing, such as precisely how syllabification and resyllabification is accomplished in a continuous speech stream (as noted above, word forms were pre-syllabified when input to the model), and questions such as speaker variability and speech rate. For Gupta & Tisdale's (2009) primary goal of examining the role of PSTM in serially ordered word-form production and learning, such phenomena were not the focus of interest, and it sufficed for the model to represent and produce serial order across syllables within a word form, but not within a syllable, as the model still would be processing and producing phonological *sequences* constituting word forms. Gupta & Tisdale (2009) also noted that there is considerable evidence to suggest that the syllable is a natural unit of phonological analysis and that there is perceptual segmentation at the level of the syllable (e.g. Massaro 1989; Menyuk *et al.* 1986; Jusczyk 1986, 1992, 1997), so that it was reasonable to treat word forms as sequences of syllables rather than as sequences of phonemes.

After the input had been presented syllable by syllable, the model attempted to produce as its output the entire word form, as the correct sequence of syllables (including correct phoneme representations within each syllable). Importantly, the model's output production was performed when there was no longer any information in the input about the word form
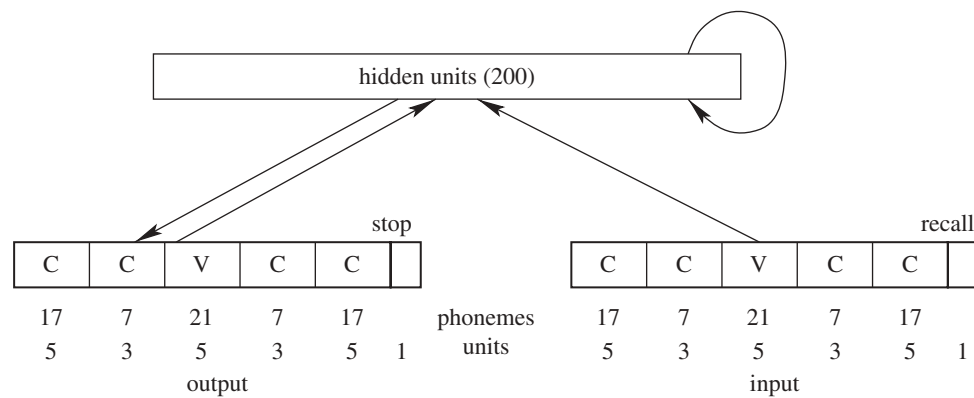
Figure 3. Architecture of the Gupta & Tisdale (2009) model.

that was presented. In order to perform the task, the model therefore had to develop an internal representation of the entire word form that included information about both its serial ordering and the phonemic structure of the syllables comprising the word form. Over the course of many such learning instances, the model learned a phonological vocabulary—the set of phonological sequences that it could correctly produce.

## (b) *Processing in the model*

Let us consider the syllable-by-syllable processing of a word form in the model. Table 1 shows the regime of presentation and desired output for the word form *flugwish*. The procedure is the same irrespective of whether the word form has been presented to the model previously. In response to 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, of course, may or may not be correct. Either way, after the model has produced an output, the actual output is compared with the target output (in this case, flug) and the discrepancy between the two is calculated (what in neural networks is often termed the *error*). Then, the activation pattern that is present at the hidden layer is transmitted across the recurrent connections from the hidden layer to itself, so that when the second syllable *wish* is presented at the next time step, the model's hidden layer will actually receive input from two sources: the input representing wish and input from its own past state.

Thus, at the second time step of processing, the input to the network is the second syllable of the word form, together with the recurrent input from the previous hidden layer activation pattern. The model's task, as for the first syllable, is to produce the current input syllable (now wish), at the output. Again, the output may or may not be correct, and the error is calculated. Again, the hidden layer activation pattern is transmitted across the recurrent connections so as to be available at the next time step.

At the next (third) time step, however, there is no phonological input, because wish was the final syllable of the word form, so that presentation of the word form has been completed. At this point the model's task is to produce at the output layer the entire sequence of syllable representations previously

Table 1. Processing regime in the Gupta & Tisdale (2009) model, illustrated for the word form flugwish.

| time step | input | target output |
|---|---|---|
| 1 | flug | flug |
| 2 | wish | wish |
| 3 | Recall | flug |
| 4 | Recall | wish |
| 5 | Recall | stop |

presented at the input layer, i.e. flug followed by wish. That is, the network's task at this point is to repeat the preceding word form, as a serially ordered sequence, when the input no longer contains any information about the word form. To indicate this to the network, the input is now a signal to recall the preceding sequence of inputs. This signal or cue is denoted by activation of the 'Recall' unit in the input layer. Thus, on the third time step, in the presence of the Recall input signal (together with the recurrent input from the previous time step), the network's task is to produce flug at the output. This target and the network's actual produced output are again used in the calculation of error. Again, the hidden layer activation pattern is transmitted so as to be available at the next (fourth) time step. At the fourth time step, the Recall signal is once again presented as input (together with recurrent input from the previous hidden layer activation pattern). The network's task is to produce the second syllable of the preceding sequence, wish. Again, the difference between this target output and the actual output produced by the network is used to determine error. Again, the hidden layer activation pattern is transmitted across the recurrent connections for availability at the next time step. At the next (fifth) time step, the input is still the Recall signal together with the recurrent input. The network's task on this time step (the final step in repetition of this word form) is to activate a specially designated 'Stop' unit at the output layer, to signify the end of its production of the word form. Once again, this target and the network's actual produced output are used in the calculation of error. At the end of this repetition process, the total error that has been calculated over all five time steps is used to adjust the model's connection weights, using a variant of the back-propagation learning

algorithm (*back-propagation through time*; Botvinick & Plaut 2006; Rumelhart *et al*. 1986). A key aspect of this algorithm is that, by using error from all steps of processing, it adjusts weights on the hidden layer recurrent connections in a way that takes into account information about the entire sequence whose presentation and attempted repetition has just ended. It may be worth emphasizing that the Recall signal carries no information about the content of any particular syllable or word form (it is the cue for recall of all syllables in all word forms) and that the model's repetition response therefore truly is in the absence of external information about the word form.

### (c) *Learning a phonological vocabulary*

A set of approximately 125 000 phonologically distinct words of one through four syllables was drawn from online corpora, and a syllabified phonemic transcription was created for each word. The transcription for each syllable in a particular word was then further translated into a set of binary vectors, one for each phoneme, according to the scheme described in the previous section. The simulation of vocabulary learning consisted in presenting the model with a set of 1000 words drawn from the overall set of 125 000, with adjustment of connection weights occurring after presentation of each word. This procedure (syllable-by-syllable presentation, and connection weight adjustment, for each of the 1000 words) was termed an *epoch*, and vocabulary learning consisted of a large number of such epochs (as discussed further shortly). The 125 000 words were intended to approximate the set of words of the language, and the sample of 1000 words in an epoch was intended to be very loosely analogous to the kind of exposure to a subset of these words of the language that a human learner might receive in a period of time such as a day. The 1000 words in each epoch were selected stochastically from the overall set of 125 000, with the probability of selection of a given word into the sample of 1000 for an epoch being based on its frequency of occurrence.

For each of the words in an epoch, the procedure described in the previous section was followed: the model was exposed to the word one syllable at a time, attempting to repeat each syllable after its presentation, and attempting to repeat the entire word one syllable at a time after presentation was complete; connection weights were adjusted at the end of this process. Following presentation and connection weight adjustment for all 1000 words in an epoch, the model's production accuracy was then assessed in a test using the same procedure (but now without any adjustment of connection weights) one word at a time, for all 125 000 vocabulary items. The model's performance in producing a given word was considered correct only if, when producing the entire word after its presentation, the model produced the correct phoneme in every slot of each syllable, and all the syllables in the correct serial order. Otherwise, the model's performance in producing that word was considered incorrect, and the word was not considered a correctly produced vocabulary item. The number of

words correct in the test was taken as a measure of the model's phonological vocabulary size at that epoch. A further stochastic sample of 1000 words was then selected, and the entire above procedure repeated, constituting another epoch. Overall, 8000 epochs of such exposure-and-test were provided to the model, for a total of 8 000 000 learning trials.

Figure 4 shows development of the model's phonological vocabulary across these epochs. The upper curve in the figure shows the number of words to which the model has been exposed at least once. The lower line shows the size of the model's vocabulary, determined as above. As can be seen, the model's phonological vocabulary exhibits steady growth as a function of exposure to the simulated linguistic environment. By about 6000 epochs, phonological vocabulary size had largely asymptoted, with that level being maintained through the remainder of the 8000 epochs of training. Clearly, the model exhibited a developmental trajectory; moreover, the trajectory exhibits the kind of power law learning often characteristic of human cognition.

### (d) *Key characteristics of the model*

The model's ability to acquire a phonological vocabulary provided a platform from which Gupta & Tisdale (2009) were able to examine how the model processed *novel* word forms to which it had not been exposed—i.e. non-words. Before describing their findings, it is worth exploring three aspects of the model that are critical to its functioning and to its performance in word-form processing.

The first of these is the fact that what the 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 information transmitted across the recurrent connections on the hidden layer, as described above. This information enables the model to know where it is in producing the current word form, by providing information about what has already been produced. Gupta & Tisdale (2009) pointed out that this information is information about the *past*, and thus indubitably constitutes *memory* information, and that 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 (Simulation 4) that it constitutes PSTM information. This indicated that PSTM functionality was crucial to serially ordered word-form production in the model.

A second aspect was that the model's PSTM functionality was not independent of its long-term knowledge. This was because the hidden-layer recurrent connections were weighted connections. Therefore, the PSTM information transmitted across these recurrent connections from one time step to the next in processing of a word form was always influenced by the weights on those connections. But these
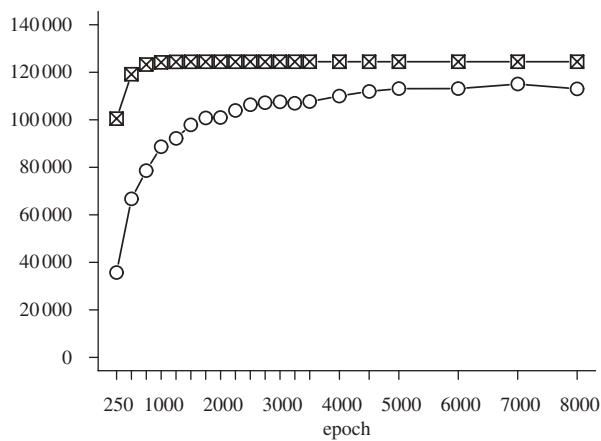
Figure 4. Simulated phonological vocabulary learning (Simulation 1, Gupta & Tisdale 2009). The *y*-axis shows the number of words that have been presented at least once, and the number of words that are correctly produced. Crosses, presented; open circles, correct.

connection weights were determined by long-term learning, just as for all connection weights in the model. Specifically, they reflected the model's long-term phonological knowledge, established through its long-term vocabulary learning. As a result, the PSTM information transmitted across these recurrent connections was always influenced by the weights on those connections—that is, it was always coloured by the model's long-term phonological knowledge.

A third aspect is one that is shared with other neural network models that employ distributed representations. As noted in the earlier discussion of systematic and arbitrary mappings, the most important characteristic of distributed representations is that they enable similar stimuli to have similar representations. If a neural network model employs distributed representations at both input and output, then after it has learned a large number of input–output pairings, if it is presented with a novel input, it will produce an output that is similar to the outputs it has learned to produce in response to inputs that are similar to the novel input. If the model simulates a domain in which the input–output mapping is systematic, this response to the novel input will likely be the correct one, as discussed earlier. Thus, in a neural network model that incorporates distributed representations and a systematic mapping, such as the Gupta & Tisdale (2009) model of input–output phonology, generalization to novel stimuli is automatic.

Gupta & Tisdale (2009) used this model to demonstrate various points about the role of PSTM functionality in word learning as well as about non-word repetition. Essentially, all of these demonstrations were the consequence of the three characteristics described above. One key demonstration was that phonological vocabulary learning is causally affected by PSTM functionality (Simulation 6), as has been proposed by Gathercole, Baddeley and co-workers (e.g. Baddeley *et al*. 1998; Gathercole *et al*. 1999). Gupta & Tisdale (2009) examined this by repeating the simulation of phonological vocabulary learning described in the previous section and depicted in figure 4. However, the model was now required to learn a

phonological vocabulary without the benefit of PSTM functionality, which was eliminated by disabling the ability of the hidden-layer recurrent connections to transmit information. The results differed dramatically from those shown in figure 4, with correct vocabulary performance remaining essentially around zero throughout the 8000 epochs of training. Gupta & Tisdale (2009) showed that PSTM functionality also causally affected non-word repetition (Simulation 5), by examining non-word repetition performance with PSTM functionality eliminated in the same manner as above, but with the elimination occurring at the end of vocabulary learning rather than at the beginning. Despite 8000 epochs of normal vocabulary growth, the model was completely unable to repeat non-words in the absence of PSTM functionality.

At the same time as demonstrating a causal effect of PSTM functionality on phonological vocabulary learning and non-word repetition, Gupta & Tisdale's (2009) investigations showed that non-word repetition was also causally affected by vocabulary size (Simulation 7). This was examined by tracking the model's non-word repetition performance across the 8000 epochs of phonological vocabulary learning. Non-word repetition accuracy improved as a function of the model's development, with no adjustment having been made to PSTM functionality, consistent with the suggestions of Edwards *et al*. (2004) and others. Furthermore, because the model's PSTM functionality was affected by its long-term knowledge as discussed above, PSTM functionality was itself improved by phonological vocabulary growth and the concomitant increase in the model's long-term phonological knowledge. But increased PSTM functionality in turn led to further improved phonological vocabulary learning and greater phonological knowledge. The pattern of causality underlying phonological vocabulary growth thus involved a feedback loop from PSTM functionality to phonological vocabulary growth (and hence increased phonological knowledge) to PSTM functionality back to phonological vocabulary growth and increased phonological knowledge, and thus included the positive feedback loop from phonological knowledge to vocabulary learning posited by Edwards *et al*. (2004) and others (for further discussion, see Gupta & Tisdale 2009).

Thus overall, the model offered a means of reconciling the debate described earlier about whether phonological vocabulary growth and non-word repetition are driven by PSTM functionality or by phonological vocabulary knowledge. The model made it unambiguously clear that they are affected by both factors. In addition, Gupta & Tisdale (2009) were able to show that the model's non-word repetition exhibited empirically observed effects of variables, such as phonotactic probability, word-form length and within-word-form syllable serial position, and also offered an account of the empirically observed correlations between non-word repetition, vocabulary size, and new word learning.

## 4. TOWARDS AN INTEGRATED FRAMEWORK
For present purposes, what is most relevant is how the Gupta & Tisdale (2009) model offers a framework for

integrating many of the issues of current focus that we highlighted earlier. Before discussing this, it is worth pointing out that the model is, in effect, an instantiation of the phonological word-form processing components of the simple lexical processing model we described in the previous section. This should be fairly clear from a comparison of figures 2 and 3. The patterns of activation at the computational model's input and output layers in figure 3 constitute the input and output phonological sequences depicted more abstractly in figure 2. The computational model's hidden layer corresponds with the abstract model's word-form phonological representation level. The correspondence is exact: the activation pattern that is present at the computational model's hidden layer at the end of presentation of a word form is indeed an internal encoding of both the phonological content and the serial order of that word form (Gupta & Tisdale 2009). Thus, the model provides a concretization of mechanisms of word-form phonological processing.

The first issue we highlighted earlier as a recent focus of research on word learning was the role of pho- notactic probability. As noted briefly above, in the Gupta & Tisdale (2009) model, non-words of higher phonotactic probability are repeated more accurately (Simulation 3). To more directly examine the *learning* of new word forms differing in phonotactic probability, we conducted a new simulation for the present article, in which high and low phonotactic probability non- words were each presented to the model multiple times, so that the model had the opportunity to *learn* them (rather than merely repeat each one once). These non-words were presented to the model after the point at which it had acquired its phonological vocabulary (i.e. at the endpoint of the learning trajec- tory shown in figure 4). Two hundred four-syllable non-words were created (100 each of high and low phonotactic probability) and the entire set of 200 non- words was presented to the model 10 times. The results of this simulation are shown in figure 5, which plots the model's repetition accuracy for high and low phono- tactic probability non-words, at initial exposure (shown in the graph as Exposure 1), after five exposures and after 10 exposures. As can be seen, at initial exposure, which corresponds to non-word repetition, accuracy is greater for high than low phonotactic prob- ability non-words, replicating Simulation 3 of Gupta & Tisdale (2009). The new result is that this advantage is maintained through the exposure trials, so that at the end of this learning, the high phonotactic probability non-words exhibit better performance (*p* < 0.001 for the high–low comparison at all three exposure points). This corresponds to the finding that high phonotactic probability non-words show better per- formance in word learning tasks at the end of several learning trials (e.g. Storkel & Rogers 2000; Storkel 2001). It should be noted that the empirical results are in the context of receptive word expressive word learning, which involves semantics, whereas the present simulation pertains to word-form learning in the absence of semantics. Nevertheless, by demonstrating an effect of phonotactic probability at the end of word-form learning, these results establish that the
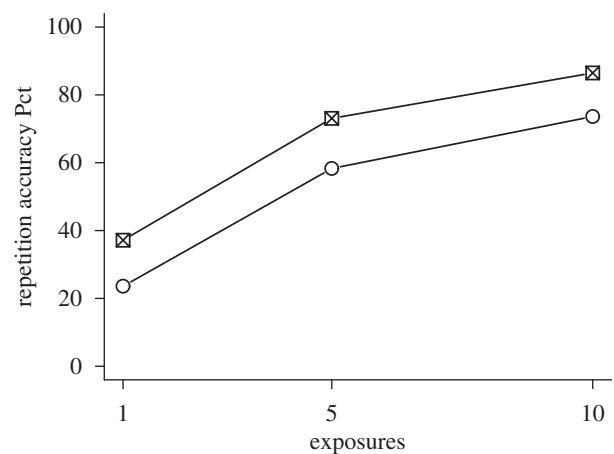


Figure 5. Simulated learning of novel word forms of high and low phonotactic probability. Crosses, high probability; open circles, low probability.

phonological level can indeed be a locus of such effects (although they do not rule out other loci in the schematic of figure 2).

The occurrence of these phonotactic probability effects in the model falls out of the very manner of its operation and learning: phones and combinations of phones that occur more frequently in the vocabulary have been processed more frequently by the model, and connection weights in the model have therefore been adjusted frequently in ways that reflect these phonotactic contingencies—thus leading to better performance in generalizing to new stimuli (i.e. non- words) that incorporate them. This is consistent with the explanation that has typically been offered for the effects of phonotactic probability in human non- word repetition (e.g. Edwards *et al.* 2004); however, the present model provides a concrete instantiation of and mechanistic basis for such an explanation and extends it to the learning of novel word forms. In addition, as noted above, the model indicates a specific locus for such effects.

As discussed above, Gupta & Tisdale (2009) also demonstrated that the model's PSTM functionality was critical to its word-learning capability. This, together with the fact that the model provided a clear operationalization of the construct of PSTM, means that the model provides an integrated account of the effects of PSTM and phonotactic probability on word learning. It is the weights on the connections that instantiate the phonotactic probability effects. It is the recurrent connections that are especially critical for PSTM functionality. However, the overall func- tioning of the model when processing/learning a word form does not depend separately on PSTM func- tionality and other aspects of the model—it depends on the entire system, which includes recurrent connec- tions that especially support PSTM, non-recurrent connections and the weights on all the connections, that yield the phonotactic probability effect. Thus, the effects of PSTM functionality and phonotactic probability are closely integrated in the functioning of the model. In addition, these results indicate that both of these effects have a locus at the level of

word-form phonological representations, in terms of figure 2.

Gupta & Tisdale (2009) were able to show that vocabulary size in the model was predictive of non-word repetition ability and, additionally, that vocabulary size was predictive of the size of the phonotactic probability effect in non-word repetition, as in the behavioural data of Munson (2001; Munson *et al.* 2005). Hierarchical regression analyses of the simulation data of Gupta & Tisdale (2009) show that vocabulary size and PSTM functionality make independent contributions to the learning of new word forms. These effects of vocabulary size on novel word-form processing and learning are also ones that are tightly integrated with the overall functioning of the model. They arise because vocabulary size in the model reflects the state of phonological knowledge encoded in the model's connection weights. Thus, the effects of vocabulary size are not the result of a separate vocabulary size variable that affects processing, but rather are inherent to the connection weights that are inherent to the model's functioning. Indeed, the model clarifies that vocabulary size is not a processing variable at all, but is merely a measured behavioural outcome of the state of the knowledge encoded in the model's connection weights. In addition, the model shows that these effects, too, have a locus at the level of word-form phonological representations.

As noted above, in the Gupta & Tisdale (2009) model, PSTM functionality is inherently affected by the weights on the recurrent connections, these weights being an aspect of long-term knowledge. The model therefore demonstrates a close coupling of short- and long-term memory/learning, and thus incorporates a very clear operationalization of how long- and short-term memory systems can interact, at least in the domain of phonological word-form learning. Once again, the model indicates that, in terms of figure 2, the level of word-form representations is one locus for such interactions. As also noted above, the kind of incremental weight adjustment incorporated in distributed connectionist networks is commonly regarded as being an instantiation of procedural learning. Therefore, the model provides an instantiation of Gupta & Dell's (1999; Gupta & Cohen 2002) proposal that the learning of phonological forms is accomplished via procedural memory/learning. It follows from the two previous points that the interaction of STM and long-term memory instantiates a close relationship between PSTM functionality and procedural learning. It is worth noting, however, that although PSTM functionality is intrinsically coloured by long-term knowledge and learning—specifically, procedural learning—it is nevertheless a distinct functionality. That is, PSTM functionality is affected by procedural learning functionality, but is not subsumed by it.

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 & Dell's (1999; Gupta & Cohen 2002) 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. In this connection, the work by Gaskell and co-workers cited earlier is very interesting: it suggests the operation of consolidation processes in the learning even of word forms that are not associated with meanings—that is, in the lower half of figure 2. Although the nature of this consolidation process is currently unclear, it raises the possibility that something more than simple procedural learning may be engaged even at the level of phonological word-form learning. This may mean, in Leach & Samuel's (2007) terms, that aspects of lexical engagement require the involvement of nonprocedural learning mechanisms. In this case, the model of Gupta & Tisdale (2009) might constitute only a partial account of the learning of phonological forms, even in the absence of links with semantics. This would not invalidate any of the demonstrations made by the model about phonological word-form learning, but would indicate that the model does not capture all aspects of such learning.

As discussed above, there has been interest in examining the interaction of PSTM functionality and long-term knowledge in other verbal domains such as list recall and list learning. It is therefore worth noting that the operationalization of PSTM functionality and its interaction with long-term knowledge in the Gupta & Tisdale (2009) model is essentially the same as that in the Botvinick & Plaut (2006) model from which it is adapted. Botvinick & Plaut (2006) examined the immediate serial recall of lists of verbal items—that is, serial ordering in lists of word forms, rather than within individual word forms. They demonstrated the ability of their model to account for a variety of results that are regarded as benchmark phenomena in immediate list recall, and in addition, the interaction of PSTM functionality with long-term knowledge in list recall as well as list learning. Although the two models are separate, and the Gupta & Tisdale (2009) model does not simulate list recall while the Botvinick & Plaut (2006) model does not simulate within-word serial ordering, the continuity of PSTM functionality across the models indicates that the same principles of PSTM functionality can account for phenomena from both domains and for the interaction of short- and long-term memory in both domains.

The integrative framework that we have outlined in this article also offers insight into two further issues in lexical processing. The first of these has to do with the relation between PSTM and language production. Acheson & MacDonald (2009) recently provided an extensive review of data that suggest commonalities in serial ordering mechanisms between the domains of verbal short-term memory and language production (see also Page Cumming Madge & Norris 2007). They argued that serial ordering in verbal working memory emerges from the mechanisms that provide for serial ordering in language production, rather than the other way around, and rather than there being separate serial ordering mechanisms. It should be clear that this

is precisely the relationship between short-term memory functionality and language production that is instantiated in the Gupta & Tisdale (2009) model, at the lexical level: the recurrent connections that critically underlie PSTM functionality are an inherent part of the phonological processing architecture. The framework thus indicates how serial-order processing is integrated across language production and verbal working memory, at least in the lexical domain. The Botvinick & Plaut (2006) model suggests a similar integration at the level of word sequences.

In addition, the Gupta & Tisdale (2009) model 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 figures 2 and 3, 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. Two caveats must be noted here. First, the computational model as currently implemented does not process word forms as sequences of phonemes, but as sequences of syllables. It therefore cannot currently address phenomena such as cohort effects that appear to be based on phoneme-level sequential processing. In addition, the issue of whether input and output phonologies are separate or overlapping has been the topic of much debate, and the present computational results certainly do not settle the debate. Nevertheless, even with these caveats in mind, the model does provide an interesting suggestion regarding the integration of lexical recognition and production.

The present framework is, of course limited: with respect to figure 2, its characterization is limited largely to the level of word-form phonological processing and learning. However, we are not aware of other frameworks that provide a comparable level of integration across word-form phonological processing, short-term memory, long-term memory, vocabulary knowledge and phonotactic knowledge (but see Page & Norris (2009) for another recent approach and also Gupta (2009) for integration of some of these aspects). We thus believe that the framework outlined in this article offers considerable promise as a means of drawing together several important issues in the study of word learning. In doing so at the phonological level, it clarifies issues and makes it potentially easier to extend an integrated view of word learning towards inclusion of semantic representations and the receptive and expressive links. Such work is currently under way and should prove fruitful in bringing further aspects of word learning into an integrated framework.

## REFERENCES

Acheson, D. J. & MacDonald, M. C. In press. Verbal working memory and language production: common approaches to the serial ordering of verbal information. *Psychol. Bull.* **135**, 50–68.

Atkins, P. W. B. & Baddeley, A. D. 1998 Working memory and distributed vocabulary learning. *Appl. Psycholinguist.* **19**, 537–552. (doi:10.1017/S0142716400010353)

Au, T. K. & Glusman, M. 1990 The principle of mutual exclusivity in word learning: to honor or not to honor? *Child Dev.* **61**, 1474–1490. (doi:10.2307/1130757)

Baddeley, A. D. 1993 Short-term phonological memory and long-term learning: a single case study. *Eur. J. Cogn. Psychol.* **5**, 129–148. (doi:10.1080/09541449308520112)

Baddeley, A. D., Papagno, C. & Vallar, G. 1988 When long-term learning depends on short-term storage. *J. Mem. Lang.* **27**, 586–595. (doi:10.1016/0749-596X(88)90028-9)

Baddeley, A. D., Gathercole, S. E. & Papagno, C. 1998 The phonological loop as a language learning device. *Psychol. Rev.* **105**, 158–173. (doi:10.1037/0033-295X.105.1.158)

Baldwin, D. A. & Markman, E. M. 1989 Establishing word–object relations: a first step. *Child Dev.* **60**, 381–398.

Botvinick, M. & Plaut, D. 2006 Short-term memory for serial order: a recurrent neural network model. *Psychol. Rev.* **113**, 201–233. (doi:10.1037/0033-295X.113.2.201)

Brown, G. D. A. & Hulme, C. 1996 Nonword repetition, STM, and word age-of-acquisition. In *Models of short-term memory* (ed. S. E. Gathercole), pp. 129–148. Hove, UK: Psychology Press.

Coady, J. A. & Aslin, R. N. 2003 Phonological neighborhoods in the developing lexicon. *J. Child Lang.* **30**, 441–469. (doi:10.1017/S0305000903005579)

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. (doi:10.1126/science.7414331)

de Saussure, F. 1916 *Cours de linguistique generale*. Paris, France: Payot.

Desrochers, A. & Begg, I. 1987 A theoretical account of encoding and retrieval processes in the use of imagery-based mnemonic techniques: the special case of the keyword method. In *Imagery and related mnemonic processes* (eds M. A. McDaniel & M. Pressley), pp. 56–77. Berlin, Germany: Springer-Verlag.

Dumay, N. & Gaskell, M. G. 2005 Do words go to sleep? Exploring consolidation of spoken forms through direct and indirect measures. *Behav. Brain Sci.* **28**, 69–70.

Dumay, N. & Gaskell, M. G. 2007 Sleep-associated changes in the mental representation of spoken words. *Psychol. Sci.* **18**, 35–39. (doi:10.1111/j.1467-9280.2007.01845.x)

Edwards, J., Beckman, M. E. & Munson, B. 2004 The interaction between vocabulary size and phonotactic probability effects on children's production accuracy and fluency in nonword repetition. *J. Speech Lang. Hear. Res.* **47**, 421–436. (doi:10.1044/1092-4388(2004/034))

Ellis, N. & Beaton, A. 1993 Factors affecting the learning of foreign language vocabulary: imagery keyword mediators and phonological short-term memory. *Q. J. Exp. Psychol.* **46A**, 533–558.

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.

Gallimore, R., Lam, D. J., Speidel, G. E. & Tharp, R. G. 1977 The effects of elaboration and rehearsal on long-term retention of shape names by kindergarteners. *Am. Educ. Res. J.* **14**, 471–483.

Gaskell, M. G. & Dumay, N. 2003 Lexical competition and the acquisition of novel words. *Cognition* **89**, 105–132. (doi:10.1016/S0010-0277(03)00070-2)

Gathercole, S. E. 1995 Is nonword repetition a test of phonological memory or long term knowledge? It all depends on the nonwords. *Mem. Cogn.* **23**, 83–94.

Gathercole, S. E. 2006 Nonword repetition and word learning: the nature of the relationship. *Appl. Psycholinguist.* **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. *J. Mem. Lang.* **28**, 200–213. (doi:10.1016/0749-596X(89)90044-2)

Gathercole, S. E. & Baddeley, A. D. 1990 The role of phonological memory in vocabulary acquisition: a study of young children learning arbitrary names of toys. *Br. J. Psychol.* **81**, 439–454.

Gathercole, S. E., Willis, C., Emslie, H. & Baddeley, A. D. 1992 Phonological memory and vocabulary development during the early school years: a longitudinal study. *Dev. Psychol.* **28**, 887–898. (doi:10.1037/0012-1649.28.5.887)

Gathercole, S. E., Hitch, G. J., Service, E. & Martin, A. J. 1997 Short-term memory and new word learning in children. *Dev. Psychol.* **33**, 966–979. (doi:10.1037/0012-1649.33.6.966)

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. *Appl. Cogn. Psychol.* **13**, 65–77. (doi:10.1002/(SICI)1099-0720(199902)13:1<65::AID-ACP548>3.0.CO;2-O)

Grossman, M. 1987 Lexical acquisition in alcoholic Korsakoff psychosis. *Cortex* **23**, 631–644.

Gupta, P. 1995. *Word learning and immediate serial recall: toward an integrated account.* PhD thesis, Department of Psychology, Carnegie Mellon University, Pittsburgh, PA.

Gupta, P. 1996 Word learning and verbal short-term memory: a computational account. In *Proc. of the Eighteenth Annual Meeting of the Cognitive Science Society* (ed. G. W. Cottrell), pp. 189–194. Mahwah, NJ: Lawrence Erlbaum.

Gupta, P. 2003 Examining the relationship between word learning, nonword repetition, and immediate serial recall in adults. *Q. J. Exp. Psychol.* (A) **56**, 1213–1236.

Gupta, P. 2005a Primacy and recency in nonword repetition. *Memory* **13**, 318–324. (doi:10.1080/09658210344000350)

Gupta, P. 2005b What's in a word? A functional analysis of word learning. *Perspect. Lang. Learn. Educ.* **12**, 4–8. (doi:10.1044/lle12.3.4)

Gupta, P. 2009 A computational model of nonword repetition, immediate serial recall, and nonword learning. In *Interactions between short-term and long-term memory in the verbal domain* (eds A. Thorn & M. Page), pp. 108–135. Hove, UK: Psychology Press.

Gupta, P. & Cohen, N. J. 2002 Theoretical and computational analysis of skill learning, repetition priming, and procedural memory. *Psychol. Rev.* **109**, 401–448. (doi:10.1037/0033-295X.109.2.401)

Gupta, P. & Dell, G. S. 1999 The emergence of language from serial order and procedural memory. In *The emergence of language, 28th Carnegie Mellon Symp. on Cognition* (ed. B. MacWhinney). Mahwah, NJ: Lawrence Erlbaum.

Gupta, P. & MacWhinney, B. 1997 Vocabulary acquisition and verbal short-term memory: computational and neural bases. *Brain Lang.* **59**, 267–333. (doi:10.1006/brln.1997.1819)

Gupta, P. & Tisdale, J. 2009 Does phonological short-term memory causally determine vocabulary learning? Toward a computational resolution of the debate. *J. Mem. Lang.*

Gupta, P., MacWhinney, B., Feldman, H. & Sacco, K. 2003 Phonological memory and vocabulary learning in children with focal lesions. *Brain Lang.* **87**, 241–252. (doi:10.1016/S0093-934X(03)00094-4)

Gupta, P., Lipinski, J., Abbs, B. & Lin, H. 2005 Serial position effects in nonword repetition. *J. Mem. Lang.* **53**, 141–162. (doi:10.1016/j.jml.2004.12.002)

Haist, F., Musen, G. & Squire, L. R. 1991 Intact priming of words and nonwords in amnesia. *Psychobiology* **19**, 275–285.

Heisler, L. 2005 The influence of neighborhood density and phonotactic probability on representational and production factors in word learning. *Diss. Abstr. Int. B Sci. Eng.* **66**, 1757.

Hollich, G. J., Hirsh-Pasek, K. & Golinkoff, R. M. 2000a Breaking the language barrier: an emergentist coalition model of the origins of word learning. *Monogr. Soc. Res. Child Dev.* **65**, v–123. (doi:10.1111/1540-5834.00090)

Hollich, G. J., Jusczyk, P. W. & Luce, P. A. 2000b Of words, birds, worms, and weeds: infant word learning and lexical neighborhoods. In *Proc. of the Twenty-Second Annual Conference of the Cognitive Science Society* (eds L. R. Gleitman & A. K. Joshi), p. 1030. Mahwah, NJ: Lawrence Erlbaum.

Hollich, G., Jusczyk, P. W. & Luce, P. A. 2002 Lexical neighborhood effects in 17-month-old word learning. In *BUCLD 26: Proc. of the 26th Annual Boston University Conf. on Language Development* (eds B. Skarabela, S. Fish & A. H.-J. Do). Somerville, MA: Cascadilla Press.

Horowitz, L. M. & Gordon, A. M. 1972 Associative symmetry and second-language learning. *J. Educ. Psychol.* **63**, 287–294. (doi:10.1037/h0032643)

Jusczyk, P. W. 1986 Toward a model of the development of speech perception. In *Invariance and variability in speech processes* (eds J. S. Perkell & D. H. Klatt). Mahwah, NJ: Lawrence Erlbaum.

Jusczyk, P. W. 1992 From general to language-specific capacities: the WRAPSA model of how speech perception develops. *J. Phonetics* **21**, 3–28.

Jusczyk, P. W. 1997 *The discovery of spoken language.* Cambridge, MA: MIT Press.

Jusczyk, P. W., Luce, P. A. & Charles-Luce, J. 1994 Infants' sensitivity to phonotactic patterns in the native language. *J. Mem. Lang.* **33**, 630–645. (doi:10.1006/jmla.1994.1030)

Leach, L. & Samuel, A. G. 2007 Lexical configuration and lexical engagement. *Cogn. Psychol.* **55**, 306–353. (doi:10.1016/j.cogpsych.2007.01.001)

Leonard, L. B., Schwartz, R. G., Morris, B. & Chapman, K. 1981 Factors influencing early lexical acquisition: lexical orientation and phonological composition. *Child Dev.* **52**, 882–887. (doi:10.2307/1129090)

Magnuson, J. S., Tanenhaus, M. K., Aslin, R. N. & Dahan, D. 2003 The time course of spoken word learning and recognition: studies with artificial lexicons. *J. Exp. Psychol. Gen.* **132**, 202–227. (doi:10.1037/0096-3445.132.2.202)

Mandel, D. R., Jusczyk, P. W. & Pisoni, D. B. 1995 Infants' recognition of the sound patterns of their own names. *Psychol. Sci.* **6**, 314–317. (doi:10.1111/j.1467-9280.1995.tb00517.x)

Markman, E. 1984 The acquisition and hierarchical organization of categories by children. In *Origins of cognitive skills* (ed. C. Sophian), pp. 371–400. Mahwah, NJ: Lawrence Erlbaum.

Markman, E. 1989 *Categorization and naming in children: problems of induction.* Cambridge, MA: MIT Press.

Markman, E. M. & Hutchinson, J. E. 1984 Children's sensitivity to constraints on word meaning: taxonomic versus thematic relations. *Cogn. Psychol.* **16**, 1–27. (doi:10.1016/0010-0285(84)90002-1)

Massaro, D. W. 1989 *Experimental psychology*. Orlando, FL: Harcourt Brace Jovanovich.

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. *Psychol. Rev.* **102**, 419–457. (doi:10.1037/0033-295X.102.3.419)

McCloskey, M. & Cohen, N. J. 1989 Catastrophic interference in connectionist networks: the sequential learning problem. In *The psychology of learning and motivation*, vol. 24 (ed. G. H. Bower), pp. 109–165. New York, NY: Academic Press.

Menyuk, P., Menn, L. & Silber, R. 1986 Early strategies for the perception and production of words and sounds. In *Language acquisition* (eds P. Fletcher & M. Garman), 2nd edn. Cambridge, UK: Cambridge University Press.

Merriman, W. & Bowman, L. 1989 The mutual exclusivity bias in word learning. *Monogr. Soc. Res. Child Dev.* **54**, 1–129.

Merriman, B. & Schuster, J. 1991 Young children's disambiguation of name reference. *Child Dev.* **62**, 1288–1301. (doi:10.2307/1130807)

Mervis, C. 1984 Early lexical development: the contributions of mother and child. In *Origins of cognitive skills* (ed. C. Sophian). Mahwah, NJ: Lawrence Erlbaum.

Mervis, C. & Pani, J. 1980 Acquisition of basic object categories. *Cogn. Psychol.* **12**, 496–522. (doi:10.1016/0010-0285(80)90018-3)

Mervis, C., Golinkoff, R. & Bertrand, J. 1994 Two-year-olds readily learn multiple labels for the same basic-level category. *Child Dev.* **65**, 1163–1177. (doi:10.2307/1131312)

Michas, I. C. & Henry, L. A. 1994 The link between phonological memory and vocabulary acquisition. *Br. J. Dev. Psychol.* **12**, 147–164.

Mishkin, M., Malamut, B. & Bachevalier, J. 1984 Memories and habits: two neural systems. In *Neurobiology of learning and memory* (eds G. Lynch, J. McGaugh & N. Weinberger), pp. 65–77. New York, NY: Guilford.

Mishkin, M., Vargha-Khadem, F. & Gadian, D. G. 1998 Amnesia and the organization of the hippocampal system. *Hippocampus* **8**, 212–216. (doi:10.1002/(SICI)1098-1063(1998)8:3<212::AID-HIPO4>3.0.CO;2-L)

Munson, B. 2001 Phonological pattern frequency and speech production in adults and children. *J. Speech Lang. Hear. Res.* **44**, 778–792. (doi:10.1044/1092-4388(2001/061))

Munson, B., Kurtz, B. A. & Windsor, J. 2005 The influence of vocabulary size, phonotactic probability, and wordlikeness on nonword repetition of children with and without specific language impairment. *J. Speech Lang. Hear. Res.* **48**, 1033–1047. (doi:10.1044/1092-4388(2005/072))

Nelson, K. E. & Bonvillian, J. D. 1973 Concepts and words in the 18-month-old: acquiring concept names under controlled conditions. *Cognition* **2**, 435–450. (doi:10.1016/0010-0277(73)90003-6)

Oviatt, S. 1982 Inferring what words mean: early development in infants' comprehension of common object names. *Child Dev.* **53**, 274–277. (doi:10.2307/1129662)

Page, M. P. A. & Norris, D. 2009 Is there a common mechanism underlying word-form learning and the Hebb repetition effect? Experimental data and a modelling framework. In *Interactions between short-term and long-term memory in the verbal domain* (eds A. Thorn & M. Page), pp. 136–156. Hove, UK: Psychology Press.

Page, M. P. A., Cumming, N., Madge, A. & Norris, D. 2007 Speech errors and the phonological similarity effect in short-term memory: evidence suggesting a common locus. *J. Mem. Lang.* **56**, 49–64. (doi:10.1016/j.jml.2006.09.002)

Paivio, A. & Desrochers, A. 1981 Mnemonic techniques in second language learning. *J. Educ. Psychol.* **73**, 780–795. (doi:10.1037/0022-0663.73.6.780)

Papagno, C. & Vallar, G. 1992 Phonological short-term memory and the learning of novel words: the effects of phonological similarity and item length. *Q. J. Exp. Psychol.* **44A**, 47–67.

Papagno, C., Valentine, T. & Baddeley, A. D. 1991 Phonological short-term memory and foreign-language learning. *J. Mem. Lang.* **30**, 331–347. (doi:10.1016/0749-596X(91)90040-Q)

Pressley, M. & Levin, J. R. 1981 The keyword method and recall of vocabulary words from definitions. *J. Exp. Psychol. Hum. Learn. Mem.* **7**, 72–76. (doi:10.1037/0278-7393.7.1.72)

Pressley, M., Levin, J. R., Hall, J. W., Miller, G. E. & Berry, J. K. 1980 The keyword method and foreign word acquisition. *J. Exp. Psychol. Hum. Learn. Mem.* **6**, 163–173. (doi:10.1037/0278-7393.6.2.163)

Pressley, M., Levin, J. R. & McDaniel, M. A. 1987 Remembering versus inferring what a word means: mnemonic and contextual approaches. In *The nature of vocabulary acquisition* (eds M. G. McKeown & M. E. Curtis), pp. 107–127. Mahwah, NJ: Lawrence Erlbaum.

Quine, W. V. O. 1960 *Word and object*. Cambridge, MA: MIT Press.

Rumelhart, D., Hinton, G. & Williams, R. 1986 Learning internal representations by error propagation. In *Parallel distributed processing*, vol. 1, Foundations (eds D. Rumelhart & J. McClelland), pp. 318–362. Cambridge, MA: MIT Press.

Samuel, A. G. & Sumner, M. 2009 Current directions in research on spoken word recognition. In *The Cambridge handbook of psycholinguistics* (eds M. Spivey, M. Joanisse & K. McCrae). Cambridge, UK: Cambridge University Press.

Samuelson, L. K. & Smith, L. B. 1999 Early noun vocabularies: do ontology, category organization and syntax correspond? *Cognition* **73**, 1–33. (doi:10.1016/S0010-0277(99)00034-7)

Samuelson, L., Newman, R., Gupta, P. & Tisdale, J. In preparation. What is that? Neighborhood effects in early word learning.

Service, E. 1992 Phonology, working memory, and foreign-language learning. *Q. J. Exp. Psychol.* **45A**, 21–50.

Service, E. & Kohonen, V. 1995 Is the relation between phonological memory and foreign language learning accounted for by vocabulary acquisition? *Appl. Psycholinguist.* **16**, 155–172. (doi:10.1017/S0142716400007062)

Squire, L. R. & Zola, S. M. 1998 Episodic memory, semantic memory, and amnesia. *Hippocampus* **8**, 205–211. (doi:10.1002/(SICI)1098-1063(1998)8:3<205::AID-HIPO3>3.0.CO;2-I)

Squire, L. R., Knowlton, B. & Musen, G. 1993 The structure and organization of memory. *Annu. Rev. Psychol.* **44**, 453–495. (doi:10.1146/annurev.ps.44.020193.002321)

Sternberg, R. J. 1987 Most vocabulary is learned from context. In *The nature of vocabulary acquisition* (eds M. G. McKeown & M. E. Curtis), pp. 89–105. Mahwah, NJ: Lawrence Erlbaum.

Storkel, H. L. 2001 Learning new words: phonotactic probability in language development. *J. Speech Lang. Hear. Res.* **44**, 1321–1337. (doi:10.1044/1092-4388(2001/103))

Storkel, H. L. 2003 Learning new words II: phonotactic probability in verb learning. *J. Speech Lang. Hear. Res.* **46**, 1312–1323. (doi:10.1044/1092-4388(2003/102))

Storkel, H. L. 2004 Do children acquire dense neighbor-hoods? An investigation of similarity neighborhoods in lexical acquisition. *Appl. Psycholinguist.* **25**, 201–221.

Storkel, H. L. & Rogers, M. A. 2000 The effect of probabil-istic phonotactics on lexical acquisition. *Clin. Linguist. Phon.* **14**, 407–425.

Swingley, D. & Aslin, R. N. 2007 Lexical competition in young children's word learning. *Cogn. Psychol.* **54**, 99–132. (doi:10.1016/j.cogpsych.2006.05.001)

Thorn, A. & Page, M. (eds) 2009 *Interactions between short-term and long-term memory in the verbal domain*. Hove, UK: Psychology Press.

Tincoff, R. & Jusczyk, P. W. 1999 Some beginnings of word comprehension in 6-month-olds. *Psychol. Sci.* **10**, 172–175. (doi:10.1111/1467-9280.00127)

Ullman, M. T. 2001 The declarative/procedural model of lexicaon and grammar. *J. Psycholinguist. Res.* **30**, 37–69. (doi:10.1023/A:1005204207369)

Ullman, M. T. 2004 Contributions of memory circuits to language: the declarative/procedural model. *Cognition* **92**, 231–270. (doi:10.1016/j.cognition.2003.10.008)

Underwood, B. J. & Schulz, R. W. 1960 *Meaningfulness and verbal learning*. Philadelphia, PA: Lippincott.

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. (doi:10.1126/science.277.5324.376)