LANGUAGE LEARNING
A Journal of Research in Language Studies

Language Learning ISSN 0023-8333

A Complementary Systems Account of Word Learning in L1 and L2

Shane Lindsay
University of York

M. Gareth Gaskell
University of York

We review a body of behavioral and neuroimaging research relating to the acquisition and integration of novel words. An important outcome from this research is that different aspects of knowledge associated with learning a new word become established over different time scales. We suggest that the temporal dissociations found in word learning are due to the application of and interaction between complementary learning systems in the brain, with rapidly acquired episodic representations stored via the medial temporal lobes and slower learning supported by neocortical systems. We discuss the implications of this model for understanding the earliest stages of learning a novel word and for learning words in a second language.

Learning New Words

Throughout your adult life you are constantly exposed to novel words. These could be encountered through exposure to a foreign language or they could be part of your native tongue. Consider the extremely rare English word *jussulent*, which means *full of soup or broth*. Whether you are a native or nonnative speaker of English, you likely experienced some confusion or surprise on encountering this word and therefore you have just participated in a word learning experience. In this article we describe behavioral and neuroimaging evidence on the consequences of learning new words, be they nonnative or native. You encountered the word *jussulent* orthographically, but in this article we will

Research in this article has been supported by a U.K. ESRC grant RES-063-27-0061 awarded to M. G. Gaskell. Thanks to Jelena Mirkovic for discussion of some of the ideas contained within.

Correspondence concerning this article should be addressed to Shane Lindsay, Department of Psychology, University of York, York, YO10 5DD, United Kingdom. Internet: shane.lindsay@gmail.com

focus our attention on the acquisition of novel words through the medium of speech. The successful storage of a novel word leads to the development of a new entry in the lexicon. Our goal in studying word learning is to understand how a novel word begins to show behavior that is characteristic of existing entries in the mental lexicon.

Recognizing and using a word appropriately depends on many kinds of information, such as knowledge of its phonology, orthography, syntax, and semantics. Correspondingly, there is a variety of different ways to determine lexical behavior. One key aspect of wordlike behavior that we have investigated is how a new word becomes part of the process of lexical competition during spoken word recognition. We take it that lexical competition is a result of a novel word being integrated into the mental lexicon and, therefore, we can assess the impact of learning a novel word by its impact on the processing of existing words. We will outline evidence for how a novel word begins to engage in this process of lexical competition, along with other aspects of behavior associated with learning a novel word, describing behavioral research in the next section and neuroimaging research in the third section. One perspective on word learning is that it is a continual process that can take place over weeks, months, or even years, with the representation of a word being gradually strengthened and enriched over this period (e.g., McMurray, 2007). Although we are supportive of this view, through the study of the development of lexical competition effects we have found behavioral and neural evidence that suggests that there are also qualitative changes in the representational status of novel words that occur over the course of hours or days that are not simply a result of exposure. In the fourth section, we will provide a theoretical account for these effects using a general systems-level model of the brain. This model suggests that the differences in lexical status found over time in our behavioral and neuroimaging research are due to the operation of two complementary learning systems with different learning properties and representational substrates. This theory provides a memory architecture that supports learning in both the first language (L1) and second language (L2), and in the final section we will outline the implications of this theory and our research for understanding how we learn words in an L2.

Behavioral Consequences of Learning Novel Words

We have focused on lexical competition as a hallmark of lexical behavior because in most (if not all) major models of spoken word recognition, the mechanism of lexical competition plays a crucial role (e.g., Luce, Pisoni, &

Goldinger, 1990). Recognizing a spoken word is a probabilistic process that relies on information that unfolds over time in the speech signal to infer the most likely intended target of the speaker. A consequence of this recognition process is that multiple candidates compete for activation as soon as information becomes available. This competition process can also occur within and across languages in bilinguals (Marian & Spivey, 2003). One candidate model for explaining lexical competition in speech recognition is Gaskell and Marslen-Wilson's (1997) distributed cohort model (DCM). In this model, lexical representations are distributed across a network as mappings between form and meaning. Lexical activation emerges in parallel during the competition process, arising from distributed overlapping patterns of activity ("blends"), with activation determined through a matching process between phonetic inputs and stored mappings to semantics and phonological codes. During word recognition, a point in the acoustic signal can be reached at which the activation of the network stabilizes at one point with maximal activation. At this point, a word can then be uniquely identified.

Given the importance of lexical competition for word recognition, participation in this process gives a benchmark for the lexical behavior of a novel word. The emergence of engagement in lexical competition was investigated in a study by Gaskell and Dumay (2003), which examined the dynamics of competition by investigating the consequences of learning novel words on the word recognition process. Participants were exposed to novel words such as *dolpheg*, which had the same onset as an existing English word with an early uniqueness point (e.g., *dolphin*). Participants were exposed to novel words using a phoneme monitoring task that involved close attention to their phonological forms (Connine & Titone, 1996). After training, participants were tested on a two-alternative forced-choice recognition task in which they had to discriminate these novel words from similar sounding forms (e.g., *dolphess*). In addition to this explicit measure of recognition, participants were tested on whether the novel words engaged in the lexical competition process. Participants performed an auditory lexical decision on existing words (e.g., *dolphin*). If the novel word *dolpheg* is integrated into the lexicon, we should expect this new onset competitor to slow down responses to existing words. Responses were compared to a control condition of a matched and counterbalanced set of existing words for which no novel neighbors had been learned. Gaskell and Dumay tested for competition effects immediately after exposure and every day for 5 days, with training every day except the last. Despite the relatively incidental nature of the learning task, participants showed near-ceiling performance in their recognition of the novel words after initial exposure. In contrast, no evidence was found for the novel words engaging in lexical competition until the fourth day. The novel words appeared to be functionally separated from the established lexicon and were not able to participate in the competition process immediately after acquisition.

One advantage of the design used by Gaskell and Dumay (2003) is that the impact of novel words on the recognition process is tested indirectly and performance on existing words can be compared with a control condition. However, a methodological concern regarding the Gaskell and Dumay design is that evidence for delayed competition with novel words was found by comparing performance of participants repeating the same task with the same stimuli at different time points. This makes it difficult to distinguish effects due to the passage of time and those that arise from task and item repetition. Evidence that the effects stem from time and not repetition was found in a study by Davis, Di Betta, Macdonald, and Gaskell (2009), in which participants were exposed to a set of novel words on one day and to a separate set of novel words a day later. Participants were then just tested once on the second day in a lexical decision task. Davis et al. found lexical competition effects only with items that were heard a day earlier, consistent with a time-based effect. A further methodological issue is whether these effects were specific to testing for competition using a lexical decision task. This was addressed by using a pause detection task instead (Gaskell & Dumay, 2003, Exp. 3). This task involves measuring response times to the detection of a silent pause inserted into a spoken word (Mattys & Clark, 2002), and the time taken to detect a pause is associated with the degree of lexical activity during processing at that point. This means that performance on this task can index the activation of multiple lexical items engaging in the lexical competition process. Using pause detection, Gaskell and Dumay found a similar pattern of delayed lexical competition effects as found with a lexical decision task. In other extensions of the findings of Gaskell and Dumay, a study by Tamminen and Gaskell (2008) looked at the time course of the emergence of lexical competition over time periods longer than several days. They found that competition effects with novel words were still present and reliable after 8 months from initial training.

These studies have consistently shown that lexical processing a day after hearing a novel word differs from the first day, without any further training, suggesting that representations of novel words undergo a process of consolidation within 24 hr. Dumay and Gaskell (2007) looked at whether sleep is involved in this consolidation process. Half of their participants came into the lab at 8 a.m., where they were exposed to novel words and tested immediately on lexical competition effects. Twelve hours later the participants returned to be tested again. This group was compared with another group trained and tested at 8 p.m. and then again at 8 a.m. the next day, after a night's sleep. Although both groups recognized the novel words accurately immediately after training, neither group showed evidence of lexical competition right away. Twelve hours later, the group without sleep still showed no evidence of lexical competition, whereas the group that had slept demonstrated engagement of novel words in lexical competition. Twenty-four hours after initial exposure, when all participants had undergone a period of nocturnal sleep, evidence for lexical competition was found in both groups.

These results stem from a paradigm that has often been used in behavioral studies of sleep-related memory consolidation (e.g., Fenn, Nusbaum, & Margoliash, 2003) and are in line with many other studies that show that sleep is associated with improvements in memory (see Walker & Stickgold, 2006, for a review). For example, sleep-associated memory enhancements can lead to enhanced skill performance (Brawn, Fenn, Nusbaum, & Margoliash, 2008) or to reduced susceptibility to interference for declarative memory (Ellenbogen, Hulbert, Stickgold, Dinges, & Thompson-Schill, 2006). The majority of studies showing sleep-associated consolidation typically demonstrate effects for learning new perceptual or motor skills, but a growing number of studies demonstrate enhancements for language learning, such as studies showing that sleep can aid generalization in artificial grammar learning tasks (Gómez, Bootzin, & Nadel, 2006; St. Clair & Monaghan, 2008). There are also several studies that now show that sleep can alter the representations of new memories allowing for abstraction (Fischer, Drosopoulos, Tsen, & Born, 2006; Wagner, Gais, Haider, Verleger, & Born, 2004) and integration of newly learned relationships (Ellenbogen, Hu, Payne, Titone, & Walker, 2007). However, what makes the Dumay and Gaskell (2007) result particularly interesting, not only for language research but also for understanding of memory consolidation, is that delayed emergence of lexical competition results from a case in which new information has to be integrated with existing knowledge. These results suggest that sleep has a role for integration of new information, which is a task crucial for language learning.

Although we have focused on the time course of the emergence of lexical competition, using it as a marker of wordlike behavior, we have also explored other indicators of lexical behavior. Davis et al. (2009) found that participants were significantly faster in a speeded oral repetition task for words trained the previous day compared with the same day, suggesting that overnight consolidation facilitates speeded access to novel forms. Snoeren, Gaskell, and Di Betta

(2009) looked at the effects of a novel word on perception of assimilation of place of articulation (e.g., where the /n/ in "*lean bacon*" is articulated more like [m]). Existing research has shown that a lexical context can affect compensation for assimilated segments. For example, people can use the fact that *leam* is not a word to uncover the underlying word *lean* (Gaskell & Marslen-Wilson, 1998). In Snoeren et al., participants learned novel words such as *decibot* and subsequently heard these words in sentential contexts in which the following speech could license assimilation of the canonical form, such as "*. . .decibop behaved. . .*." They found that compensation for assimilation could happen immediately after learning a novel word, indicating a lexical influence without the need for consolidation. However, the detection of underlying forms of assimilated segments (such as /t/ in "*. . .decibop behaved. . .*") was nonetheless quicker on the day after learning, indicating that form-based information was more easily accessed after a period of offline consolidation. In other studies, Lindsay, Sedin, and Gaskell (2010) looked at whether a novel word could exert a lexical bias on the perception of ambiguous phonemes (Ganong, 1980) and found that novel words exerted a top-down bias both immediately and a week later.

These lines of research suggest that some aspects of a novel word's ability to influence lexical processing are immediately available, whereas some aspects take longer to materialize. This pattern is supported by the findings of other research groups. Bowers, Davis, and Hanley (2005) found that competition effects with visually presented words occurred only a day after training, not on the first day. Clay, Bowers, Davis, and Hanley (2007) looked at whether novel words with meanings could exhibit picture-word interference effects and found that these effects emerged only a week after training, with no additional practice. Other research has found that the type of training is also an important factor in the emergence of lexical behavior. In a similar vein to the work on the Ganong effect in novel words, Leach and Samuel (2007) used the fact that whereas words can exert a top-down influence on the boundaries of ambiguous categories in perceptual learning, pseudowords do not (Norris, McQueen, & Cutler, 2003). They investigated the time course of how a novel word could influence phonetic categorization over 5 days of testing and whether the type of training mattered. They found that a critical determiner of lexicality effects was the type of training. No effects were found on the first day when words were trained without semantics, but when trained with meanings, the novel words showed lexically based perceptual learning effects. Additionally, the strength of the effect appeared to be modulated by the richness of the semantics associated with the words.

In our research, semantic training was not sufficient for lexical competition effects to show on the same day of training (Dumay, Gaskell & Feng, 2004), but we have investigated whether other kinds of training have an impact on the time course of consolidation of novel forms. In recent work we have used spaced learning within a single day to see whether lexical competition effects could emerge on a time scale that does not involve sleep (Lindsay & Gaskell, 2009). Spaced learning refers to distributing learning episodes over time and has long been shown to be advantageous for memory performance compared with the equivalent exposure all at once (see Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006, for a review). Spaced learning might also be considered a more naturalistic training method and has been shown to benefit foreign vocabulary acquisition (Dempster, 1987). Our previous studies using phoneme monitoring for novel word training (e.g., Dumay & Gaskell, 2007) have used a single session of repeated exposure to novel forms, a form of massed learning. Lindsay and Gaskell (2009) found that by spacing out learning over the course of a day, lexical competition emerged within a time period that did not include sleep. This study shows that along with sleep, spaced learning can also promote lexical integration. These results can both be explained using our theoretical approach, which will be elucidated in the fourth section.

Neural Consequences of Learning Novel Words

Further work has explored the neural changes that take place when learning a novel word. Davis et al. (2009) used functional magnetic resonance imaging (fMRI) to examine neural responses to novel words learned the same day as scanning, compared with novel words learned the previous day. Activation elicited by trained novel words was compared with activation for existing words and completely novel words heard for the first time in the scanner. Increased activity in response to unfamiliar novel words was principally found in the anterior and posterior superior temporal gyrus and left cerebellum. Elevated responses in these regions were also shown with novel words that participants had heard earlier that day. However, there was no significant difference between novel words trained the previous day and words with an existing lexical representation. Strikingly, a period of 24 hr from initial exposure to a novel word was enough to produce a brain response similar to that of real words, whereas learned words that had not undergone overnight consolidation were responded to in the neocortex as if they were completely novel.

Along with these consolidation-related differences, Davis et al. (2009) also looked at brain structures involved in the initial encoding of words prior to sleep. They found significant activation of the hippocampal region of the medial temporal lobe (MTL) in response to novel words encountered for the first time, a response that was significantly weaker for words encountered during training on the same day. Furthermore, the size of this hippocampal response was predictive of performance on the novel items in a later recognition task. In a follow-up fMRI experiment, we explored the learning of novel orthographic forms rather than phonological forms and additionally looked at the consolidation of meaning (Gaskell et al., 2009). Results from this study also show activation within the hippocampus in response to the initial encoding of novel forms.

Increased activity in hippocampal regions has also been found in other neuroimaging studies of word learning. Breitenstein et al. (2005) exposed participants to novel words that were either consistently or inconsistently paired with pictures that provided a meaning. Over successive consistent form-meaning pairings, activity in the left hippocampus reduced, with no such reduction found for inconsistent form-meaning pairings. Like the study by Davis et al. (2009), hippocampal activity was able to significantly predict performance on a subsequent memory test. More evidence comes from a study by Mestres-Misse, Camara, Rodriguez-Fornells, Rotte, and Münte (2008), which investigated the development of form-meaning pairings using novel words in sentences, and activity in the MTL (anterior parahippocampal gyrus) was associated with the development of lexical representations for these novel words.

To summarize, neuroimaging data support the behavioral evidence that consolidation of novel words over time alters their representation so that they more closely resemble existing words in the lexicon. The neuroimaging evidence suggests that brain structures in the MTL, particularly the hippocampus, are important for the initial acquisition of novel words.

A Complementary Learning Systems Account of Word Learning

In this section we provide an account of word learning based on the idea that we have two interdependent learning systems in the brain: a slow learning cortical network and a rapid learning MTL system. We base our approach on one of the most influential models in the memory literature, the Complementary Learning Systems (CLS) account of learning and memory (McClelland, McNaughton, & O'Reilly, 1995; O'Reilly & Norman, 2002), although the main tenets of the approach are compatible with a number of theories that share similar assumptions (e.g., Marr, 1971). A more detailed and comprehensive application of the CLS theory to word learning can be found in Davis and Gaskell (2009).

The CLS approach was developed in part to deal with a problem known as the stability-plasticity dilemma (Carpenter & Grossberg, 1987). Our memory systems need to be plastic to acquire new knowledge, but if the learning rate of a memory system is set too high, then the new information may interfere with or overwrite existing knowledge. This type of catastrophic interference is a particular problem for distributed connectionist models (McCloskey & Cohen, 1989). The CLS account provides a solution to the catastrophic interference problem by characterizing learning as the operation of multiple memory systems with different architectures and computational principles. In this framework, the neocortex forms the basis of our long-term internal model of the environment, and its primary memory function is the abstraction of experience using distributed and overlapping representations. Slow and gradual neocortical learning is supplemented by areas of the MTL (particularly the hippocampus) that are specialized for the rapid storage of cortical patterns, using computational principles of sparse coding and pattern separation, which allow new memories to be segregated and distinguished from each other, supporting memory of unique episodes. Rather than updating new information in neocortical memory systems while they are online, the separate but interacting MTL system can index cortical patterns, which can then be interleaved into long-term cortical memory when the system is offline (see also Marr, 1971). The theory suggests that sleep provides an opportunity for regions of the MTL to reciprocally activate the neocortex, resulting in neocortical consolidation. Such a theory is consistent with animal studies that show hippocampal "replay" of new memories during sleep (e.g., Skaggs & McNaughton, 1996) and with neurophysiological studies showing coordination through synchronized firing patterns between the hippocampus and the neocortex during sleep (e.g., Sirota, Csicsvari, Buhl, & Buzsáki, 2003). This reactivation allows for new memories to be consolidated via a process of interleaving new information with existing knowledge. Exactly how this interleaving process works is currently unclear, but connectionist modeling has suggested that it might occur by rehearsal of existing or randomly generated patterns ("pseudopatterns"), whereas new knowledge is integrated (French, Ans, & Rousset, 2001), allowing for the perseveration of the existing knowledge in the network.

The CLS theory helps to explain the data we have collected on delayed consolidation of novel phonological forms. When encountering a novel word, the new representation is not immediately added to the mental lexicon. Due to the problem of catastrophic interference, a novel word might disrupt the representation of existing words (e.g., learning *cathedruke* might render the listener unable to recognize *cathedral*). Instead, MTL systems allow for rapid indexing of the new form information, as well as allowing for an arbitrary meaning to be rapidly linked to the new form. These MTL-mediated representations support the recognition of new words without the need for consolidation. However, given the opportunity for offline consolidation, the new form information is, over time, interleaved with existing lexical representations. The role of sleep in the CLS framework naturally provides an explanation for our data on sleep-associated changes in novel word representation (Dumay & Gaskell, 2007). The integration during offline states allows the emergence of more fully lexical behavior, such as the ability to participate in lexical competition (Dumay & Gaskell, 2007). In addition, integration allows for neural responses similar to existing words (Davis et al., 2009). The CLS account also helps to interpret the presleep lexical competition effects using spaced learning in Lindsay and Gaskell (2009). We explain these effects as the operation of online learning reliant on neocortical systems. It suggests that repeated exposure to novel forms spaced out over time allows for a degree of integration of new forms into the lexicon while the system is still online, reducing the need for offline consolidation.

The CLS account is a dual-systems model for both word learning and word recognition, with two routes involved in the word recognition process: a neocortical route and a MTL route, for consolidated long-term lexicophonological knowledge and unconsolidated rapidly learned knowledge, respectively (Davis & Gaskell, 2009). After consolidation of novel forms in the neocortex, they can participate quickly and automatically in the probabilistic word recognition process, whereas access to unconsolidated information may be slower, more indirect, and more automatic. One of the greatest challenges for current models of word recognition is accounting for evidence that word recognition relies on both fine phonetic detail in the speech signal and the extraction of more abstract context invariant units (such as phonemes or syllables). The CLS framework provides an account of the architecture that supports the storage of both these sources of information, with detailed unique phonetic information associated with hearing a word rapidly acquired via MTL systems and more abstract information gradually acquired in neocortical areas (see Goldinger, 2007, for a similar view). We have preliminary evidence for the prediction that access to unconsolidated form information takes more time, as described earlier, such as the finding that the repetition of unconsolidated novel words is slower (Davis et al., 2009) and that it takes longer to detect assimilated segments in these words (Snoeren et al., 2009). Evidence that the MTL route may be slower comes from observations that episodic talker-specific representations take longer to influence processing than more abstract representations

(McLennan & Luce, 2005). However, these predictions need further empirical investigation, as there is likely to be differential weighting of these two routes depending on particular task demands.

In summary, the application of the CLS approach to word learning provides an explanatory architecture that supports the earliest stages of word learning. An initial encounter with a novel word leads to the rapid storage of a unique record of that experience, dependent on the MTL memory system. This record may include fine phonetic detail and a variety of nonphonetic information, such as the identity of the speaker, the situation it was encountered in, and links to an initial meaning. Whereas storage could potentially occur automatically, we expect that the strength and detail of this memory trace is dependent on the level of attention during encoding. During offline periods following initial exposure, these information sources are integrated into the distributed long-term storage systems supporting word recognition. Cortico-MTL connectivity, necessary for access to unconsolidated information and associations, is gradually supplanted by cortico-cortico connectivity as novel forms are consolidated, allowing for more efficient mappings between form and meaning.

Implications for L1 and L2 Word Learning

We have described a body of behavioral and neuroimaging research showing that some aspects of lexical behavior require consolidation before emerging, such as engagement in lexical competition with other words (Dumay & Gaskell, 2007; Gaskell & Dumay, 2003). Other aspects of lexical behavior are available immediately after exposure to a novel word, such as the top-down biasing of phoneme discrimination (Lindsay et al., 2010) and compensation for coarticulated segments (Snoeren et al., 2009). These findings have potential implications for the study of L2 acquisition. When studying knowledge of a new language and its impact on the existing language system, the time between initial exposure and testing is of crucial importance. Testing for knowledge of a new language may produce different results immediately after exposure compared with a day or week later, due to the time course of consolidation processes.

A CLS account of L2 acquisition allows for some basic predictions to be made regarding how we learn an L2. We should expect that initial representations of an L2 word will be mediated by MTL memory systems and that access to representations in this system should be slower than for consolidated neocortical representations. Following offline consolidation, neocortical representations of L2 should be strengthened, with sleep playing an important role. Initial support for this last claim comes from a study showing that sleep can enhance the learning of foreign vocabulary (Gais, Lucas, & Born, 2006). Alternatively, gradual learning of language via online interleaving can accelerate the development of new L2 representations in neocortical memory stores (Lindsay & Gaskell, 2009), and spaced learning should enhance language learning, as has been shown by studies indicating that spaced learning can improve vocabulary learning (Dempster, 1987).

Underlying this account, our theoretical perspective involves a commitment to explaining psycholinguistic behavior with a domain-general account of learning and memory, as far as is possible. Applying this to L2 word learning, this means that there should be no hard divisions, cognitively or neurally, between the lexicons for the L1 and L2. Nonetheless, differences such as the age of acquisition, frequency of exposure to the L1 and L2, proficiency, and lexicophonological structure may lead to more graded representational differences.

With respect to age of acquisition, an L2 is often learned later in life, when the L1 is already well established (Seidenberg & Zevin, 2006), and this factor may have significant implications for the functional separability of the L1 and L2. At present we know little about the extent to which new vocabulary remains dependent on the hippocampus in the longer term. We have observed striking changes in lexical organization over the course of a single night, but this does not mean that hippocampal consolidation is complete by this point. Furthermore, the changes that we observe are only half of the story, as they tell us about what the neocortex has gained rather than what the hippocampus has lost. Studies of hippocampal amnesia suggest that the hippocampus remains important for years or even decades after a memory has first formed (Nadel & Moscovitch, 1997), although some neuroimaging work shows a gradient of hippocampal activity that drops to near baseline after months rather than years (Takashima et al., 2006). A potentially influential factor in determining the longevity of a hippocampal trace is the degree to which it fits in with a preexisting schema or knowledge base. Tse et al. (2007) have found that for associations between odors and locations in rats, the duration of hippocampal dependence is determined by the extent to which a prior set of similar stimulus-location mappings has been learned. It is undoubtedly a substantial leap to relate this kind of study to L1 and L2 learning in humans; nonetheless, such behavior fits with a CLS account of memory. In the case of language learning, there is a well-established set of mappings between form and meaning for L1, and so, in many cases, incorporating a new word should be relatively straightforward, which could mean that hippocampal dependence is quite short-lived. In the case of L2 learning, where the L2 schema is weakly established and the L1 schema is a poor fit (perhaps because of phonological differences), there may be a reliance on the hippocampus as a mediating structure for a longer period of time.

These speculations are most obviously related to the process of lexical competition in the L1 and L2. A variety of studies, using a range of techniques, have demonstrated that under certain circumstances, words from both languages can compete with each other during spoken word recognition. In single-word paradigms or when words are embedded in fairly neutral phrases, there is clear evidence of lexical competition involving both L1 and L2 words (e.g., Marian & Spivey, 2003; Weber & Cutler, 2004). In other cases, particularly when the linguistic context is strong or immersive, some selectivity can be found (e.g., FitzPatrick & Indefrey, 2010). Interestingly, the extent to which words from the nonpresented language contribute to the recognition process seems to depend in part on the listener's proficiency in that language. In other words, if a language has been learned relatively late or relatively weakly, words from that language play a weaker part in recognition processes (e.g., Blumenfeld & Marian, 2007; Silverberg & Samuel, 2004). Although many factors may influence this asymmetry, it is conceivable that part of the effect is carried by a neural dissociation in recognition for the late-acquired language, with a greater reliance on hippocampal mediation leading to slower activation of recently acquired words. In the most extreme case, we would predict that the hippocampal system is dominantly involved in the immediate effects of learning L2 words. Our work on lexical competition and the CLS approach suggests that if L2 words are learned on the same day of testing, we should expect access to their representations to be primarily hippocampally mediated and therefore less able to influence the recognition of existing L1 and L2 words. Over time, these words will gradually get consolidated into neocortical word recognition systems and have greater influence on the word recognition process.

With regard to lexicophonological differences between L1 and L2 lexicons, our research has relied on the use of word learning paradigms using nonsense words. Although our main goals have been to understand the native language system, one of the advantages of such studies is that the methodology can be informative about both L1 and L2 learning. In our studies of lexical competition, we have used novel words that were closely related to existing forms in a participant's L1 (e.g., *cathedral/cathedruke*), but learning of these forms could be a model for the learning of L2 cognates or false cognates. In other cases, experiments have used words that follow phonological rules of an L1, but with forms designed to be very weakly related to existing words or morphemes (Gaskell et al., 2009) or with sparse phonological neighborhoods in L1

(Magnuson, Tanenhaus, Aslin, & Dahan, 2003). All of these types of stimuli have some validity for studying mechanisms of L2 lexical acquisition. However, irrespective of whether one is studying L1 or L2 language learning, we suggest that one of the most important influences on word learning is the similarity of a new form to existing lexicophonological representations. As discussed earlier, the degree of similarity of new forms to existing forms may influence the ease and speed of a hippocampally mediated memory trace being integrated into preexisting knowledge or schemas.

An important advantage of a complementary systems theory is that it provides a means of incorporating novel words quickly into a distributed system while avoiding catastrophic interference. It is clear that catastrophic interference does not occur in the acquisition of individual L1 or L2 words when both languages remain in use, but there are documented examples of L2 immersion causing more or less complete L1 loss. One such case is that of Korean adults who were adopted by French families in early childhood. These participants were exposed to Korean during an fMRI study, and they showed a neural response that was no different from control participants who had not been previously exposed to Korean (Pallier et al., 2003). This kind of loss can indeed be described as catastrophic, and so could be viewed as evidence against a CLS account of word learning. In fact, however, this kind of effect fits rather well within a CLS model. The hippocampal system provides a buffer for new L2 words, supporting the interleaving of new linguistic knowledge in offline states. This allows the system to avoid any catastrophic interference in the short term. Nonetheless, the interleaving of the new language in the absence of any new exposure to the L1 causes changes in the main neocortical store. Over time, this neocortical mapping becomes more and more attuned to L2, and the influence of L1 becomes weaker. Thus, the CLS model does not prevent loss of old knowledge altogether in these rare cases when there is a dramatic switch between two languages in terms of input. Rather, it ensures that the consequences of the switch in inputs are gradual, leading to a more graceful form of interference over the course of years rather than days. A connectionist account of such changes also provides a simple explanation for preserved implicit knowledge of the lost L1, in that the weights in the network could still contain some semblance of L1 knowledge even when explicit tests of that knowledge (e.g., vocabulary recognition) fail. This latent knowledge could then be uncovered when the participant is retrained on their lost language as an adult (Bowers, Mattys, & Gage, 2009).

We are in the early stages in the application of the CLS approach to understanding word learning and L2 acquisition. Along with the application of the theory to L2 learning, fundamental challenges remain in our understanding of the coordination between cortical areas involved with speech perception and the MTL during learning and perception. We need to further investigate the mechanisms behind online learning and offline consolidation that allow lexical integration to occur. An additional issue, not touched upon here, is the role of short-term and working memory in word learning and how this relates to the CLS account, given evidence for the importance of the phonological loop in vocabulary acquisition (Baddeley, Gathercole, & Papagno, 1998). Although many challenges remain, we hope that the CLS approach to word learning in L1 and L2 may help stimulate research in L1 and L2 acquisition and draw attention to the important role for consolidation in acquiring lexical representations.

References

Baddeley, A. D., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review, 105*, 158–173.

Blumenfeld, H. K., & Marian, V. (2007). Constraints on parallel activation in bilingual spoken language processing: Examining proficiency and lexical status using eye-tracking. *Language and Cognitive Processes, 22*, 633–660.

Bowers, J. S., Davis, C. J., & Hanley, D. A. (2005). Interfering neighbours: The impact of novel word learning on the identification of visually similar words. *Cognition, 97*, 45–54.

Bowers, J. S., Mattys, S. L., & Gage, S. H. (2009). Preserved implicit knowledge of a forgotten childhood language. *Psychological Science, 20*, 1064–1069.

Brawn, T. P., Fenn, K. M., Nusbaum, H. C., & Margoliash, D. (2008). Consolidation of sensorimotor learning during sleep. *Learning and Memory, 15*, 815–819.

Breitenstein, C., Jansen, A., Deppe, M., Foerster, A.-F., Sommer, J., Wolbers, T., et al. (2005). Hippocampus activity differentiates good from poor learners of a novel lexicon. *NeuroImage, 25*, 958–968.

Carpenter, G. A., & Grossberg, S. (1987). A massively parallel architecture for a self-organizing neural pattern recognition machine. *Computer Vision, Graphics, and Image Processing, 37*, 54–115.

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin, 132*, 354–380.

Clay, F., Bowers, J. S., Davis, C. J., & Hanley, D. A. (2007). Teaching adults new words: The role of practice and consolidation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 970–976.

Connine, C. M., & Titone, D. (1996). Phoneme monitoring. *Language and Cognitive Processes, 11*, 635–645.

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

Davis, M. H., & Gaskell, M. G. (2009). A complementary learning systems account of word learning: Neural and behavioural evidence. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*, 3773–3800.

Dempster, F. N. (1987). Effects of variable encoding and spaced presentations on vocabulary learning. *Journal of Educational Psychology, 79*, 162–170.

Dumay, N., & Gaskell, M. G. (2007). Sleep-associated changes in the mental representation of spoken words. *Psychological Science, 18*, 35–39.

Dumay, N., Gaskell, M. G., & Feng, X. (2004). A day in the life of a spoken word. In K. Forbus, D. Gentner & T. Regier (Eds.), *Proceedings of the twenty-sixth annual conference of the Cognitive Science Society* (pp. 339–44). Mahwah, NJ: Erlbaum.

Ellenbogen, J. M., Hu, P. T., Payne, J. D., Titone, D., & Walker, M. P. (2007). Human relational memory requires time and sleep. *Proceedings of the National Academy of Sciences United States of America, 104*, 7723–7728.

Ellenbogen, J. M., Hulbert, J. C., Stickgold, R., Dinges, D. F., & Thompson-Schill, S. L. (2006). Interfering with theories of sleep and memory: Sleep, declarative memory and associative interference. *Current Biology, 16*, 1290–1294.

Fenn, K. M., Nusbaum, H. C., & Margoliash, D. (2003). Consolidation during sleep of perceptual learning of spoken language. *Nature, 425*, 614–616.

Fischer, S., Drosopoulos, S., Tsen, J., & Born, J. (2006). Implicit learning-explicit knowing: A role for sleep in memory system interaction. *Journal of Cognitive Neuroscience, 18*, 311–319.

FitzPatrick, I., & Indefrey, P. (2010). Lexical competition in non-native speech comprehension. *Journal of Cognitive Neuroscience, 22*, 1165–1178.

French, R. M., Ans, B., & Rousset, S. (2001). Pseudopatterns and dual-network memory models: Advantages and shortcomings. In R. French & J. Sougné (Eds.), *Connectionist models of learning, development and evolution* (pp. 13–22). London: Springer.

Gais, S., Lucas, B., & Born, J. (2006). Sleep after learning aids memory recall. *Learning and Memory, 13*, 259–262.

Ganong, W. F., III (1980). Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance, 6*, 110–125.

Gaskell, M. G., & Dumay, N. (2003). Lexical competition and the acquisition of novel words. *Cognition, 89*, 105–132.

Gaskell, M. G., Lindsay, S., Tamminen, J., Chen, Y., Li, R., Wolfson, J., et al. (2009, October). *The time-course of form and meaning consolidation in novel word acquisition*. Poster presented at the Neurobiology of Language Conference, Chicago, IL. Retrieved November 10, 2009, from http://www.nlc2009.angularis.org/ScientificProgramNLC2009.pdf

Gaskell, M. G., & Marslen-Wilson, W. D. (1997). Integrating form and meaning: A distributed model of speech perception. *Language and Cognitive Processes, 12*, 613–656.

Gaskell, M. G., & Marslen-Wilson, W. D. (1998). Mechanisms of phonological inference in speech perception. *Journal of Experimental Psychology: Human Perception and Performance, 24*, 380–396.

Goldinger, S. D. (2007, August). *A complementary-systems approach to abstract and episodic speech perception*. Paper presented at the 16th International Congress of Phonetic Sciences, Saarbrücken, Germany. Retrieved October 25, 2009, from http://www.icphs2007.de/conference/Papers/1781/1781.pdf

Gómez, R. L., Bootzin, R. R., & Nadel, L. (2006). Naps promote abstraction in language-learning infants. *Psychological Science, 17*, 670–674.

Leach, L., & Samuel, A. G. (2007). Lexical configuration and lexical engagement: When adults learn new words. *Cognitive Psychology, 55*, 306–353.

Lindsay, S., & Gaskell, M. G. (2009). Spaced learning and lexical integration of novel words. In N. A. Taatgen & H. van Rijn (Eds.), *Proceedings of the 31st Annual Conference of the Cognitive Science Society* (pp. 2517–2522). Austin, TX: Cognitive Science Society.

Lindsay, S., Sedin, L., & Gaskell, M. G. (2010). *Acquiring novel words and their tenses: Evidence from lexical effects on phonetic categorisation*. Manuscript in preparation.

Luce, P. A., Pisoni, D. B., & Goldinger, S. D. (1990). Similarity neighborhoods of spoken words. In G. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp. 122–147). Cambridge, MA: MIT Press.

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. *Journal of Experimental Psychology: General, 132*, 202–227.

Marian, V., & Spivey, M. (2003). Competing activation in bilingual language processing: Within- and between-language competition. *Bilingualism: Language and Cognition, 6*, 97–115.

Marr, D. (1971). Simple memory: A theory for archicortex. *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences, 262*, 23–81.

Mattys, S. L., & Clark, J. H. (2002). Lexical activity in speech processing: Evidence from pause detection. *Journal of Memory and Language, 47*, 343–359.

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning-systems in the hippocampus and neocortex. *Psychological Review, 102*, 419–457.

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

McLennan, C. T., & Luce, P. A. (2005). Examining the time course of indexical specificity effects in spoken word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 306.

McMurray, B. (2007). Defusing the childhood vocabulary explosion. *Science, 317*, 631.

Mestres-Misse, A., Camara, E., Rodriguez-Fornells, A., Rotte, M., & Münte, T. F. (2008). Functional neuroanatomy of meaning acquisition from context. *Journal of Cognitive Neuroscience, 20*, 2153–2166.

Nadel, L., & Moscovitch, M. (1997). Memory consolidation, retrograde amnesia and the hippocampal complex. *Current Opinion in Neurobiology,* 7, 217–227.

Norris, D., McQueen, J. M., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive Psychology, 47*, 204–238.

O'Reilly, R. C., & Norman, K. A. (2002). Hippocampal and neocortical contributions to memory: Advances in the complementary learning systems framework. *Trends in Cognitive Sciences, 6*, 505–510.

Pallier, C., Dehaene, S., Poline, J. B., Le Bihan, D., Argenti, A. M., Dupoux, E., et al. (2003). Brain imaging of language plasticity in adopted adults: Can a second language replace the first? *Cerebral Cortex, 13*, 155.

Seidenberg, M. S., & Zevin, J. D. (2006). Connectionist models in developmental cognitive neuroscience: Critical periods and the paradox of success. In Y. Munakata & M. H. Johnson (Eds.), *Attention and performance XXI: Processes of change in brain and cognitive development* (pp. 315–347). Oxford: Oxford University Press.

Silverberg, S., & Samuel, A. G. (2004). The effect of age of second language acquisition on the representation and processing of second language words. *Journal of Memory and Language, 51*, 381–398.

Sirota, A., Csicsvari, J., Buhl, D., & Buzsáki, G. (2003). Communication between neocortex and hippocampus during sleep in rodents. *Proceedings of the National Academy of Sciences United States of America, 100*, 2065–2069.

Skaggs, W. E., & McNaughton, B. L. (1996). Replay of neuronal firing sequences in rat hippocampus during sleep following spatial experience. *Science, 271*, 1870–1873.

Snoeren, N. D., Gaskell, M. G., & Di Betta, A. M. (2009). The perception of assimilation in newly learned novel words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*, 542–549.

St. Clair, M. C., & Monaghan, P. (2008). Language abstraction: Consolidation of language structure during sleep. In B. C. Love, K. McRae, & V. M. Sloutsky (Eds.), *Proceedings of the 30th annual conference of the Cognitive Science Society* (pp. 727–732). Austin, TX: Cognitive Science Society.

Takashima, A., Petersson, K. M., Rutters, F., Tendolkar, I., Jensen, O., Zwarts, M. J., et al. (2006). Declarative memory consolidation in humans: A prospective functional magnetic resonance imaging study. *Proceedings of the National Academy of Sciences United States of America, 103*, 756–761.

Tamminen, J., & Gaskell, M. G. (2008). Newly learned spoken words show long-term lexical competition effects. *Quarterly Journal of Experimental Psychology, 61*, 361–371.

Tse, D., Langston, R. F., Kakeyama, M., Bethus, I., Spooner, P. A., Wood, E. R., et al. (2007). Schemas and memory consolidation. *Science, 316*, 76–82.

Wagner, U., Gais, S., Haider, H., Verleger, R., & Born, J. (2004). Sleep inspires insight. *Nature, 427*, 352–355.

Walker, M. P., & Stickgold, R. (2006). Sleep, memory, and plasticity. *Annual Review of Psychology, 57*, 139–166.

Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language, 50*, 1–25.