Novel Word Integration in the Mental Lexicon: Evidence from Unmasked and Masked Semantic Priming

Jakke Tamminen and M. Gareth Gaskell

Department of Psychology, University of York

Author note

Jakke Tamminen, M. Gareth Gaskell, Department of Psychology, University of York, United Kingdom.

Jakke Tamminen is now at the Department of Psychology, Royal Holloway, University of London, United Kingdom.

Address correspondence to: Jakke Tamminen, Department of Psychology, Royal Holloway, University of London, Egham, TW20 0EX, United Kingdom. E-mail: jakke.tamminen@gmail.com.

Abstract

We sought to establish whether novel words can become integrated into existing semantic networks by teaching participants new meaningful words, and then using these new words as primes in two semantic priming experiments, in which participants carried out a lexical decision task to familiar words. Importantly, at no point in training did the novel words co-occur with the familiar words that served as targets in the primed lexical decision task, allowing us to evaluate semantic priming in the absence of direct association. We found familiar words were primed by the newly related novel words, both when the novel word prime was unmasked (Experiment 1) and masked (Experiment 2), suggesting that the new words had been integrated into semantic memory. Furthermore, this integration was strongest after a 1-week delay and was independent of explicit recall of the novel word meanings: forgetting of meanings did not attenuate priming. We argue that even after brief training newly learned words become an integrated part of the adult mental lexicon rather than being episodically represented separately from the lexicon.

Keywords: word learning, semantic priming, memory consolidation.

Acknowledgements

Jakke Tamminen was supported by an Economic and Social Research Council (ESRC) doctoral studentship, an ESRC postdoctoral fellowship (PTA-026-27-2540), and ESRC grant RES-062-23-2268. Gareth Gaskell was supported by an ESRC fellowship (RES-063-27-0061) and a Leverhulme Trust grant (F/00 224/AO).

Novel Word Integration in the Mental Lexicon: Evidence from Unmasked and Masked Semantic Priming

As we attempt to keep up with the flow of new information and ideas that surround us, we rely on the flexibility and receptiveness of the human mind. One cognitive faculty that is central in allowing us to learn about the world is the mental lexicon. The lexicon provides vital shortcuts to ideas and objects, allowing us to both receive and impart knowledge. Therefore it is necessary for it to remain open and to rapidly integrate new words and concepts throughout the life span. While word learning in children has occupied researchers for decades, a number of studies have begun to shed light on the events that occur in the adult mental lexicon as a new word is encountered.

Recently, Davis and Gaskell (2009) have put forward an account of adult word learning, drawing on the principles of complementary learning systems (CLS) theories of memory (e.g., McClelland, McNaughton, & O'Reilly, 1995). According to this account newly learned words are initially stored as distinct episodic representations (thought to be reliant on medial temporal structures of the brain, the hippocampus in particular), isolated from existing bodies of knowledge, including the long-term lexical representations of the mental lexicon (mediated by neocortical structures). It is only later, as a consequence of either external repeated exposure over extended periods of time (i.e., spaced learning, Lindsay & Gaskell, in press) or internal replay of the episodic representations through memory consolidation during offline periods such as sleep (e.g., Dumay & Gaskell, 2007; Tamminen, Payne, Stickgold, Wamsley, & Gaskell, 2010), that stable cortical representations emerge, fully integrated in the mental lexicon. Given that the central distinction between early episodic representations and mature lexical representation concerns their integration in the lexicon, under this account one key diagnostic of a "lexicalised" word representation is its ability to engage with long-term stores of lexical knowledge (see also Leach & Samuel, 2007, for a similar argument). Below we briefly review existing evidence regarding the emergence of novel lexical representations in the light of this CLS framework.

Non-semantic diagnostics of lexical integration

Early studies showed that new items can be added to the mental lexicon after relatively little training in the laboratory. Salasoo, Shiffrin, and Feustel (1985) used visual threshold identification tasks to show that identification performance converged between real words and newly learned words just after five presentations of the new words, and that this was the case even one year after training. Novel words have also been shown to elicit identity priming after but not before training (Forster, 1985; Rajaram & Neely, 1992). This provides a useful measure of lexicality as nonwords are known not to reliably enable such priming. While these data inform us about the time course and manner in which a novel lexical representation is created in the lexicon, only more recent studies have begun to examine the degree to which these new representations are integrated with existing representations. This is significant because according to the CLS accounts, a new word that has become a part of the lexicon (as opposed to an isolated episodic representation) should in a measurable way influence the activity of the lexical units it is connected to, and vice versa.

One instance of interaction between words can be seen during spoken or visual word recognition, where the number of phonologically or orthographically overlapping neighbours influences the recognition time of a given word (Marslen-Wilson, 1987; McClelland & Rumelhart, 1981). Gaskell and Dumay (2003) showed that the introduction of a novel spoken word with no trained meaning (e.g., *cathedruke*) affects the recognition time of an overlapping existing phonological neighbour (e.g., *cathedral*), relative to control words (e.g., *dolphin*) for which no new overlapping competitor was trained. Interestingly, this effect was not observed immediately after training, but only in a delayed test some days later, as predicted by CLS accounts. These inhibitory effects provide evidence of novel words participating in the dynamics of spoken word recognition, suggesting that the novel word has been integrated into the lexicon. Bowers, Davis, and Hanley (2005) showed a similar effect using written words. Leach and Samuel (2007) approached the issue of lexical integration from a different perspective, taking advantage of the finding that ambiguous phonemes contained in familiar words, but not in nonwords, can retune phoneme categories after sufficient exposure (Norris, McQueen, & Cutler, 2003). Leach and Samuel found that newly learned words too can retune phoneme categories, although only in a training regime that provided the novel words with meaning.

Semantic diagnostics of lexical integration

The studies discussed above focused on lexical interaction that does not directly involve semantics (although Leach and Samuel's effect was dependent on training involving meaning). While such interaction is predicted by the CLS accounts, these accounts also predict that similar interaction should be seen at the semantic level. In addition, given that the goal of communication is to convey meaning, it is clearly crucial to understand the adult word learning process from the perspective of acquiring new meaningful words. Previous research has measured interaction between newly learned words and existing lexical-semantic knowledge through semantic priming measures.

Semantic priming (see McNamara, 2005, for an extensive review) refers to the finding that a target word (e.g., *doctor*) is recognised faster if it is preceded by a semantically related word (e.g. *nurse*) compared to an unrelated word (e.g. *teacher*). Such priming effects are usually taken to reflect the spread of activation in a semantic network: as the prime word is encountered, its semantic representation is activated and this activity spreads to semantically related representations which in turn will benefit from this preactivation when a related target word is encountered and activated (McNamara, 2005). Applied to the case of novel word learning the prediction would be that if a novel word has been integrated in such lexical-semantic networks, it should come to act as an effective prime when presented shortly before a related familiar target word.

The use of semantic priming as a diagnostic for lexical integration is complicated by the fact that priming effects can arise from at least two sources: semantic association and semantic relatedness (e.g., Lucas, 2000; Hutchison, 2003). Association refers to those instances in which two words are associated through co-occurrence, but are not semantically related (e.g., *mouse* and *cheese*). Semantic relatedness on the other hand relies on overlap of properties such as semantic features (e.g. *cherry* and *apple*). Therefore it is possible to observe priming between words that are associated in semantic memory through co-occurrence but not linked by virtue of their semantic properties such as semantic features. An extreme example of such cases is provided by studies looking at experimentally created associations. For example, Pecher and Raaijmakers (1999) taught participants semantically unrelated word pairs, and found in a subsequent test phase priming between the paired words (see McKoon & Ratcliff, 1979, for a similar finding, but also see Durgunoglu & Neely, 1987, for some boundary conditions for this effect). The implication of this is that a novel word might come to prime a familiar word if the two words were presented repeatedly together prior to the priming test, for example during a training phase. For example, if participants were taught that the novel word *feckton* refers to a "cat-like animal", *feckton* could be expected to come to prime the word *cat*. Such priming based on co-occurrence would provide only weak evidence of lexical integration, as the link between a novel word and a frequently co-occurring familiar word could be accommodated in the CLS framework by the episodic hippocampal system and does not require integration in the neocortical semantic network. The CLS accounts furthermore predict that such associative priming should become observable immediately after training as no consolidation is needed for episodic associations to be formed.

A third way in which semantic priming can arise is the case of mediated (or indirect) semantic priming (see Jones, 2010, for a review). This refers to the finding that processing of a target (e.g., *cheese*) can be facilitated by a semantically unrelated prime (e.g., *cat*) that is related indirectly through a mediator (e.g., *mouse*). Such a mechanism may be particularly relevant when seeking evidence of a novel word priming a related but unassociated existing word. For example, if one learns that *feckton* is a cat-like animal, in a subsequent priming task the activation of *feckton* would spread to the representation of *cat*, and spread further from *cat* to *kitten*. This would then give rise to priming effects between *feckton* and *kitten*. As far as we are aware, mediated priming has not been reported using newly learned words, but the plausibility of such a mechanism is supported by a study by Hayes and Bissett (1998) in which participants learned A-B and A-C nonword pairs (with no meaning provided for the nonwords). After training priming between the pairs was evaluated by presenting two words simultaneously on the screen, and asking participants to decide whether both words were trained nonwords ("yes" response), or whether one of them was an untrained nonword ("no" response). The authors found faster "yes" responses (i.e., priming) when the words on screen were trained pairs as well as when they were C-B and B-C pairs, an instance in which A probably acted as a mediator, compared to an unprimed condition in which one of the trained words had no association (direct or indirect) with the other trained word on screen.

Based on these considerations of semantic priming it is possible to conclude that priming between a novel word and a familiar word that was at an earlier stage associated with the novel word (e.g., through association during training) would not provide strong evidence for the novel word having been fully integrated in semantic network because such association could be supported by episodic hippocampal links. What is needed is a demonstration of a novel word (e.g. *feckton*) priming a semantically related familiar word that has never before been seen in association with that novel word (e.g. *kitten*). There is no way to definitively determine whether such priming would operate through direct priming (*feckton* triggering activation of *kitten,* a case of nonassociative semantic priming relying on featural overlap as long as *feckton* was never presented in association with *kitten* prior to the priming test) or mediated priming (*feckton* triggering activation of *cat*, which further triggers activation of *kitten*, a case which permits a combination of semantic associative and nonassociative priming, as *feckton* would have been associated with *cat* during training but not with *kitten*). However, in both cases priming would reflect semantic activity triggered by the novel word exerting an influence over units of the network which have not previously been associated with the novel words. This would constitute evidence in favour of the view that the novel word has been integrated in neocortical semantic memory. CLS models furthermore make the prediction that such priming should not be observed immediately after training, but only after a consolidation opportunity. Below we will review existing data on novel word learning and semantic priming, and argue that previously reported effects of novel words affording semantic priming are likely to reflect purely associative mechanisms that could be supported by the episodic hippocampal system.

Semantic priming in newly learned words

Dagenbach, Horst, and Carr (1990) demonstrated episodic semantic priming in a series of experiments in which participants learned the definition of obscure words (e.g., *drupe*, which refers to a fleshy fruit with a hard stone), followed by training on an episodic study list which paired a newly learned word with a semantically related familiar word (e.g., *drupe - cherry*). The newly learned words were found to prime lexical decision to the paired familiar words from the study list after extensive training over several weeks. This study however does not necessarily speak directly to the issue of new words being integrated into the existing lexicon, as the priming was between words that were episodically associated in a study list. Perfetti, Wlotko, and Hart (2005) taught participants the definition of 60 obscure words over a 45-minute training session that involved repeated presentation of the words and their definitions (e.g. *clowder* is a collection or group of cats). The training was immediately followed by a test phase in which participants carried out a semantic decision task while event-related potentials (ERPs) were recorded. The semantic decision involved presentation of a trained novel word and a semantically related familiar word (which was in fact in many cases a word that had appeared in the definition of the newly learned word), and required participants to decide whether the two words were related. Faster reaction times (RTs) were found for related than unrelated trials and the N400, an ERP component thought to index access to word meanings (see Kutas & Federmeier, 2011, for a review), showed higher amplitude in unrelated than related trials.

A similar result was reported by Mestres-Misse, Rodriguez-Fornells, and Munte (2007) in another ERP study, in which participants learned the meaning of a fictitious novel word through reading a sequence of three sentences using the novel word. An immediately following semantic decision task in which a novel word was paired with a semantically related (i.e., its own meaning) or unrelated word (i.e., a meaning of a different novel word) showed no behavioural advantage of relatedness, although there was an N400 effect similar to that obtained by Perfetti et al. (2005). A second study by Mestres-Misse, Camara, Rodriguez-Fornells, Rotte, and Munte (2008) on the other hand did find a behavioural effect in the same semantic decision task. A more commonly used form of priming was chosen by Breitenstein et al. (2007), who trained participants over five days on novel words by pairing them with different drawings of a familiar objects. Each word was assigned one drawing with which it occurred more frequently and consistently than any other drawing, thus allowing participants to deduce the meaning (drawing) of each word through statistical learning. At the end of the 5-day training period, cross-modal priming was tested by presenting participants with the drawings as targets of a living/non-living categorisation task, preceded by either the previously consistently paired or an inconsistently paired novel word. Priming was found after training, but not before.

We argue that priming based on a direct word-to-word association is likely to be the mechanism underlying many of the novel word priming results reported above. For example, as in the study by Breitenstein et al. (2007) the novel word and the drawing became strongly associated during training, the cross-modal priming seen in the test may well have been due to this association rather than the novel word having been integrated into existing lexical-semantic networks. The same issue applies to the other studies outlined above. In the Perfetti et al. (2005) study the related familiar target word was often part of the definition given to the novel word in training, and in Mestres-Misse et al. (2008) the novel word was always paired with its own definition. In other words, these studies show that a novel word can prime an existing word, but only when the novel word and the existing word have co-occurred during training. As discussed earlier, in CLS terms this link between a novel word and its trained definition can be accommodated by the hippocampal episodic memory system and does not necessarily imply that a novel lexical representation has been fully integrated in existing semantic memory beyond its trained definition.

Existing evidence along these lines comes from a study by Clay, Bowers, Davis, and Hanley (2007) who used the picture-word interference task to measure a novel word's influence on naming of semantically related pictures. In this paradigm the task is to name the picture of an object, while the picture is superimposed by a word that is supposed to be ignored. The typical finding is that picture naming times are slower if the superimposed letter string is a real word rather than a nonword. In addition to this lexicality effect, a further inhibitory effect is seen if the superimposed word is semantically related to the object. Applied to the issue of word learning, a newly learned word should interfere with picture naming only if it has been integrated in the lexicon and/or semantic memory, at least after consolidation. In this study participants learned 12 novel words and their meanings, presented both in the form of a drawing and a visually presented definition (e.g., *kosla* = bitter and spiky fruit). In training participants saw a novel word and a potential meaning (on any given trial either in the form of a definition or a drawing) and were asked to decide whether the pairing was correct (which it was half of the time), followed by feedback. There were 16 study trials for each word. Following the training phase, participants were tested by presenting them with drawings of familiar objects to be named. Each object was superimposed by a written word which the participant was to ignore. Clay et al. found the lexicality effect with the novel words immediately after training, suggesting that the novel words had developed lexical representations. However, the semantic interference effect using novel words superimposed on semantically related pictures was only found in their second test session which took place one week after training. This implied that the novel words had been eventually integrated into semantic networks, but that this integration required consolidation. Note that the novel words and the pictures of familiar objects that participants named in the test phase had not co-occurred prior to the test, therefore meeting our criterion for priming (in this case interference) without prior association.

However, a problem with this task is that it is difficult to know how much of the effect is driven by strategic processes and to what degree automatic spread of semantic activation contributes to the effect. Although no response is required to the superimposed word, it is nonetheless fully visible and there is nothing stopping participants from attempting to evaluate its relatedness to the picture. Furthermore, in the Clay et al. study each picture to be named was repeated 24 times in both test sessions, each time with a new superimposed word. This repetition makes it increasingly likely that attention may have shifted from the pictures to the words. An advantage of the semantic priming paradigm relative to picture-word interference is that the prime can be presented outside of consciousness by using a very short prime duration in combination with masking procedures (see our Experiment 2). This minimises potential strategic influences more reliably than in the picture-word interference paradigm. In Experiment 2 of the current report we take advantage of these techniques.

The current study

The main goal of the current study was to establish whether novel words become integrated into semantic networks in adult learners by using newly learned words as primes and semantically related familiar words had never before been associated with the novel words as targets. This is an important question as it is possible that newly learned words in laboratory conditions will not be treated as real words by the lexicon. Some accounts of adult word learning make exactly this claim. For example, Jiang and Forster (2001; see also Qiao, Forster, & Witzel, 2009) have suggested that second language (L2) words learned in adulthood can only be represented episodically outside of the lexicon. This was motivated by the finding that bilinguals showed *masked* L2-L1 priming in an episodic memory recognition task for studied targets, but no priming when the task tapped non-episodic memory (lexical decision). The authors argued that there is a point in time in development when new L2 words no longer enter the lexicon, possibly due to reduced neural plasticity. Qiao et al. (2009) extended this argument to novel word learning in L1.

A second aim was to elucidate the time course of lexical integration. Evidence from lexical competition (e.g., Dumay & Gaskell, 2007) and the picture-word interference task (Clay et al., 2007) suggest that newly learned words are initially stored separate from the lexicon and/or semantic networks, and become integrated with these existing stores of knowledge only after a period of memory consolidation, a process consistent with CLS accounts of memory. For this reason we tested participants immediately after training and after a short or a long delay.

We report two semantic priming experiments using novel word primes: one using visible primes that participants were explicitly aware of, and another using masked primes in which participants remained unaware of the presence or the identity of the primes. These two experiments were contrasted in order to vary the degree of automaticity of semantic access, with the first experiment allowing strategic effects to operate, while the masked version was intended to rely more heavily on automatic activation and to minimise strategic priming effects. Note that we do not claim that masked semantic priming is exclusively a result of automatic semantic activation and does not recruit strategic or conscious processing; we merely intend it as a task which is less subject to strategic influences than the visible primes version.

To summarise our hypotheses, if novel words become integrated in existing semantic networks, we predicted that we should see priming in both our experiments where the novel word primes and familiar word targets have not been associated prior to test. Furthermore, the CLS accounts of memory predict that such priming should not be seen immediately after training, but only after consolidation, as newly learned words are initially stored separate from the neocortical stores of lexical-sematic knowledge. Finally, observing masked semantic priming would for the first time provide evidence of integration of the new lexical item with existing semantic knowledge in the absence of exclusively strategic processing (cf. Clay et al., 2007).

Experiment 1

The primary measure of semantic integration in Experiment 1 was, as outlined above, semantically primed lexical decision to familiar word and nonword targets. In one version of the task the primes were real words; this condition was intended as a control condition to ensure our task parameters allow priming effects to be seen. This real prime version is also important in giving us an idea of the likely magnitude of the priming effect we can expect to observe with novel primes. Due to practical constraints in the selection of primes and targets in a word learning study such as this (see below for a detailed discussion), we were forced to accept prime-target pairs whose association strength was not as high as normally seen in semantic priming experiments. Therefore the real prime condition should be viewed as a "baseline" for quantifying the realistic magnitude of priming that can be expected to be seen with prime-target pairs of relatively low association strength. In the critical version of the task the primes were newly learned words. As discussed earlier, Clay et al. (2007) showed that integration of novel words with semantic neighbours may under some circumstances be observed only in a delayed test condition, after memory consolidation has occurred. Therefore we tested novel word primes that were learned either on the day of testing, henceforth referred to as "recent" novel words, or learned one or seven days before testing, henceforth referred to as "remote" novel words. As illustrated in Figure 1, we also varied the duration of the consolidation opportunity as it is unclear how long consolidation of novel word meanings might take: half of the participants had a 1-day consolidation opportunity, while the other half had a 7-day consolidation opportunity.

(FIGURE 1 ABOUT HERE)

Method

Participants. Sixty native English speaking participants drawn from the University of York and York St. John University student and staff populations participated in the experiment. Thirty were allocated to the short (1-day) consolidation opportunity group (11 male, one left-handed, mean age = 21.2, range = 17-42), and 30 to the long (7-day) consolidation opportunity group (seven male, three left-handed, mean age = 20.6, range = 18-28). No participants reported language disorders. Participants were paid or received course credit.

Materials and Design. Training of 34 novel words and meanings for all participants took place on day 1 (Figure 1). Half of the participants then returned on day 2 to learn another set of 34 novel words and meanings, followed by a test including both word sets. The other half returned one week after day 1 (on day 8) to learn the second set of novel words, again followed by the test session. The test session included two semantic priming tests; in one the primes were real (familiar) words, and in the other they were the trained novel words.

The novel word stimuli consisted of 102 written pronounceable nonwords[1] (e.g., *feckton*) created by Deacon, Dynowska, Ritter, and Grose-Fifer (2004), with average length of 6.5 letters. These nonwords were not derived from real words and were designed not evoke the meanings of real words, as confirmed by a norming study by Deacon et al. (2004). The set of novel words was divided into three lists of 34 words to be used in the two time-of-training conditions and in one untrained control condition (the untrained condition was not used in the tasks reported here). The words in the three lists were matched in number of letters (6.4, 6.4, and 6.5) and rotated through both time-of-training conditions. Similar sounding words were avoided within a list, to avoid confusability between words. All lists were equally often used in all conditions across all participants.

Each novel word was coupled with a meaning during training. A novel word meaning was made up of a familiar noun referring to an existing object, coupled with two semantic features that set the concept apart from already familiar concepts (e.g., participants would learn that "*feckton is a type of cat that has stripes and is bluish-grey*"), hence making the learning task more comparable to real world situations in which adults learn new names in their L1 for unfamiliar concepts. The novel word meanings (e.g., *cat*) were selected from the University of South Florida Free Association Norms (Nelson, McEvoy, & Schreiber, 2004). Each word in this corpus has a list of associated words, and each associate has a numerical association strength value based on the proportion of people who produced that associate in response to the stimulus word. Sixty-eight words were chosen from this set to act as the meanings assigned to the novel words. All selected meanings were nouns of medium frequency (M = 73.7), as listed in the CELEX database (Baayen, Piepenbrock, & van Rijn, 1995) with several associates. Three associates were selected for each meaning to act as targets in the priming task while avoiding overlap with other associates. For example, *cat* was selected as one of the meanings as it has several associates, three of which were selected to act as targets in the priming task (*dog, mouse, kitten*). The selection of three associates allowed the same novel word to act as a prime three times during the priming task, thus increasing the number of observations and statistical power without the need to train participants on a very large number of novel words. The associates of the meanings could be nouns, verbs, or adjectives (mean frequency = 151.6, mean number of letters = 5.8, mean association strength = 0.16). The 68 novel word meanings were divided into two lists of 34 meanings for use in the two time-of-testing trained conditions. The two lists were matched on frequency (72.3 and 75.0, mean frequency of associates = 150.0 and 152.8), mean association strength (both 0.16), number of letters (4.74 and 4.85), and number of letters of associates (both lists 5.77). Across all participants, both meaning lists occurred in both time-of-training conditions an equal number of times.

A further 34 nouns were selected from the free association corpus to be used as primes in the real word prime condition in the semantic priming task, acting as a baseline for the novel word prime condition. These primes were chosen based on the same criteria as the novel word meanings, except that these items were of lower frequency (mean frequency = 9.0, mean number of letters = 5.74, mean number of letters in associates = 6.02). The reason for choosing lower frequency primes for this control condition was to try to better match them with the novel word primes, which are bound to have extremely low frequencies (determined by the number of exposures in training). As before, three associates were selected for each prime (mean frequency = 66.0, mean association strength = 0.16) to act as targets in the priming task so that each real word prime too could be used three times. For example, for the prime *wasp* three associates were selected: *sting, bee, nest*. Care was taken in the selection of all items to ensure that there was no overlap: no word was repeated across the primes or the associates.

The primed lexical decision task required generation of 306 nonword targets (204 with novel word primes, 102 with real word primes). These were created by changing one letter in the real word targets to generate legal nonwords (e.g., *stong* derived from *sting*). This generated nonword targets that were carefully matched with the word targets, and likely increased task difficulty, maximising semantic influence of the primes (e.g., Joordens & Becker, 1997; Evans, Lambon Ralph, & Woollams, 2012). Roughly equal numbers of nonwords were created by changing letters in different positions of the words, with consonants always replaced with consonants, and vowels with vowels. As the order of presentation of trials in the lexical decision tasks was newly randomised for each participant by the presentation software, we did not control for the distance between the occurrence of a target words and the nonword that was derived from that target word. However, this randomisation means that the likelihood of a target word appearing in close proximity with its derived nonword was small and, more importantly, equal in the primed and unprimed conditions.

Procedure. ***Training session.*** Each training session consisted of four different training tasks: word-to-meaning matching, meaning-to-word matching, meaning recall, and sentence plausibility judgment. In the word-to-meaning matching task a novel word (e.g., *feckton*) was presented on the screen with two meaning alternatives. The participant was asked to indicate which of the two meanings was the correct one for the given word. The meaning was always presented in the form of "*is a type of X*" where X refers to one of the meanings described above (e.g., *is a type of cat that has stripes and is bluish-grey*). On each trial, one of the options was correct, and the incorrect option was randomly picked from the pool of meanings used in the current session by the experimental software. The allocation of the correct option to the left or right side was also randomly assigned for each trial by the software. After a response was made by pressing a key on a button box, the incorrect option disappeared from the screen, and the correct one remained on screen for 1500 ms and was followed by a new trial. Unlimited time was given to make the response. Participants were asked to pay close attention to the correct meaning remaining on the screen, in order to start learning the meanings at the beginning of the training. The meaning-to-word matching task was identical, except that in this task a meaning was presented in the top half of the screen, and two novel word alternatives were presented below. Participants were asked to pick the correct novel word for the given meaning.

In the meaning recall task a novel word was presented on the screen, and the task was to type in the complete meaning of the word. Unlimited time was given for responding. The correct answer was always presented after the response was completed. Order of trials in each block was randomized.

The fourth training task was the sentence plausibility judgment task. Here a sentence was presented on the screen, using one of the novel words (e.g., *"The girl was woken up by the paws of her hungry feckton"*). The participant's task was to indicate whether the novel word fit in the context of the sentence in terms of its meaning or not. Unlimited time was given to make a response. A response was followed by a feedback screen providing accuracy feedback, the novel word, and its meaning. Each novel word was presented four times in this task, always in a different sentence, three times in the correct usage condition, and once in the incorrect usage condition (e.g., *"The jubilant man raised his arm and punched the air with his feckton"*). This imbalance was intentionally built into the design to minimise the presentation of novel words in an incorrect context to avoid interference during learning. The presentation order of the sentences was randomized.

Training started with three blocks of word-to-meaning matching, with participants having no exposure to the novel words prior to this task, and with each novel word occurring once in each block. Participants then carried out one block of meaning recall with one exposure to each novel word, followed by two more blocks of word-to-meaning matching. At this point a second block of meaning recall was carried out, followed by three blocks of meaning-to-word matching. A last block of meaning recall was carried out before completing two more blocks of meaning-to-word matching. The final task of the session was sentence plausibility judgment, which included four blocks. Hence the total number of exposures to each novel word in this training regime was 17. The meaning recall task was interleaved with the other tasks as outlined above as it allowed the participants to identify words which they had not yet learned, and to focus on these words as training progressed. Participants were informed prior to training that they would be tested on the novel words later.

Test session. The test session followed the second training session. Participants were offered a chance to take a rest break before beginning the test. Test tasks consisted of meaning recall and primed lexical decision. Meaning recall was carried out first in order to allow participants to explicitly access their knowledge of the novel word meanings once before doing the primed lexical decision task. It was hoped that this would cancel out any possible recency effects that might benefit the recently learned novel words, and result in a purer measure of speed of access to novel word semantics. The meaning recall task required participants to type in the full meaning of a trained novel word presented on screen. No time limit was imposed, and no feedback was provided. Words from both training sessions were randomly intermixed.

The next task was primed lexical decision. Each trial began with the presentation of a fixation cross for 500 ms. This was replaced by the prime in lowercase letters for 200 ms, followed by a blank screen for 250ms, after which the target appeared also in lowercase letters for 200 ms. Timing started from the onset of the target, and 2000 ms was allowed for a response to be made. The participant's task was to decide whether the target was a real word or a nonword by pressing a key on a button box. Half of the participants responded to real words with their right hand, and the other half with their left hand. Once a response was made, feedback was provided both in terms of response accuracy and RT. The purpose of this feedback was to encourage fast and accurate responding. A rest break was offered half way through the trials, together with a summary of accuracy statistics. Participants were informed that in many trials the prime and target would be related, to encourage them to attend to the prime as well as the target.

The order in which the real word priming and novel word priming tasks were given was counterbalanced across participants. As stated earlier, because non-associative semantic priming effects are known to be small and require a large number of trials to be statistically reliable, it was necessary to use each novel word prime more than once to ensure enough data were collected. Hence both versions (real prime and novel prime) of the task were carried out in three consecutive blocks, with each prime occurring twice within each of the three blocks, once with a word target and once with a nonword target. Since there were an odd number of blocks, it was impossible to balance the primes so that each participant would see each prime an equal number of times with a semantically related and unrelated word target. However, the stimuli were counterbalanced so that across all participants each prime occurred in each priming condition an equal number of times, and within participants half of the primes appeared twice in the related condition and once in the unrelated condition, and the other half of the primes appeared once in the related condition and twice in the unrelated condition. For the targets a typical split-plot design was implemented in which each participant saw each word or nonword target only once, but across all participants all targets appeared in both priming conditions an equal number of times. This was achieved by randomly dividing the targets into two lists, and presenting one list with related primes and the other with unrelated primes. Unrelated prime-target pairs were created by rescrambling primes and targets. Novel word primes from both training sessions were intermixed, and the order of trials was pseudorandomised using Mix (van Casteren & Davis, 2006). At least 15 trials separated the repetition of any prime. A maximum of three consecutive trials from the same time-of-training or priming condition were allowed, and a maximum of four consecutive trials from the same lexicality condition were allowed. A new pseudorandomised order was generated for each participant.

Training and test tasks were carried out on a Pentium PC connected to a CRT monitor (17" Iiyama Vision Master Pro 454) with a high refresh rate (85 Hz). Manual responses were collected by a Cedrus button box. E-prime (Psychology Software Tools Inc. www.pstnet.com/eprime) was used as the experimental software in all tasks.

Results

All data reported henceforth were analysed using mixed-effects modelling (Baayen, 2008; Baayen, Davidson, & Bates, 2008; Janssen, 2012) with subjects and items entered as random effects, thus eliminating the need for separate by-items and by-subjects analyses. The reaction time data were analysed using the MIXED procedure in SPSS 19 (using the maximum-likelihood procedure), and we report type III tests of fixed effects. Random slopes were included when model comparison indicated that they significantly improved the fit of the model (measured by comparing the log-likelihood values of competing models[2]), and their inclusion is reported in the text when appropriate.

Explicit recall of novel word meanings in the test session. As recommended by Jaeger (2008), explicit recall of novel word meanings was analysed with mixed-effects logistic regression, as these data are binary (a word meaning will either be recalled or not). Mixed-effects logistic regression has not been implemented in the MIXED procedure in SPSS 19, therefore these data were analysed using the *lmer* procedure in R (see Jaeger, 2008). In all explicit meaning recall tasks reported in this paper a response was scored as correct if the participant recalled the object[3] to which the novel word refers to. Figure 2 shows the proportion of meanings recalled correctly by both participant groups in the test session. A model with the fixed factors of time-of-training (remote vs. recent) and length of consolidation opportunity (1-day vs. 7-day) showed that an interaction between the factors significantly contributed to the model, $\chi^2(1) = 14.5$, $p < .001$. Interaction contrasts showed that both the 1-day-consolidation group ($z = 14.6$, $p < .001$) and the 7-day-consolidation group ($z = 20.4$, $p < .001$) recalled more recent meanings (i.e., words learned immediately before test) than remote meanings (i.e., words learned one or seven days before test). Furthermore, while both groups recalled equally many meanings that were trained immediately before test (recent), the 1-day-consolidation group recalled more remote meanings than the 7-day-consolidation group, $z = 3.1$, $p = .002$.

(FIGURE 2 ABOUT HERE)

Lexical decision with real word primes. Only accurate responses were analysed. As recommended by Baayen (2008), RTs faster than 150 ms or slower than 1500 ms were removed as extreme outliers, followed by a log transformation in order to reduce the effect of remaining outliers and to reduce the positive skew typical of RT data. The same data trimming procure was used in all RT analyses reported in this paper.

Table 1 shows mean RTs to word targets in the primed and unprimed trials with real word primes. We entered priming (primed vs. unprimed) and length of consolidation opportunity group (1-day vs. 7-day) as fixed factors. The interaction between the two did not significantly benefit the model, and consolidation length had no significant effect, showing that both groups exhibited similar priming effects when the primes were real words. Priming showed a significant main effect whereby RTs to target words were faster in the primed condition than in the unprimed condition, $F(1, 5350) = 43.81$, $p < .001$.

(TABLE 1 ABOUT HERE)

Error rates for this task are also presented in Table 1. A mixed effects logistic regression analysis used the same factors as the RT analysis and revealed no interaction or an effect of consolidation length group. There was an effect of priming with primed trials yielding significantly fewer errors than unprimed trials, $\chi^2(2) = 22.8$, $p < .001$.

Lexical decision with novel word primes. Table 1 shows the mean RTs to word targets in each priming and time-of-training condition. A model with the fixed factors of priming (primed vs. unprimed), time-of-training (remote vs. recent) and length of consolidation opportunity (1-day vs. 7-day) revealed no interactions that contributed significantly to the model. While the main effect of consolidation length also proved non-significant, the model showed a significant main effect of time-of-training whereby responses to trials with recent primes were significantly slower than trials with remote primes, F(1, 10807) = 5.78, p = .016. Critically, a main effect of priming showed that RTs to primed trials were significantly faster than RTs to unprimed trials, $F(1, 10810) = 5.84$, $p = .016$. Although the interaction between priming and time-of-training was not significant ($F = 2.50$), the data in Table 1 suggest that the priming effect was numerically larger for remote than recent primes. Hence we also evaluated the priming effect separately for remote and recent primes, and found that a significant priming effect was found only with remote primes, $F(1, 5307) = 7.19$, $p = .007$ (for recent, $F(1, 5276) = 0.40$, $p = .53$).

Error rates are also presented in Table 1, and were analysed as above, using mixed-effects logistic regression. This analysis revealed no statistically significant effects.

Discussion

The key finding from Experiment 1 was that novel words can prime semantically related familiar word targets to such a degree as to facilitate lexical decision to those targets, as demonstrated by a significant main effect of priming, despite the fact that these stimuli were not directly associated or otherwise explicitly linked during training. This result goes beyond previous studies looking at the generation of new lexical semantic representations, as most of those studies have included a degree of episodic association. Priming effects following direct association can be accommodated by the formation of a single new associative link (e.g., Pecher & Raaijmakers, 1999) and need not depend on the main store of lexical knowledge. Here, however, the facilitation effect suggests that the semantic influence of the novel word generalises to the wider lexical environment of the novel word's new associate, suggesting that the novel word has been incorporated into full semantic network.

We did not find a statistically reliable interaction between time-of-training and priming, suggesting that even those novel words that were learned immediately before testing may have allowed some degree of semantic priming. It is nonetheless worth noting that the priming effect was nonsignificant for recently learned words and only showed significant priming in the remote condition, suggesting that the experiment may have been underpowered for the detection of the interaction. This numerical trend however is consistent with the hypothesis that semantic integration may benefit from a period of offline consolidation, as concluded by Clay et al. (2007). In Experiment 2, we show that this consolidation effect is replicable. There are two alternative explanations for the finding that priming was restricted to remote words however. First, it might be due to fatigue effects that may have impaired learning in the second (recent) training session. However, this explanation is not supported by explicit recall of novel word meanings, as this was superior for the recent words, suggesting that their meanings were learned reliably. Also, even in the short consolidation opportunity group participants had 24 hours to recover from fatigue. Second, these effects might be due to proactive interference, whereby the words learned in the first training session (remote) may have interfered with learning of the words in the second training session (recent). While our data do not allow us to conclusively reject this possibility, the explicit recall data again do not suggest any disadvantage for learning in the second training session.

The real word prime condition was included to give us a baseline of the magnitude of the priming effect we could expect with stimuli in which the prime-target association strength is quite low, and the primes are of low frequency. Although the novel word priming effects are small, in the remote condition (7 ms) they are almost half the size of the priming effect in the real word prime condition (18 ms). The real word items will have typically been encountered numerous times over the preceding years. Furthermore, as the real word prime-target pairs were derived from free association norms, the stimulus set is likely to include items that are semantically related and associated. Semantic association may give priming a “boost” (Lucas, 2000), something that the novel word prime-target pairs would be lacking. In this light the novel word priming effects are close to the maximum level that could reasonably be expected. The low prime-target association strength was unavoidable as we needed three different associates for each novel word meaning in order to have enough data points to provide sufficient statistical power for detecting a priming effect, and very few words have three strong associates. So far we have discussed association strength in terms of forward association strength (FAS), i.e., the proportion of participants in word association norms who produce the target in response to the prime. Another factor that affects priming strength in the lexical decision task is backward associative strength (BAS) which refers to the proportion of participants who produce the prime when cued with the target (see Kahan, Neely & Forsythe, 1999; Thomas, Neely & O'Connor, 2012). While our real word and novel word priming conditions were matched in FAS, we did not match them in BAS. In fact, BAS was much lower in the real prime condition (0.06) than in the novel prime condition (0.16). Hutchison, Balota, Cortese, and Watson (2008) showed that higher BAS predicts stronger priming effects, at least with a long SOA. This may partially explain our relatively small priming effects in the real word prime condition. Finally, considering that the novel word training was not nearly as extensive as in some earlier studies and took place in one session lasting no more than two hours, the emergence of a statistically reliable priming effect with novel word primes is cause for increasing confidence in the power and validity of laboratory-based novel word learning paradigms.

It is also instructive to contrast these priming data with participants' explicit recall of the novel word meanings. It was clear (see Figure 2) that participants forgot a significant number of words between the first training session and the test session. Furthermore, the longer this delay was, the more they forgot. What is surprising is that this explicit forgetting did not weaken the priming effect, quite the opposite. Remote novel words which were subject to forgetting showed, if anything, numerically stronger priming effects than the words learned immediately before testing. This implies that the semantic priming effect here was independent of explicit recall of the meanings, and that the priming task tapped into more implicit semantic activation. Our demonstration of declining explicit recall rates is informative also with respect to the issue of whether participants rehearsed the novel words between training and testing. For example, the picture-word interference effects Clay et al. (2007) showed emerging over time may have been facilitated by rehearsal prior to their second test. In our case rehearsal appears unlikely as explicit recall declined significantly, despite participants knowing that they would be tested in the later session. In sum, the central finding from Experiment 1 was the significant main effect of priming with novel word primes, suggesting that novel words have been integrated into semantic networks. In the next experiment we sought even stronger evidence of this by minimising strategic influences in the priming task.

Experiment 2

The priming effect obtained in Experiment 1 was with clearly visible primes and participants were encouraged to make use of the prime in responding to the target as quickly as possible, hence allowing strategic influences to operate. Relevant to this issue is a study reported recently by Qiao et al. (2009) who pointed out that masked form priming is typically only seen when the prime is a nonword (contrapt-CONTRACT), and not when the prime is a real word (contrast-CONTRACT), an effect known as the prime lexicality effect. Therefore an untrained novel word should prime an orthographically overlapping familiar word, while a fully integrated novel word should not. However, they found priming both with trained and untrained novel words, leading them to argue that the novel words, which were borrowed from Bowers et al. (2005) and trained in an identical way, were represented in episodic memory and had not been integrated into the lexicon, at least under the particular training regime used in their experiment. Interestingly, Elgort (2011) did find the prime lexicality effect using newly learned words. However, there the prime duration was very long (522 ms compared to Qiao et al's 50 ms), allowing for the possibility that strategic influences may have contributed to this effect and hence it may reflect novel words residing in episodic (hippocampal) rather than in lexical memory. Qiao et al. (2009) argued that in order to best probe lexical memory, strategic effects should be excluded by using masking procedures, allowing the prime to be presented outside of awareness. Therefore, in Experiment 2, we attempted to find a semantic priming effect using masked novel word primes instead of fully visible primes, and used a much reduced stimulus onset asynchrony between the prime and target (450 ms in Experiment 1, 47 ms in Experiment 2). If a priming effect was found even under these demanding conditions that minimise strategic influences, it would provide us strong evidence that the novel words are represented in the lexicon rather than in episodic memory, and that the novel words have been integrated into semantic networks.

The use of masked semantic priming (Marcel, 1983) as the primary task measuring semantic integration provided an opportunity to enhance the design of our study and optimise statistical power. In standard priming experiments the explicit nature of the prime would provide a clear instance of the novel word and thus repeated testing on the same primes over time would be expected to result in artificially inflated priming effects. On the other hand, if one is interested in taking into account the emergence of the priming effect over time, then testing the same participants repeatedly provides statistically the most powerful design. In masked priming experiments the problem of repeated exposure is alleviated as the masking means participants are unable to read the prime, and most of the time remain consciously unaware of its existence. In order to take advantage of the power offered by masking and within-participants designs, in Experiment 2 all participants were initially trained and tested immediately on day 1, followed by a second test on day 2, and a third test on day 8 (Figure

3). Hence we could track the emergence of the priming effect in each participant as a function of time. An explicit meaning recall task was retained for approximately half of the participants. This was done to rule out effects of explicit testing on priming, as explicit recall might act as additional training.

(FIGURE 3 ABOUT HERE)

Method

Participants. Forty-eight native English-speaking students from the University of York took part in the experiment (8 male, 9 left-handed, mean age = 20.0, range = 18-40). No participants reported language disorders, or had participated in Experiment 1. Participants were paid or received course credit.

Materials and Design. The experiment was carried out on three days, spanning one week (Figure 3). On day 1, participants were trained on the novel words. This was immediately followed by a test session on all items. For about half of the participants (n = 21) a test session consisted of the explicit meaning recall task, followed by primed lexical decision using real (familiar) word primes in one task and novel word primes in another, and identical test sessions were attended on day 2 and day 8. For the remaining participants (n = 27) the test session on day 1 and day 2 included only the primed lexical decision tasks, and the day 8 session included the lexical decision tasks followed by a meaning recall task. This allowed us to see if the repeated administration of the meaning recall task was needed to maintain memory of the word meanings, with potential consequences for explicit recall and semantic priming. This would appear to be the case if priming were found only in the group of participants who experienced the meaning recall task in the beginning of each testing session.

For the purposes of the primed lexical decision task a split-plot design was used, and the same counterbalancing measures were taken as in Experiment 1, to ensure that all targets appeared in primed and unprimed conditions across participants, while each individual participant saw each target only once. As in Experiment 1, the task consisted of three blocks with each prime occurring in each block, and repeated twice within each block, once with a word target and once with a nonword target. Note that due to experimenter error two novel word primes were removed from the lexical decision analyses because they were paired with wrong targets.

The materials were drawn from the stimulus pool used in Experiment 1. However, since only one set of words was trained in the current experiment, only 34 novel words and 34 meanings were needed. Novel words were selected so that the full range of word lengths from the original set was represented (M = 6.4 letters, range = 5-8), and that the words were as dissimilar as possible from each other to minimise their confusability. A set of 34 meanings and corresponding associates was also selected (frequency = 49.4, frequency of associates = 99.6, mean length = 6.0 letters, mean association = 0.18). The same set of 34 real word primes and their three associated targets used in the real prime condition of Experiment 1 was used here. Each novel word was paired with a meaning, taking care not to pair meanings with words that sounded like the novel word.

Procedure. ***Training session.*** The procedure of the training session was identical to that used in Experiment 1, with meaning-to-word matching, word-to-meaning matching, meaning recall, and semantic plausibility tasks. The number of exposures was also the same as before (17 in total).

Test session. Meaning recall in the test sessions was identical to Experiment 1, with a novel word presented on screen and responses typed on the keyboard, without a time limit or feedback. The order of trials in the primed lexical decision task was pseudorandomised individually for each participant in each test session using the same constraints as in Experiment 1. A trial started with the presentation of a mask (##########) for 500 ms at the centre of the screen. The number of #s was the same on each trial, and was determined by the length of the longest stimulus used in the experiment (i.e., ten letters). The prime word in lowercase letters appeared at the offset of the mask for 47 ms, and was then replaced by the target presented in uppercase letters. The target remained on screen until a response was made, or until 2000 ms had elapsed. At this point accuracy and RT feedback was presented. A key press initiated a new trial with a delay of 500 ms. Participants were not told about the presence of the prime, and were instructed to make the lexical decision as quickly and as accurately as possible by pressing a key on the button box. Half of the participants made word responses using their right hand, and the other half used the opposite mapping. At the end of the last testing session participants were asked if they had noticed the prime on the screen between the presentation of the mask and the target. Out of 48 participants only seven reported noticing that a word would sometimes appear briefly, but none reported being able to read it. Although this does not mean that the primes were fully outside of consciousness (e.g., Kouider & Dupoux, 2004), it does suggest minimal opportunity for strategic processing. The same equipment and software was used as in Experiment 1.

Results

Explicit recall of novel word meanings. Figure 4 shows the accuracy data in the meaning recall test task for proportion of objects recalled. These data are shown both for the group which was tested in the beginning of each test session, and for the group which was tested only once, in the end of the last test session. Data from the repeatedly tested group were analysed first. A mixed-effects logistic regression model with time-of-testing (day 1, day 2, day 8) as the fixed factor was fitted. Comparing the model with the fixed effect to an empty model with no fixed effect showed that time-of-testing had a significant effect, $\chi^2(2) = 87.0$, $p < .001$. Planned comparisons between the three days showed that while no significant difference was seen between day 1 and day 2, accuracy rates on day 8 were significantly lower than either on day 1 ($z = 7.6$, $p < .001$) or day 2 ($z = 6.7$, $p < .001$). Next, the difference between the repeatedly tested and once only tested groups on day 8 was compared, using an identical model as above but adding testing group as a fixed factor and removing the factor of day. The factor of testing group contributed significantly to the model over an empty model, $\chi^2 (1) = 7.0$, $p = .001$, confirming that the repeatedly tested group recalled more word meanings.

(FIGURE 4 ABOUT HERE)

Lexical decision with masked real word primes. Table 2a shows RTs in the real word prime condition. The analysis was carried out with priming (primed vs. unprimed), time-of-testing (day 1, day 2, day 8), and presence of repeated semantic test of novel words (test vs. no test) as fixed factors. Subject-specific random slopes for the factor of time-of-testing significantly improved the fit of the model. No significant interactions between the fixed factors were found (largest $F = 2.40$). Of the main effects only the critical main effect of priming was significant, $F(1, 13268) = 23.85$, $p < .001$.

Error rates are also shown in Table 2a. No effects reached significance in a mixed-effects logistic regression model with priming and time-of-testing as fixed factors.

(TABLE 2A ABOUT HERE)

Lexical decision with masked novel word primes. Table 2b shows data from the primed lexical decision task with novel word primes. The analysis was carried out as before, with priming (primed vs. unprimed), time-of-testing (day 1, day 2, day 8), and presence of repeated semantic test (test vs. no test) as fixed factors. Subject-specific random slopes for the effect of time-of-testing significantly benefitted the model, and were therefore included. No interactions between the fixed effects reached significance. Importantly, the main effect of priming did reach significance, with RTs faster to primed trials compared to unprimed trials, $F(1, 12519) = 9.89$, $p = .002$.

Although there was no significant interaction between priming and day of testing, suggesting that the priming effect was equally strong on each day, we also evaluated the priming effect separately on each day. The effect reached significance on day 8, $F(1, 4146) = 12.21$, $p < .001$, but not on day 1, $F(1, 4132) = 2.55$, $p = .11$, or day 2, $F(1, 4187) = 0.16$, $p = .69$.

Error rates are also presented in Table 2b. The mixed-effects logistic regression analysis with the same fixed and random effects as the RT analysis showed no significant effects of any of the fixed factors or their interactions.

(TABLE 2B ABOUT HERE)

Discussion

In Experiment 2 we wished to see if activation of a meaningful novel word allowed automatic facilitation of the recognition of semantically related familiar words by using a masked priming paradigm. As demonstrated by a significant main effect of priming, this was found to be the case, again under circumstances where there was no direct association between the novel word and the semantically related familiar target words. As the prime was masked and presented extremely briefly, it is unlikely that it was available for conscious report and in this way potential strategic effects were minimized. As such, this experiment provides the strongest evidence to date of novel words having been integrated into existing semantic networks. Thus these data run counter to Qiao et al.'s (2009) arguments that novel words learned in laboratory circumstances with a brief training regime remain separate from the mental lexicon. The precise reason for the discrepancy between our data and those of Qiao et al. is difficult to identify because they used a different test for lexicalisation. It is however possible that their training was not extensive enough, either with regard to the meanings of their novel words, or to the amount of orthographic exposure (see McCormick & Rastle, 2012, for evidence that this is the case).

As in Experiment 1, we did not find a reliable interaction between time-of-testing and priming, suggesting that the priming effect may have been present to some degree from the first test session. However, again following the pattern seen in Experiment 1, the effect was numerically strongest after a consolidation opportunity. In fact, when evaluated in each test session individually, the priming effect failed to reach significance on day 1 and day 2, but was statistically significant on day 8. This may imply that the effect benefits from consolidation; in particular as a trend supporting the consolidation hypothesis was also seen in Experiment 1.

The same caveats from Experiment 1 concerning the size of the priming effect need to be considered here. The real word prime condition was again used to establish the likely magnitude of the priming effect with stimuli that have relatively low prime-target association strength and low frequency primes. It is striking that by the end of the experiment the size of the priming effect in the novel word condition (8 ms) is almost identical to the real word condition (9 ms). This suggests that the novel words produced priming as strong as that for familiar words, a surprising finding once again considering the relatively brief training involved. Furthermore, the size of our priming effect is not out of line compared to other masked semantic priming data reported in the literature using familiar word primes. For example, Balota, Yap, Cortese, and Watson (2008) recently reported an experiment with similar prime duration to ours (42 ms) that showed a statistically significant masked priming effect of 13 ms in lexical decision, despite a much stronger average forward association strength (0.66) than in our stimulus set (0.16 in the familiar word prime condition, 0.18 in the novel word prime condition). In this experiment, as in Experiment 1, backward association strength was not matched between the real word and novel word prime conditions (0.20 in novel word primes, 0.06 in real word primes), although BAS may have a smaller effect in a short SOA experiment such as the current one (Hutchison et al., 2008).

Unlike Experiment 1, in Experiment 2 participants were tested repeatedly on three occasions in the priming task. Could the repeated exposure to the prime-target pairs in this test have created an episodic association between the related primes and targets that might account for the priming effect? We believe this not to be the case, as such episodic link would also form between unrelated primes and targets (which were also repeated), hence cancelling out any priming effect. This would not allow us to observe the significant priming effect which we report here. Furthermore, if repeated exposure accounted for the late emergence of the priming effect we would expect to see the effect growing stronger over time also in the real word prime condition. There was however no evidence of priming in this condition being stronger on day 8 than on the earlier test times.

Semantic priming was unaffected by the manipulation of the explicit test: the presence or absence of the test did not interact with priming in either the novel word or familiar word prime conditions. The lack of an influence of explicit meaning recall in this experiment also shows that the priming effect observed with remote novel word primes in Experiment 1 could not have been caused by the presence of the explicit recall test. The rate of accurate explicit recall of novel word meanings on the other hand was unambiguously affected by the presence of the recall test in each session. Participants who were tested recalled significantly more novel word meanings in the final session compared to those who were not tested prior to day

8. This suggests that repeated administration of the recall task maintained explicit memory of the meanings and protected it from forgetting. A similar phenomenon was reported by Roediger and Karpicke (2006) who showed that repeated testing of newly memorised materials immediately after study resulted in increased recall rates several days after study relative to just increasing study trials without providing repeated testing. As in Experiment 1, we found no evidence of the priming effect decaying with forgetting. Instead numerically the strongest priming effect was seen in the last session in which also the explicit recall rates were the lowest.

Combined analysis of Experiments 1 and 2

As discussed above, we failed to find an effect of time-of-testing on priming, despite numerically larger priming effects on day 8 compared to day 1. The same numerical pattern was seen in Experiment 1 although the interaction failed to reach significance there too. Given the subtlety of the priming effects, it is however possible that we did not in either experiment have enough statistical power to reliably observe the subtle consolidation effect. We therefore carried out an analysis combining the data from the novel word prime conditions of Experiment 1 and 2[4]. For this analysis we contrasted priming using novel words learned immediately before the test, and novel words learned one week earlier (as this is the strongest time-of-testing manipulation). Therefore we combined data from the two experiments, coding the recent primes condition of Experiment 1 (long delay group only) and the day 1 priming data of Experiment 2 as “recent”, and the remote primes condition of Experiment 1 (long delay group only) and the day 8 priming data of Experiment 2 as “remote”. All participants were included in the analyses (including participants in Experiment 2 who whose meaning recall was repeatedly or only once tested). The resulting factor of time-of-testing was entered into a mixed-effects model as a fixed factor, together with the factors of priming (primed vs. unprimed) and experiment (experiment 1 vs. experiment 2). Subjects and items were entered as random factors, and subject-specific slopes were fitted for priming, time-of-testing, and their interaction. We observed a significant effect of priming, $F(1, 13639) = 6.84$, $p = .01$, and a significant effect of experiment, $F(1, 79) = 21.15$, $p < .001$. The latter was caused by overall RTs being slower in Experiment 2 than in Experiment 1 (see Tables 1 and 2b). Crucially, the interaction between priming and time-of-testing also reached significance, $F(1, 13595) = 3.92$, $p = .048$, while no other interactions were significant. To further examine the priming x time-of-testing interaction we analysed data separately from the remote and recent conditions, with priming and experiment as fixed factors. In the remote condition we found a significant effect of priming $F(1, 6758) = 10.75$, $p = .001$ and experiment $F(1, 80) = 16.63$, $p < .001$, and no interaction between the two. In the recent condition the effect of priming remained non-significant, $F(1, 6733) = 0.20$, $p = .66$, while the effect of experiment was again significant, $F(1, 80) = 23.78$, $p < .001$. No interaction was found. These results suggest that the numerical pattern of priming increasing over time observed in both experiments is statistically reliable given sufficient statistical power.

General Discussion

We have presented what we believe to be the strongest evidence to date of laboratory-trained novel words becoming engaged in the existing semantic network of related words in adult learners. We used newly learned words as primes in two semantic priming experiments which required a lexical decision to familiar target words that could be semantically primed or unprimed by the novel word meanings. Experiment 1 used visible primes and participants were encouraged to attend to the primes to take advantage of the related primes. We found a statistically significant main effect of priming, suggesting that the novel word meanings had been linked with related concepts. Experiment 2 provided even stronger evidence of this, as there we used masked novel word primes of which most participants were not aware, or at a minimum were not able to read consciously. A statistically significant main effect of priming effect was found even under these demanding conditions, again suggesting that the novel word meanings had been integrated into semantic networks. While earlier studies using priming paradigms have shown that newly learned words can prime words or pictures they were paired with during training, those studies can be accommodated by assuming that a single associative link is formed between a novel word and its given meaning during training. In the CLS framework, such a link could be accommodated by the short-term hippocampal system and does not necessarily imply that the novel word has been integrated with existing knowledge in the long-term neocortical store. Our study goes further, in that as far as we know, it provides the first demonstration of priming between novel words and semantically related but unassociated existing words.

Although the mechanism underlying our novel word priming effects may rely on either direct or indirect (mediated) priming (or a mixture of both), in both cases this result implies a greater level of semantic integration for the novel words in that they must be able to utilise the existing semantic structure (e.g., associated forms) of the newly acquired meanings. In any case, there is reason to believe that mediated priming plays a relatively small role in our novel word priming condition. This is because the size of the novel word priming effect was close to the real word priming effect, which is likely to rely more on direct priming. This is relevant because mediated priming effects tend to be smaller than direct priming effects, particularly in lexical decision (Balota & Lorch, 1986). Secondly, masked mediated semantic priming has not been found, to our knowledge, in a lexical decision task (see Pecher, Zeelenberg, & Raaijmakers, 2002, for data from a perceptual identification task), again suggesting that our masked priming effect is unlikely to rely heavily on mediated priming.

Our masked priming data are also an important finding in light of the growing interest in laboratory techniques for word learning. It has been argued, for example by Qiao et al. (2009; see also Jiang & Forster, 2001), that newly learned words trained in an experimental context may be stored in episodic memory rather than in the lexicon, at least unless extensive training is provided. Our data are encouraging in suggesting that, at least when the novel words are trained with meaning, they become embedded with other lexical semantic representations and interact with them even when these representations have never been explicitly linked before (during training for example). Furthermore, Experiment 2 showed the priming effect when the prime was masked, replicating the effect under circumstances where there is minimal opportunity for strategic processing which might tap into episodic (hippocampal) knowledge..

Related evidence comes from a study by Elgort (2011) who also found semantic priming using novel word primes, but in a variant of the priming task in which a lexical decision is made to all stimuli, including the primes, which are fully visible. Also of note are the data reported by Clay et al. (2007) discussed earlier, but as the picture-word interference task involves overt presentation of the to-be-ignored novel word distractor, we argue that our masked priming paradigm offers still stronger evidence of automatic activation, although we acknowledge that even in this task the extent of automaticity is not generally agreed upon (see Van den Bussche, Van den Noortgate, & Reynvoet, 2009, for a review). Nonetheless, the demonstration of such semantic integration effects provides an important addition to the earlier demonstrations of lexical integration effects, as shown in studies looking at lexical competition (Gaskell & Dumay, 2003; Bowers et al., 2005; Dumay & Gaskell, 2007; Tamminen & Gaskell, 2008; Davis, Di Betta, Macdonald, & Gaskell, 2009; Fernandes, Kolinsky, & Ventura, 2009; Tamminen et al., 2010), and perceptual learning (Leach & Samuel, 2007).

In both experiments we measured priming both immediately after training when the novel words were likely to be in an unconsolidated state, and after one or seven days, when the words had had an opportunity to begin undergoing extensive offline memory consolidation. Although neither experiment showed a statistically significant interaction involving priming and time-of-testing/training when analysed separately, in both experiments the priming effect increased numerically with time and the interaction between priming and time-of-testing was significant when data from the two experiments were combined. This is consistent with CLS accounts of memory and the data reported by Clay et al. (2007) whose interference effect specific to novel word meaning was only found in their second test session which took place a week after training.

The idea that novel word meanings benefit from time to integrate in the semantic system may seem to be contradicted by a number of recent studies reporting various word-like effects in new words learned only minutes earlier. For example, Shtyrov, Nikulin, and Pulvermuller (2010) found a word-like EEG response to novel words trained for only 10 minutes. Borovsky, Kutas, and Elman (2010) and Mestres-Misse et al. (2007) showed word-like N400 effects with novel words immediately after training. While these studies provide interesting insight into the very first moments of word learning in adults and the impressive speed with which novel words begin to behave like established familiar words, it would be premature to conclude that no further developments take place during the hours and days that follow. It is possible that the novel words at this very early time following training were stored in the form of transient, possibly episodic, memory traces rather than in the established lexicon.

This view is supported by a study by Batterink and Neville (2011) which looked at the N400 in semantic priming using newly learned words as primes. These authors found no behavioural or N400 priming effect immediately after a training task which involved ten exposures to the novel words in a story context. In contrast, they did find an N400 effect in a task measuring explicit recall of novel word meanings (cf., Jiang & Forster, 2001), supporting the dissociation we observed between an initially weak semantic priming effect growing more reliable over time, and a parallel decay of explicit recall of novel word meanings. To evaluate the degree to which novel words have been integrated into the lexicon one needs evidence of integration at later time points, such as provided in the current experiments, or retests after extended periods of time such as weeks or months (Tamminen & Gaskell, 2008). The delayed lexical competition effects reviewed earlier and the preliminary evidence from the priming experiments here that showed the most reliable priming only after a delay lead us to suggest that novel words integrate in the lexicon and semantic networks over time, possibly as a result of hippocampal-neocortical dialogue as suggested by complementary learning systems theories (see Davis & Gaskell, 2009, for a review).

Further evidence for a qualitative change over time in newly learned linguistic information comes from a study reported by Tamminen, Davis, Merkx, and Rastle (2012). Here participants were trained on novel words made up of familiar stem words and a novel affix (e.g., *-nule* in words like *buildnule* and *climbnule*). These authors found in a speeded shadowing task that after training participants processed words with a trained affix (e.g. *buildnule*) faster than words with an untrained affix (e.g. *buildhoke*), suggesting that affix learning had occurred. However, when tested immediately after training, this evidence of learning was restricted to stimuli where the stem had also occurred during training (e.g., *buildnule*). Two days after training the learning effect extended to stimuli where the stem had not been used in training (e.g. *sleepnule*). This suggests that the new affix representations became generalisable to stem contexts that the participant had not encountered during training only after a consolidation opportunity and that novel affix representations require time to become abstract. In a striking similarity to the contrast we observed between gradually declining explicit memory for our word meanings and gradually emerging priming effects, suggesting that priming effects were independent of participants' explicit memory of the novel word meanings, Tamminen et al. (2012) found that the process of abstraction was independent of participants' explicit memory of the novel affixes in that recognition memory for the novel affixes was very high immediately after training, in the absence of evidence for abstraction, as well as two days later when abstraction finally was observed.

In sum, we have demonstrated for the first time that newly learned words can prime semantically related familiar words in both masked and unmasked semantic priming paradigms. These effects were found in spite of declining explicit recall over time of the novel word meanings, and suggest that the novel words had been integrated with existing words and conceptual knowledge. As research into adult word learning in L1 gathers momentum (see Lindsay & Gaskell, 2010, for a discussion of how this research might translate into L2 word learning), it is vital to demonstrate that words learned as a result of brief training in laboratory conditions become a part of the established lexicon, as opposed to being represented as transient short-term memory traces. Our data provide an important advance in this endeavour, while at the same time highlighting the remarkable plasticity of the adult lexicon.

References

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R.* Cambridge: Cambridge University Press.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59,* 390-412.

Baayen, R. H., Piepenbrock, R., & van Rijn, H. (1995). The CELEX Lexical Database (CD-ROM). Linguistic Data Consortium, University of Pennsylvania, Philadelphia, PA.

Balota, D. A., & Lorch, R. F. (1986). Depth of automatic spreading activation: Mediated priming effects in pronunciation but not in lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12,* 336-345.

Balota, D. A., Yap, M. J., Cortese, M. J., & Watson, J. M. (2008). Beyond mean response latency: Response time distributional analyses of semantic priming. *Journal of Memory and Language, 59,* 495-523.

Batterink, L., & Neville, H. (2011). Implicit and explicit mechanisms of word learning in a narrative context: An event-related potential study. *Journal of Cognitive Neuroscience, 23,* 3181-3196.

Borovsky, A., Kutas, M., & Elman, J. (2010). Learning to use words: Event-related potentials index single-shot contextual word learning. *Cognition, 116,* 289-296.

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,* B45-B54.

Breitenstein, C., Zwitserlood, P., de Vries, M. H., Feldhues, C., Knecht, S., & Dobel, C. (2007). Five days versus a lifetime: Intense associative vocabulary training generates lexically integrated words. *Restorative Neurology and Neuroscience, 25,* 493-500.

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.

Cousineau, D. (2007). Confidence intervals in within-subjects designs: A simpler solution to Loftus and Masson's method. *Tutorials in Quantitative Methods for Psychology, 1,* 42-45.

Dagenbach, D., Horst, S., & Carr, T. H. (1990). Adding new information to semantic memory: How much learning is enough to produce automatic priming? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 581-591.

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 systems account of word learning: Neural and behavioural evidence. *Philosophical Transactions of the Royal Society, 364,* 3773-3800.

Deacon, D., Dynowska, A., Ritter, W., & Grose-Fifer, J. (2004). Repetition and semantic priming of nonwords: Implications for theories of N400 and word recognition. *Psychophysiology, 41,* 60-74.

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

Durgunoglu, A. Y., & Neely, J. H. (1987). On obtaining episodic priming in a lexical decision task following paired-associate learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13,* 206-222.

Elgort, I. (2011). Deliberate learning and vocabulary acquisition in a second language. *Language Learning, 61,* 367-413.

Evans, G. A. L., Lambon Ralph, M. A., & Woollams, A. M. (2012). What's in a word? A parametric study of semantic influences on visual word recognition. *Psychonomic Bulletin & Review, 19,* 325-331.

Fernandes, T., Kolinsky, R., & Ventura, P. (2009). The metamorphosis of the statistical segmentation output: Lexicalization during artificial language learning. *Cognition, 112,* 349-366.

Forster, K. I. (1985). Lexical acquisition and the modular lexicon. *Language and Cognitive Processes, 1,* 87-108.

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

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

Hayes, S. C., & Bissett, R. T. (1998). Derived stimulus relations produce mediated and episodic priming. *The Psychological Record, 48,* 617-630.

Hutchison, K. A. (2003). Is semantic priming due to association strength or feature overlap? A microanalytic review. *Psychonomic Bulletin & Review, 10,* 785-813.

Hutchison, K. A., Balota, D. A., Cortese, M. J., & Watson, J. M. (2008). Predicting semantic priming at the item level. *The Quarterly Journal of Experimental Psychology, 61,* 1036-1066.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59,* 434-446.

Janssen, D. P. (2012). Twice random, once mixed: Applying mixed models to simultaneously analyse random effects of language and participants. *Behavior Research Methods, 44,* 232-247.

Jiang, N., & Forster, K. I. (2001). Cross-language priming asymmetries in lexical decision and episodic recognition. *Journal of Memory and Language, 44,* 32-51.

Jones, L. L. (2010). Pure mediated priming: A retrospective semantic matching model. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36,* 135-146.

Joordens, S., & Becker, S. (1997). The long and short of semantic priming effects in lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 1083-1105.

Kahan, T. A., Neely, J. H., & Forsythe, W. J. (1999). Dissociated backward priming effects in lexical decision and pronunciation tasks. *Psychonomic Bulletin & Review, 6,* 105-110.

Kouider, S., & Dupoux, E. (2004). Partial awareness creates the “illusion” of subliminal semantic priming. *Psychological Science, 15,* 75-81.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology, 62,* 621-647.

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. (in press). Lexical integration of novel words without sleep. *Journal of Experimental Psychology: Learning, Memory & Cognition*.

Lindsay, S., & Gaskell, M. G. (2010). A complementary systems account of word learning in L1 and L2. *Language Learning, 60, S2,* 45-63.

Lucas, M. (2000). Semantic priming without association: A meta-analytic review. *Psychonomic Bulletin & Review, 7,* 618-630.

Marcel, A. J. (1983). Conscious and unconscious perception: Experiments on visual masking and word recognition. *Cognitive Psychology, 15,* 197-237.

Marslen-Wilson, W. (1987). Functional parallelism in spoken word recognition. *Cognition, 25,* 71-102.

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning-systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review, 102,* 419–457.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review, 88,* 375–407.

McCormick, S. F., & Rastle, K. (2012). Lexicalisation of trained new words: Evidence from masked form priming. Manuscript in preparation.

McKoon, G., & Ratcliff, R. (1979). Priming episodic and semantic memory. *Journal of Verbal Learning and Verbal Behavior, 18,* 463-480.

McNamara, T. P. (2005). *Semantic priming: Perspectives from memory and word recognition.* New York: Psychology Press.

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

Mestres-Misse, A., Rodriguez-Fornells, A., & Munte, T. F. (2007). Watching the brain during meaning acquisition. *Cerebral Cortex, 17,* 1858-1866.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (2004). The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods, Instruments, & Computers, 36,* 402–407.

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

Pecher, D. & Raaijmakers, J. G. W. (1999). Automatic priming effects for new associations in lexical decision and perceptual identification. *Quarterly Journal of Experimental Psychology, 52A,* 593-614.

Pecher, D., Zeelenberg, R., & Raaijmakers, J. G. W. (2002). Associative priming in a masked perceptual identification task: Evidence for automatic processes. *Quarterly Journal of Experimental Psychology, 55A,* 1157-1173.

Perfetti, C. A., Wlotko, E. W., & Hart, L. A. (2005). Word learning and individual differences in word learning reflected in event-related potentials. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 1281-1292.

Pollatsek, A., & Well, A. D. (1995). On the use of counterbalanced designs in cognitive research: A suggestion for a better and more powerful analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 785–794.

Qiao, X, Forster, K. I., & Witzel, N. (2009). Is banara really a word? *Cognition, 113,* 254-257.

Rajaram, S., & Neely, J. H. (1992). Dissociative masked repetition priming and word frequency effects in lexical decision and episodic recognition tasks. *Journal of Memory and Language, 31,* 152-182.

Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning – Taking memory tests improves long-term retention. *Psychological Science, 17,* 249-255.

Salasoo, A., Shiffrin, R. M., & Feustel, T. C. (1985). Building permanent memory codes: Codification and repetition effects in word identification. *Journal of Experimental Psychology: General, 114,* 50-77.

Shtyrov, Y., Nikulin, V., & Pulvermuller, F. (2010). Rapid cortical plasticity underlying novel word learning. *Journal of Neuroscience, 30,* 16864-16867.

Tamminen, J., Davis, M. H., Merkx, M., & Rastle, K. (2012). The role of memory consolidation in generalisation of new linguistic information. *Cognition, 125,* 107-112.

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

Tamminen, J., Payne, J. D., Stickgold, R., Walmsley, E. J., & Gaskell, M. G. (2010). Sleep spindle activity is associated with the integration of new memories and existing knowledge. *Journal of Neuroscience, 30,* 14356–14360.

Thomas, M. A., Neely, J. H., & O'Connor, P. (2012). When word identification gets tough, retrospective semantic processing comes to the rescue. *Journal of Memory and Language, 66,* 623-643.

Van Casteren, M., & Davis, M. H. (2006). Mix, a program for pseudorandomization. *Behavior Research Methods, 38,* 584-589.

Van den Bussche, E., Van den Noortgate, W., & Reynvoet, B. (2009). Mechanisms of masked priming: A meta-analysis. *Psychological Bulletin, 135,* 452-477.

Footnotes

[1]Only 68 of these were trained, the remaining 34 were untrained and used in a shadowing task not reported in this paper. This task always took place after the priming task, and therefore could not have affected the data reported in the present paper. All 102 nonwords were used as trained novel words across all the participants, as the nonwords were rotated through the trained/untrained conditions.

[2]There is currently little consensus among users of mixed-effects models in the field of psycholinguistics on the use of random slopes. Here we adopted the practice of including them only when they significantly improve the fit of the model, as recommended by Baayen et al. (2008) for example. However, in those cases in which random slopes did not significantly improve the fit of the model, we have checked that the effects we report being significant remain so even in the presence of subject-specific random slopes for the effect in question (unless including the random slope prevented the model from converging).

[3]We have also analysed the number of features recalled by participants; this analysis provided identical results to the analysis based on object recall reported here in both experiments.

[4]We thank Jim Neely for suggesting this analysis.

Figure captions

Figure 1. Diagram showing the timing of the training and testing sessions in Experiment 1. Participants learned two different sets of novel words (A and B), the timing of which depended on group allocation (long or short consolidation opportunity).

Figure 2. Percentage of novel word meanings recalled accurately in the test session of Experiment 1. Error bars represent within-participant standard error of the mean (Cousineau, 2007).

Figure 3. Diagram showing the timing of the training and testing sessions in Experiment 2. Participants learned one set of novel words (A), which was subsequently tested at three points in time.

Figure 4. Percentage of novel word meanings recalled accurately in the test sessions of Experiment 2. Error bars in the repeatedly tested group represent within-participants standard error of the mean (Cousineau, 2007), and conventional standard error in the once-tested group.

Figure 1.

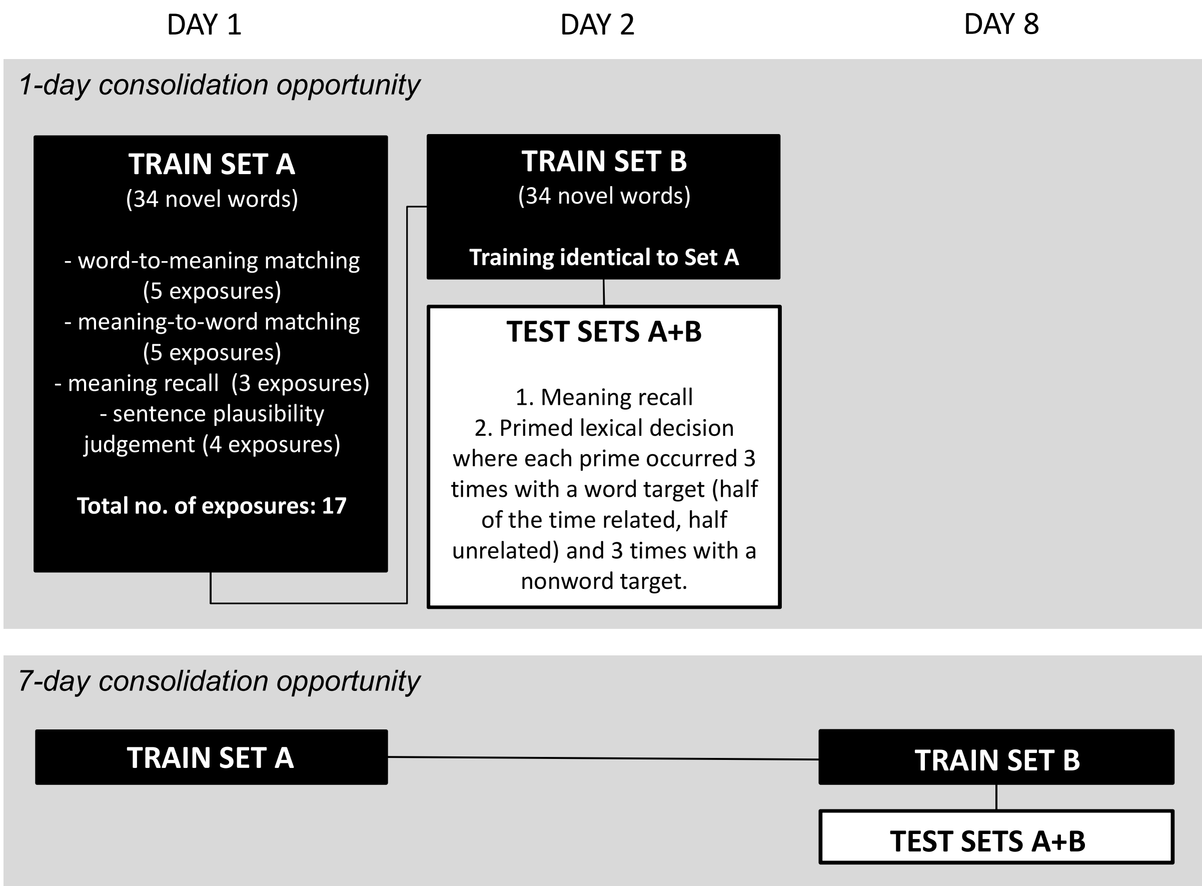
DAY 1
DAY 2
DAY 8
1-day consolidation opportunity
TRAIN SET A
(34 novel words)
- word-to-meaning matching (5 exposures)
- meaning-to-word matching (5 exposures)
- meaning recall (3 exposures)
- sentence plausibility judgement (4 exposures)
Total no. of exposures: 17
TRAIN SET B
(34 novel words)
Training identical to Set A
TEST SETS A+B
1. Meaning recall
2. Primed lexical decision where each prime occurred 3 times with a word target (half of the time related, half unrelated) and 3 times with a nonword target.
7-day consolidation opportunity
TRAIN SET A
TRAIN SET B
TEST SETS A+B

Figure 2.

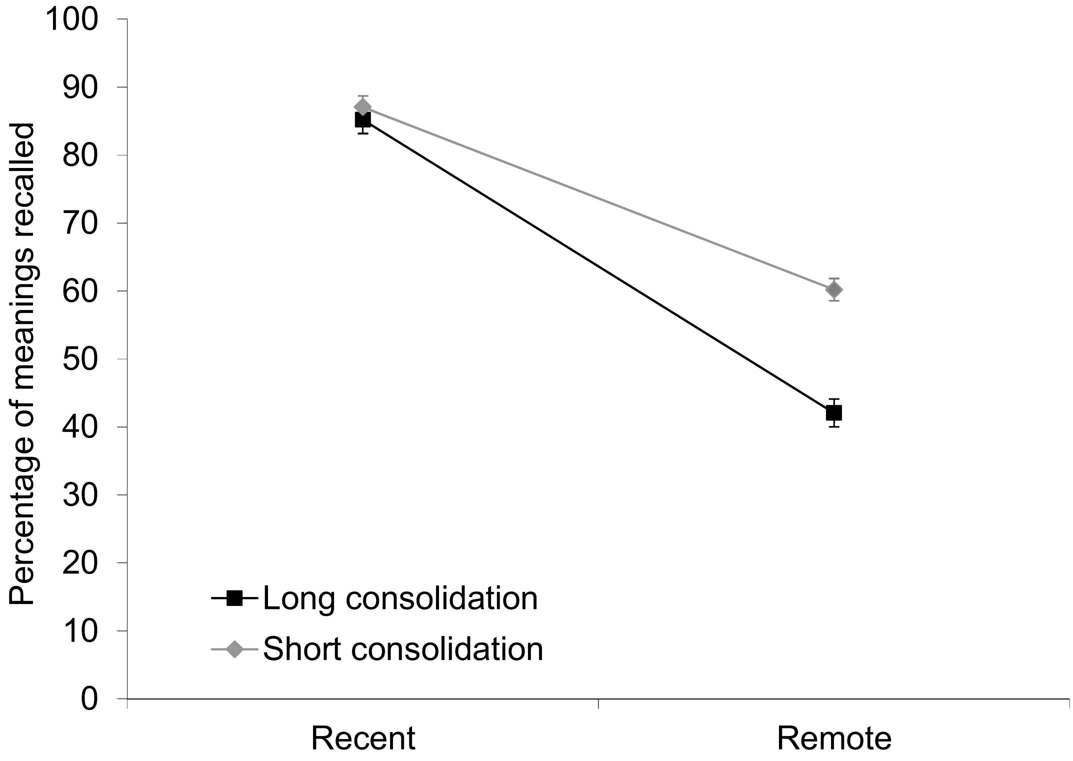
100
90
80
70
60
50
40
30
20
10
0
Percentage of meanings recalled
Long consolidation
Short consolidation
Recent
Remote

Figure 3.

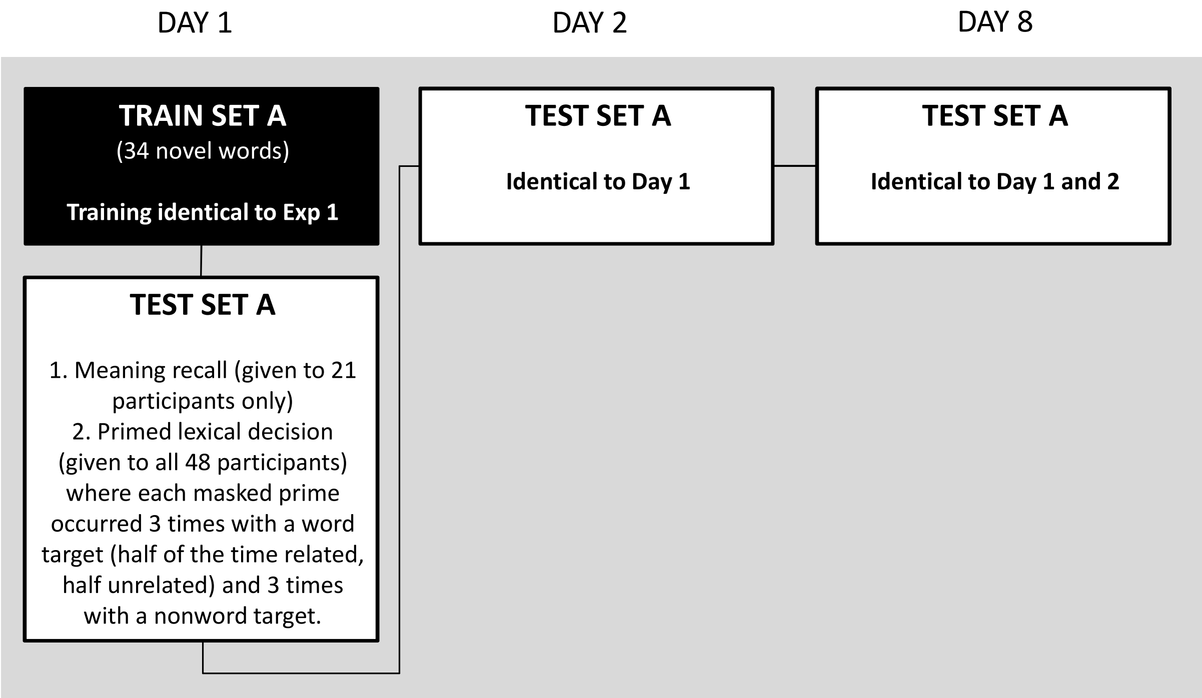
DAY 1
DAY 2
DAY 8
TRAIN SET A
(34 novel words)
Training identical to Exp 1
TEST SET A
1. Meaning recall (given to 21 participants only)
2. Primed lexical decision (given to all 48 participants) where each masked prime occurred 3 times with a word target (half of the time related, half unrelated) and 3 times with a nonword target.
TEST SET A
Identical to Day 1
TEST SET A
Identical to Day 1 and 2

Figure 4.

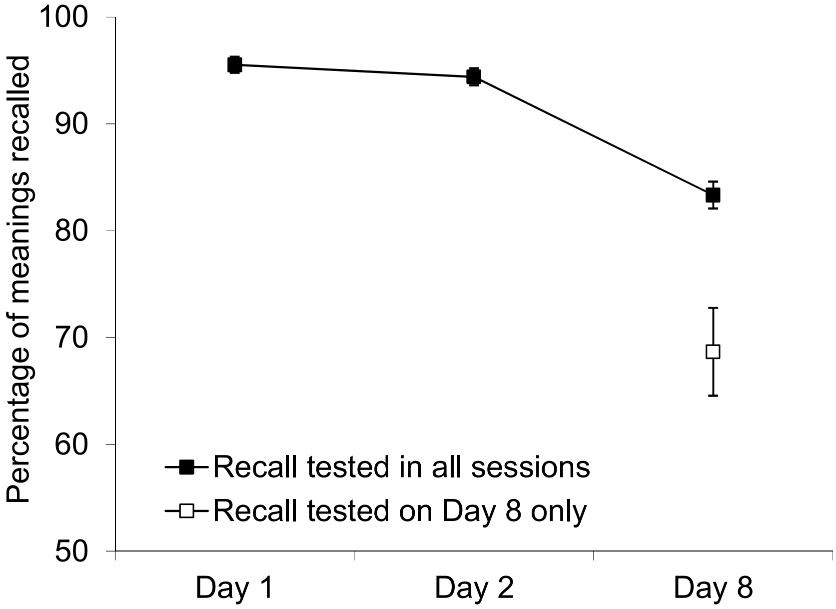
Percentage of meanings recalled
100
90
80
70
60
50
Recall tested in all sessions
Recall tested on Day 8 only
Day 1
Day 2
Day 8

Table 1. Mean RTs (in ms) to each novel word prime and real word prime condition in Experiment 1 ± within-participant standard error of the mean (Cousineau, 2007). Percentage of errors in parentheses, upper and lower limits of 95% confidence intervals for the priming effect given in brackets.

Consolidation length		Primed	Unprimed	Priming effect (in ms)
1-day	Remote primes	470±2.7 (9.2±0.7)	478±2.0 (8.3±0.8)	8 [1, 16]
	Recent primes	477±2.7 (9.3±0.8)	482±2.5 (9.3±0.6)	5 [-3, 14]
	Real word primes	468±2.4 (7.2±0.5)	485±2.4 (11.7±0.5)	17 [7, 27]
7-day	Remote primes	468±2.4 (8.5±0.7)	473±1.7 (10.2±0.7)	5 [0, 13]
	Recent primes	476±2.7 (9.9±0.8)	475±2.3 (9.9±0.7)	-1 [-10, 7]
	Real word primes	463±2.7 (9.0±0.6)	482±2.7 (11.6±0.6)	19 [8, 29]
Combined	Remote primes	469±1.8 (8.8±0.5)	476±1.3 (9.2±0.5)	7 [3, 12]
	Recent primes	476±1.9 (9.6±0.5)	478±1.7 (9.6±0.4)	2 [-4, 8]
	Real word primes	466±1.8 (8.1±0.4)	484±1.8 (11.7±0.4)	18 [10, 25]

Table 2a. Mean lexical decision RTs (in ms) using real word primes in Experiment 2 ± within-participant standard error of the mean (Cousineau, 2007). Percentage of errors in parentheses, upper and lower limits of 95% confidence intervals for the priming effect given in brackets.

Recall of novel word meanings	Time of testing	Primed	Unprimed	Priming effect (in ms)
Tested	Day 1	528±6.0 (8.6±0.8)	542±5.6 (8.8±1.0)	14 [3, 26]
	Day 2	532±5.6 (6.8±0.7)	533±6.4 (8.4±0.6)	1 [-9, 10]
	Day 8	516±5.1 (8.2±0.7)	525±5.8 (8.5±1.3)	9 [2, 17]
Not tested	Day 1	512±4.7 (7.6±0.7)	516±5.4 (7.8±0.9)	4 [-4, 10]
	Day 2	498±3.6 (7.6±0.9)	504±3.9 (8.5±0.7)	6 [-3, 15]
	Day 8	510±4.4 (7.0±0.7)	518±4.4 (8.0±0.7)	8 [1, 15]
Combined	Day 1	519±3.7 (8.0±0.6)	527±3.9 (8.2±0.7)	8 [2, 15]
	Day 2	513±3.3 (7.2±0.6)	517±3.6 (8.5±0.5)	4 [-3, 10]
	Day 8	512±3.4 (7.6±0.5)	521±3.6 (8.2±0.7)	9 [3, 13]

Table 2b. Mean lexical decision RTs (in ms) using novel word primes in Experiment 2 ± within-participant standard error of the mean (Cousineau, 2007). Percentage of errors in parentheses, upper and lower limits of 95% confidence intervals for the priming effect given in brackets.

Recall of novel word meanings	Time of testing	Primed	Unprimed	Priming effect (in ms)
Tested	Day 1	533±6.7 (10.7±1.5)	541±5.9 (9.9±1.0)	8 [-5, 21]
	Day 2	539±5.5 (7.1±0.8)	536±6.6 (7.1±0.9)	-3 [-9, 4]
	Day 8	530±4.8 (7.9±1.1)	537±4.6 (9.6±1.4)	7 [-2, 18]
Not tested	Day 1	520±4.2 (7.9±0.8)	523±3.9 (7.8±1.0)	3 [-7, 12]
	Day 2	505±3.2 (6.9±0.6)	506±3.2 (9.0±0.9)	1 [-6, 12]
	Day 8	510±3.4 (7.3±0.6)	519±3.8 (8.8±0.8)	9 [2, 16]
Combined	Day 1	525±3.8 (9.2±0.8)	530±3.4 (8.7±0.7)	5 [-3, 13]
	Day 2	519±3.1 (7.0±0.5)	520±3.4 (8.2±0.7)	1 [-5, 6]
	Day 8	519±2.8 (7.6±0.6)	527±2.9 (9.2±0.7)	8 [3, 14]