Newly-acquired words are more phonologically robust in verbal short-term memory when they have associated semantic representations

Nicola Savill, Andrew W. Ellis and Elizabeth Jefferies

Department of Psychology, University of York, UK

To appear in ***Neuropsychologia***

doi: 10.1016/j.neuropsychologia.2016.03.006

Corresponding Author:

Nicola Savill

Department of Psychology,

University of York,

Heslington,

York

YO10 5DD

UK

Email: nicola.savill@york.ac.uk

Tel: +44 (0)1904 322937

Fax: +44(0)1904 323181

Keywords: novel word learning; phonological representations; semantic binding; short-term memory; lexicalisation; meaning

Abstract

Verbal short-term memory (STM) is a crucial cognitive function central to language learning, comprehension and reasoning, yet the processes that underlie this capacity are not fully understood. In particular, although STM primarily draws on a phonological code, interactions between long-term phonological and semantic representations might help to stabilise the phonological trace for words (“semantic binding hypothesis”). This idea was first proposed to explain the frequent phoneme recombination errors made by patients with semantic dementia when recalling words that are no longer fully understood. However, converging evidence in support of semantic binding is scant: it is unusual for studies of healthy participants to examine serial recall at the phoneme level and also it is difficult to separate the contribution of phonological-lexical knowledge from effects of word meaning. We used a new method to disentangle these influences in healthy individuals by training new ‘words’ with or without associated semantic information. We examined phonological coherence in immediate serial recall (ISR), both immediately and the day after training. Trained items were more likely to be recalled than novel nonwords, confirming the importance of phonological-lexical knowledge, and items with semantic associations were also produced more accurately than those with no meaning, at both time points. For semantically-trained items, there were fewer phoneme ordering and identity errors, and consequently more complete target items were produced in both correct and incorrect list positions. These data show that lexical-semantic knowledge improves the robustness of verbal STM at the sub-item level, even when the effect of phonological familiarity is taken into account.

1. Introduction

Communication, thought and vocabulary acquisition draw on verbal short-term memory (STM) – i.e., the ability to actively maintain verbal information for brief periods. Theoretical accounts of this function have suggested it largely reflects temporary activation of a phonological code (Baddeley & Hitch, 1974; Baddeley, 1986, 2000). However, STM plays a crucial role in extracting and conveying semantic information through language. Semantic representations might, therefore, influence the stability of the phonological trace and this effect might be crucial for understanding word learning and comprehension at a sub-item level.

We know that speech sounds are maintained better in STM when they are meaningful. When participants reproduce a sequence of items in order, as in immediate serial recall (ISR), performance is better for lists of words that are higher in imageability/concreteness (Acheson, Postle, & MacDonald, 2010; Allen & Hulme, 2006; Caza & Belleville, 1999; Hoffman, Jefferies, Ehsan, Jones, & Lambon Ralph, 2009; Jefferies, Frankish, & Lambon Ralph, 2006a; Majerus & van der Linden, 2003; Roche, Tolan, & Tehan, 2011; Romani, McAlpine, & Martin, 2008; Walker & Hulme, 1999; Wilshire, Keall, & O'Donnell, 2010), related in meaning (Poirier & Saint-Aubin, 1995; Saint-Aubin et al., 2013; Wilshire et al., 2010), or when word meaning has been emphasised at encoding (Campoy & Baddeley, 2008; Savill, Metcalfe, Ellis, & Jefferies, 2015).

Furthermore, patients with semantic dementia – who show progressive loss of semantic knowledge associated with atrophy of the anterior temporal lobe – have difficulty maintaining the correct phonological forms of words that are poorly understood (Hoffman et al., 2009; Jefferies, Crisp, & Lambon Ralph, 2006; Jefferies, Jones, Bateman, & Lambon Ralph, 2004; Knott, Patterson, & Hodges, 1997; Majerus, Norris, & Patterson, 2007; Patterson, Graham, & Hodges, 1994). As first noticed by Patterson et al. (1994), these patients frequently recombine the phonological elements of different items in ISR, particularly when repeating words with more degraded meanings, despite fluent speech production and generally intact phonological performance. ISR impairments related to compromised semantic function have since been observed in patients with Alzheimer's disease (Peters, Majerus, De Baerdemaeker, Salmon, & Collette, 2009) and those with focal frontal and temporal lesions (Forde & Humphreys, 2002; Jefferies, Hoffman, Jones, & Lambon Ralph, 2008; Wilshire et al., 2010).

Nevertheless, there is disagreement regarding the way in which semantic information influences verbal STM; whether beneficial effects of word meaning occur at a lexical level or sub-lexical phoneme level. Some accounts have proposed that semantic information might aid STM via a process of 'redintegration', whereby the phonological trace that has degraded over time is actively reconstructed from long-term lexical knowledge – for example, semantic information would enable participants to establish if they had been presented with the word 'man' or 'map', allowing the complete item to be produced accurately, but not necessarily in the correct location within the list (Poirier & Saint-Aubin, 1995; Walker & Hulme, 1999). Alternatively, 'linguistic' accounts have proposed that semantic information influences the stability of the phonological trace more directly, by virtue of bidirectional connections between these systems (Acheson & MacDonald, 2009; Jefferies, Frankish, & Noble, 2009; Patterson et al., 1994). That is, in both speech production and auditory comprehension tasks, activation in the phonological system is thought to represent a sequence of phonemes in order, and this phonological processing co-occurs with semantic activation. As a consequence, semantic activation might benefit serial order memory in the phonological system. Most models of language processing allow interactive-activation between semantics and phonology (e.g., Dell, Schwartz, Martin, Saffran, & Gagnon, 1997; Patterson et al., 1994; Plaut & Kello, 1999; Plaut, McClelland, Seidenberg, & Patterson, 1996) – although these models rarely include an explicit phoneme sequencing mechanism. Nevertheless, the finding of better ISR for more meaningful verbal information can potentially emerge naturally from accounts of STM that explain this capacity in terms of temporary activation or weight changes within the language system (see Majerus, 2013, for an overview).

If we assume that the phonological system has an in-built capacity to maintain a string of speech sounds in order, we might envisage that semantic information can constrain the order of phonemes in verbal STM. The "semantic binding hypothesis" proposed by Patterson et al. (1994) suggests that the phonemes of words can be maintained in the correct configuration more easily than phonemes corresponding to nonwords for two reasons: (i) the phonological system learns familiar sequences and thus develops "pattern completion properties" for words; (ii) the phonological system receives further stabilising input from co-activation with semantic representations, strengthening this pattern completion effect for words[1]. This theoretical framework predicts that semantic information contributes to order memory but specifically at the level of individual phonemes – i.e., the constituents of words should be *less likely* to split apart, to migrate to a different place in the sequence and be recombined with the elements of other list items. Importantly, this account differs from the current redintegration perspective in its assumption that the position of phonemes in STM is inherently unstable and vulnerable to migration or loss, and it is the availability of long-term representations that help to bind phonemes together and reduce such movement. Recall-based accounts (such as redintegration) assume the quality/stability of the available phonological trace in STM itself is not directly influenced by long-term activation –and, as such, do not include a mechanism to predict patterns in sub-item phoneme movement. At the whole-item level, in contrast, both the

[1] Patterson et al.'s (1994) hypothesis discusses the influence of long-term representations on STM in terms of the strength of ongoing temporary activation, but more recent work suggests that STM can occur via synaptic priming (Lewis-Peacock, Drysdale, Oberauer, & Postle, 2012), which would yield similar predictions.

redintegration and semantic binding hypotheses predict a potential *increase* in order errors with greater semantic support, since strong binding (or redintegration of an item) should encourage all of the elements of a word to be recalled together, even when the location of the item in the list is incorrect.

The predictions arising from the semantic binding hypothesis have important implications for our understanding of word learning and comprehension but these predictions have not been adequately tested because (i) most studies of immediate serial recall have examined performance at the level of whole items and have not looked for a semantic influence on order memory at the phoneme level and (ii) it is difficult to examine the influence of semantic information on phonological maintenance independently of phonological-lexical familiarity. Studies that have examined phoneme-level recall have largely failed to separate these factors. The semantic binding hypothesis was originally proposed to explain the tendency of patients with semantic dementia to recombine phonemes across different items in immediate serial recall for semantically-degraded words (Hoffman et al., 2009; Jefferies et al., 2004; Knott et al., 1997; Majerus et al., 2007; Patterson et al., 1994). This phoneme error pattern is also seen for nonwords compared to words in healthy individuals (Hoffman et al., 2009; Jefferies, Frankish, et al., 2006a; Jefferies, Frankish, & Lambon Ralph, 2006b; Jefferies et al., 2009). However, semantic binding might not be necessary to explain this ISR advantage for 'known' words relative to semantically degraded items or nonwords, since familiarity with the phonological form of words alone, without an accompanying semantic representation, also benefits phonological stability (Benetello, Cecchetto, & Papagno, 2015; Savill, Ashton, et al., 2015). We recently showed that familiarisation with the phonological form of nonwords *without* meaning increased the coherence of these items in ISR (Savill, Ashton, et al., 2015). Thus, it has been proposed that differences in phoneme migrations for 'known' and ‘semantically degraded’ words in semantic dementia might relate to greater phonological familiarity and frequency of usage for words that are still understood (Papagno, Vernice, & Cecchetto, 2013).

One way to determine if semantic information benefits short-term phonological maintenance independently of phonological-lexical familiarity is to test ISR for trained sets of items matched for exposure to the phonological-lexical form, but varying in the availability of semantic support. This broadly constituted the approach in a recent study that found no difference in ISR between foreign words familiarised with and without semantics (Benetello et al., 2015), leading the authors to conclude that, in the context of supporting recently familiarised phonological-lexical forms in STM, ‘meaning is useless’. However, the study’s design might have restricted sensitivity to semantic effects, since there was a small set of items for each condition (N=10) and each condition was tested separately; thus the ISR test was characterised by massed repetition. Lexical and semantic effects on ISR are strongest in open sets of non-repeated items (Jefferies et al., 2004; Roodenrys & Quinlan, 2000): small, closed sets maximise demands on whole-item order memory, while the use of more open sets taps the retention of phoneme order and item identity, which are most likely to benefit from semantic support. In addition, Benetello et al. tested ISR immediately after training yet previous research suggests that lexical and semantic learning effects may emerge or strengthen after a period of consolidation. For example, Davis and Gaskell’s complementary systems account of word learning characterises a process of ‘lexicalisation’ over time, particularly after a period of sleep (Davis & Gaskell, 2009; see also McClelland, McNaughton, & O’Reilly, 1995). Evidence for the acquisition and integration of semantic representations for new words also emerges after sleep (e.g., Clay et al., 2007; Lin & Yang, 2014; Tamminen & Gaskell, 2013).

1.1. *The present study*

To examine whether semantic information increases the stability of the phonological trace, independently of phonological familiarity, we considered 1) if nonwords trained with a semantic association would have higher stability than nonwords trained without meaning – i.e., if more whole items would be correctly recalled and if there would be fewer phoneme ordering errors in ISR following semantic training – and 2) if lexical or semantic effects on phonological coherence would be apparent in ISR from the day of training, or if they would emerge more strongly on the subsequent day. We used a phonological familiarisation procedure developed by Savill, Ashton, et al. (2015), in which 48 novel phonological forms were paired with either pictures of objects (in the semantically-trained condition; SEM) or blurred images in which there were no discernible features (in the no semantics condition; FAM). The paired images were of previously-unfamiliar objects that did not have a pre-existing name. Prior to phonological familiarisation, participants were taught 'facts' about the objects to make them more meaningful. ISR was then tested on the day of training and the subsequent day alongside sets of untrained nonwords (NEW).

These manipulations could yield several possible outcomes, with differing interpretations: For example, an effect of phonological familiarisation with no additional effect of semantic training (i.e., ISR accuracy: NEW < FAM = SEM) would imply pattern completion within the phonological representations, irrespective of semantics; an additional effect of semantic training (i.e., ISR accuracy: NEW < FAM < SEM or NEW = FAM < SEM) would imply an added contribution of word meaning to the phonological stability of the nonwords. If either lexical or semantic training effects require a period of consolidation (such that effects emerge on Day 2) we would also expect day by training interactions.

2. Method

2.1 Study overview

The study employed a within-subjects design. Participants took part in two sessions, at the same time of day, on consecutive days (Figure 1). On Day 1, meaning was provided for a set of images of novel objects in a semantic training task. In this task, participants learnt the functions, contexts and operations ascribed to the pictured objects. We used unfamiliar objects without known names so that pre-existing lexical and semantic knowledge would not interfere with the effects of training. Semantic training trials (involving the presentation of clear photographs of novel objects paired with semantic facts) were randomly interspersed with meaningless blurred images of objects, without distinct features, which were also used in the following phonological familiarisation phase. These blurred images were presented without any facts and required only a button press in acknowledgment, with exposure frequency and duration matched to the clear object images.

Immediately after the semantic training phase, also on Day 1, participants were familiarised with the phonological forms of nonwords in two sets: in one set, nonwords were paired with the semantically-trained objects (SEM condition) and in the other set, nonwords were presented alongside the blurred images with no associated meaning (FAM condition).

Learning was subsequently assessed by asking participants to freely recall as many of the nonwords as they could; through an immediate serial recall (ISR) task, in which lists to be repeated were composed of SEM, FAM or untrained NEW items and on a subsequent picture naming task. These three tests were repeated on Day 2, after an opportunity for consolidation (with the items from the first ISR task reordered to form new test lists for ISR). On Day 2 only, these tests were followed by a paced reading task (similar to Wilshire, 1998), involving written forms of the nonwords from all three conditions, which is not discussed further. Finally, a phonological discrimination task consisting of previously presented items and phonological neighbours as distractors was used to assess the recognition of nonwords in all three exposure conditions on Day 2.

FIGURE 1 ABOUT HERE

2.2. Participants

Thirty-six native British English undergraduate students who had not been participants in the earlier study by Savill, Ashton, et al. (2015), took part in the testing sessions held over two days. Participants all had normal vision and hearing, and were paid for participation.

2.3. Stimuli

The nonwords and images used for training were selected from Savill, Ashton, et al. (2015). 72 spoken disyllabic (CVCVC) nonwords (C = consonant, V = vowel, e.g., 'vaitag' /vetæg/) were grouped into three sets of 24 items, which were allocated to one of three training conditions: familiarisation with semantic features (SEM), familiarisation without semantic features (FAM), and no training (NEW). The sets of 24 items were assigned to groups in a way that allowed the creation of four-item ISR lists that did not contain repetition of phonemes in the same syllable position. This allowed us to track the majority of phoneme migrations at recall, since phonemes tend to preserve syllable position when they are recalled as part of the wrong item (Ellis, 1980). The corresponding six unique sets of four nonwords were reordered twice to produce 18 ISR test lists per condition, each containing four nonword items. The phonotactic properties of the nonwords, which had been designed to obey English phonology but such that the whole phonological forms avoided close similarity with existing English words (i.e., had no phonological neighbours), were matched between nonword sets

(Table 1). The presentation order of the lists and order of items within lists were reordered for testing on the second day. Allocation of nonword stimuli to conditions was rotated across participants.

TABLE 1 ABOUT HERE

Images were of novel objects with no clear name and blurred versions of object images, taken from the 'referent' and 'no-referent' conditions of Savill, Ashton, et al. (2015). Twenty four 'referent' objects were selected for semantic training: these were given plausible meanings consistent with the visual features visible in the photographs. Three facts were trained per object, pertaining to function (i.e., what it is used for), operation (i.e., how the object works), and the context in which it would be found. These facts formed a unique semantic representation for each image (i.e., akin to the way that a thin object with bristles at one end that is moved in a repetitive left-right motion, is found in bathrooms, and is used for keeping teeth clean, would converge on the concept of a toothbrush; see Figure 2). These semantically-trained images were assigned to the SEM set of nonwords in the phonological familiarisation phase, while the blurred 'no referent' images were assigned to the FAM set of nonwords.

In addition to the 72 spoken nonword stimuli used in the phonological familiarisation and ISR tasks, 72 phonological neighbours of these stimuli (that were recorded and edited in the same conditions as the original stimuli recordings) were collated for use in a phonological discrimination task, tested at the end of the experiment. These neighbour stimuli differed by one phoneme in one of any of the five possible phoneme positions (e.g., 'vaitag' /vetæg/ had the neighbour 'vaitang' /vetæŋ/, but each neighbour could have been one of a multitude of alternatives, such as 'baitag' /betæg/or 'vottag' /vɒtæg/).

2.4. Procedures

2.4.1. Semantic training task procedure

In this first task, participants were told that they would see photographs of some unusual objects with two descriptions displayed beneath them: they were asked to decide which of the two descriptions applied to the object, and pressed a key accordingly (SEM-training trials). They were advised that blurred images would also randomly appear with no descriptions (FAM-image trials; see Figure 2) and that these required only a button press in acknowledgment. The incorrect descriptions displayed in the SEM-training trials were taken from other test objects. These trials remained on the screen until participants responded, after which feedback on the accuracy of their response was displayed for 1s. The function, context, and operation facts were presented three times. Thus, each object underwent nine semantic training trials. The FAM-images were presented for a minimum fixed duration of 2 s and were then accompanied by a visual prompt to continue until participants pressed a key. Each of these images was also presented nine times. Images were presented via E-prime in a pseudo-random order such that each item appeared equally in both halves of the experiment. Participants had a rest break half way through the trials.

FIGURE 2 ABOUT HERE

2.4.2. Phonological familiarisation task

The phonological familiarisation task was adapted and extended from Savill, Ashton, et al. (2015; see Figure 4): Each image from the semantic training task was displayed for 500 ms prior to the onset of an auditory nonword. The image stayed on screen until participants pressed one of two keys to indicate whether the pairing was correct or incorrect. This elicited immediate visual feedback regarding the accuracy of the response, displayed for 1 s, followed by the next trial. Each nonword was presented six times with its correct image, and four times with a different image. Trials were presented in a pseudo-random order, such that half of the correct and incorrect trials for a given nonword were presented in each half of the task. Participants were trained with one of three versions of the phonological familiarisation task, in which different sets of nonwords were paired with the SEM, FAM and NEW conditions.

2.4.3. Free recall

Immediately after phonological familiarisation on the training day, and also at the beginning of the Day 2 session, participants were given two minutes in which to verbally recall all the items they remembered from training. This task assessed the ability of the participants to independently generate the phonological form of the nonwords and whether this was facilitated by the availability of semantic information. Experimenters noted all responses, and scored them against an item checklist. Responses were subsequently phonemically transcribed.

2.4.4. Immediate Serial Recall task

In this task, participants attempted to repeat lists of spoken items in order. This was our primary assessment of the effects of training on phonological stability. Both ISR assessments, on the day of training and the following day, consisted of 54 recall lists in a fixed pseudo-random order. These lists were composed of items from the SEM, FAM and untrained NEW conditions, and were presented in three mixed blocks separated by rest breaks (six unique lists per condition, with items in each list reordered to produce 18 lists per condition in both versions of the task). The ISR procedure was identical to Savill, Ashton et al. (2015; Figure 3). Participants wore a headset to listen to spoken nonwords presented in lists of four items, at a rate of 1.25 s per nonword. An exclamation mark appeared on screen 250 ms prior to the onset of the first spoken nonword and remained until the offset of the last item, when it was replaced by a question mark to cue the recall attempt. A digital recorder recorded vocal responses. Participants pressed a key after each recall attempt to cue the next trial. The allocation of ISR task version to testing session was counterbalanced across participants.

FIGURE 3 ABOUT HERE

2.4.5. Picture Naming

After ISR, the SEM images were presented on screen one at a time and the participant was asked to attempt to generate the object's name, or to say 'pass' if they had no idea what the name was.

2.4.6. Phonological discrimination task

To verify whether the phonological-lexical forms had been acquired, a final task tested the participants' speeded recognition of the experimental stimuli, which had to be distinguished from their phonological neighbours. Auditory stimuli (NEW, FAM and SEM nonwords and their respective novel phonological neighbours) were randomly presented 200 ms after an exclamation mark prompt appeared on the screen. Participants were asked to indicate whether the presented nonword had been previously presented or not, as quickly and accurately as possible. Our rationale for this task was that well-learned items should be recognised more easily (and quickly) than poorly-learned ones, while the phonological neighbours of well-learned items should be harder (and slower) to reject.

2.5. ISR coding

The transcription procedure and coding scheme were identical to those detailed in Savill, Ashton, et al. (2015). Verbal responses were transcribed phoneme-by-phoneme by two independent coders blind to the experimental conditions. Three data sets (8.3% of the transcripts) were transcribed by both coders to test for inter-rater reliability and these resulted in a high level of agreement at the phoneme level (91% overall)[2].

The coding scheme applied to the transcription automatically categorised each target phoneme as being (1) correct in position within an item that was entirely correct (Whole CIP), (2) correct in position in the context of a partially correct nonword (Individual phoneme CIP), (3) part of a whole item order error (Whole ORD) – i.e., phonemes corresponding to an entire nonword produced out of sequence, or (4) recalled out of position but not as part of a whole correct nonword (Individual phoneme ORD). This latter category consisted of phoneme migration errors – i.e., target phonemes in the list that were produced in the wrong position but in the same *syllabic position* as the target, and that did not migrate as a result of an entire *item* being produced out of sequence[3] – and phoneme repetitions – i.e., incorrectly-positioned target phonemes that were produced more than once, in the same relative syllabic position, and that were not repeated as part of an entire target item. These repetitions were identified through automatic categorisation of phoneme responses that replaced target phonemes that had not been recalled at all. Together these phoneme responses form the target phonemes correctly recalled in or out of position (CAP) measure used to calculate percentages in target-related analyses. After accounting for repetition errors, the remaining 'not recalled' target phonemes were replaced by either (5) phoneme intrusions (INT) – i.e., non-target responses that were not CIP or ORD errors or (6) omissions (OM). Phoneme intrusions

[2] In the three data sets that were transcribed by both coders, a third rater (the lead author) assessed and resolved the conflict in the few cases of disagreement. The majority of the 9% disagreements were due to mumbled or rapid articulation by the participant, which primarily affected vowels (43% of disagreement), and final consonants (30% of disagreement).

[3] Since stimuli were bisyllabic nonwords, migration and repetition errors could include phonemes misplaced across the two syllables within individual items in addition to phonemes misplaced across items, as in Savill, Ashton, et al. (2015).

were quantified by subtracting the number of repetition and OM errors from the total count of target phonemes that were not recalled.

2.6. *Data Analysis*

Unless otherwise stated in the results, the data from each of the tasks were submitted to repeated measures ANOVAs, with post-hoc pairwise comparisons performed where appropriate. For data analysed by ANOVA, partial-eta squared (η_p^2) effect sizes are reported, while Cohen's d (*d*) values are reported for two sample tests.

3. Results

3.1. Semantic feature training task

Improvements in accuracy and faster reaction times (for correct trials) over the course of the semantic training task confirmed that participants successfully learned the associations between the object images and their semantic features, with almost perfect accuracy by the end of training [**Effect of image trial number on accuracy, F (8,280) = 58.60, p < .001**, **η_p^2 = .63; and RT, F (8,280) = 181.23, p < .001**, **η_p^2 = .83**; **see Figure 2**]. Analysis of presentation times confirmed that participants had *at least* the same length exposure to the blurred object images than the semantically-trained clear object images [**trend for longer blurred image exposures, F (1, 35) = 3.16, p = .084, η_p^2 = .08**; **SEM clear images M = 2482.48 ms; SD = 478.84, FAM blurred images M = 2622.52 ms; SD = 272.11**]. Their on-screen time decreased over presentation trials [**F (8,280) = 9.55, p < .001, η_p^2 = .21**].

FIGURE 4 ABOUT HERE

3.2. Phonological familiarisation task

Improvements in accuracy and faster reaction times (for correct trials) over the course of the phonological familiarisation task confirmed that participants also learned to link the images with their associated nonwords by the end of the training [**effects of nonword trial number on accuracy, F (5, 170) = 9.84, p < .001, η_p^2 = .22, and RT, F (5,170) = 12.88, p < .001, η_p^2 = .28, see Figure 4**]. Image referent type (i.e., SEM clear or FAM blurred referent) also influenced performance on this task [**accuracy: F (1, 34) = 40.63, p < .001, η_p^2 = .54; RT: F (1,34) = 20.01, p < .001, η_p^2 = .37**]: the higher distinctiveness of the clear images presumably made it easier to match these stimuli to their phonological forms. However, learning occurred at a similar rate for the two conditions (i.e., there was no interaction with presentation number) [Accuracy: **F (5, 170) = 0.34, p = .88, η_p^2 = .10**].

3.3 *Free Recall*

Immediately post-training and again at the beginning of the second session, participants successfully generated very few trained items, if any at all (four participants failed to produce any entirely correct targets; a further three participants produced only a single FAM item correctly, while another three participants produced only a single SEM item correctly). However, when nonwords were produced, these were more likely to be from the semantically-trained condition [**main effect of training condition, F (1, 35) = 12.99, p < .001, η_p^2 = .27; Table 2**]. Recall was also higher on the second day of testing [**F (1, 35) = 4.58, p < .05, η_p^2 = .12; Table 2**] but there was no interaction between type of training and day [**F (1, 35) = 0.58, p = .45, η_p^2 = .02**][4]. Analyses of near-misses (nonwords recalled with one incorrect phoneme) showed that

[4] Since free recall responses were rare and did not approximate a normal distribution, supplementary non-parametric (Wilcoxon signed rank) analyses were performed: These confirmed better free recall in the semantic training condition on both days (SEM vs. FAM: Day 1 Z = -2.99, p = .003, d = -0.65; Day 2 Z = -2.61, p = .009, d = -0.46), however the individual improvements from Day 1 to Day 2 were only significant for the FAM items (Day 1

these responses more frequently related to semantically-trained nonwords [***F* (1, 35) = 4.83, *p* < .05, η_p^2 = .12**]. The increase in near-misses with Day was a trend [***F* (1, 35) = 3.94, *p* = .055, η_p^2 = .10**] and did not interact with training condition [**F (1, 35) = 0.58, *p* = .45, η_p^2 = .02**].

TABLE 2 ABOUT HERE

3.4. ISR

We first considered whether recent exposure to the phonological forms of nonwords was sufficient to improve ISR *without* a period for consolidation, since a recent study showed that phonological familiarisation improved the accuracy of ISR when this was tested the subsequent day, but the time-course of this effect was not examined (Savill, Ashton, et al., 2015). We compared ISR performance for phonologically familiarised nonwords (FAM) with NEW nonwords on Day 1 and 2 using a 2 × 2 repeated measures ANOVA (examining the factors 'training' and 'day'), permitting a direct comparison with this previous study. Secondly, we examined whether our semantically trained nonwords (SEM) had an added recall advantage over familiarised-only nonwords (FAM) and whether this was contingent upon an opportunity for training consolidation, using a second 2×2 repeated-measures ANOVA (examining the factors 'type of training' and 'day'). The main effect of day in itself is not informative since it would be remarkable if additional phonological exposure did not improve accuracy. Of more interest is the possibility of an interaction between day and the training manipulations. For simplicity, we only report interactions of day with training that reached p < .1. We also provide paired *t*-tests for each training contrast, separately for each day in the Supplementary Materials.

---

vs. Day 2 SEM: *Z* = -1.09, *p* = .277, *d* = -0.14; FAM *Z* = -2.35, *p* = .019, *d* = -0.44). Wilcoxon analyses showed that near-misses were significantly influenced by training condition only on Day 2 (SEM vs. FAM: Day 1 *Z* = -0.863, *p* = .39, *d* = -0.24; Day 2 *Z* = -2.12, *p* = .034, *d* = -0.23).

Since phoneme order errors are contingent on the numbers of phoneme targets recalled overall in any position (Jefferies, Frankish, et al., 2006a), we expressed each of the target-related phoneme responses used to assess the phonological coherence of the response (Whole CIP, Individual phoneme CIP, Whole ORD and individual Phoneme ORD) as a percentage of the overall total of target phonemes recalled in any position (CAP) (i.e., the individual's sum of phonemes recalled across the four target-related categories; Figure 5). This approach is in line with other studies of verbal short-term memory, which typically express order memory as a proportion of target items recalled in any position (e.g., Baddeley, Hitch, & Allen, 2009; Jefferies, Frankish, & Noble, 2011; Jefferies, Frankish, et al., 2006a; Miller & Roodenrys, 2009; Murdock, 1976; Poirier & Saint-Aubin, 1996; Saint-Aubin et al., 2013; Saint-Aubin & Poirier, 1999; Savill, Ashton, et al., 2015; Tse, Li, & Altarriba, 2011). For completeness, OM and INT errors (non-target responses which cannot be analysed in the same way) were analysed as a percentages of total *target* phonemes. To demonstrate that our key results are not an artefact of our analysis method, the supplementary materials provide an analysis of all responses expressed as a percentage of the number of targets.

3.4.1. Analyses of total target phonemes recalled in any position (CAP) (as a percentage of total phoneme targets)

Prior phonological familiarisation helped overall target phoneme recall [**main effect of phonological familiarisation, F (1, 35) = 28.39, p < .001, η_p^2 = .45; Figure 5**]**.** More target phonemes were recalled in any position for familiarised nonwords, consistent with Savill, Ashton, et al. (2015).

Semantic training had an additional impact on overall target phoneme recall, with more target phonemes recalled in the SEM condition than the FAM condition [**main effect of semantic training, *F (*1, 35) = 6.79, *p* < .05, η_p^2 = .16; Figure 5**].

For both comparisons the recall improvement from Day 1 to Day 2 did not significantly interact with training condition.

FIGURE 5 ABOUT HERE

3.4.2. Phonemes correct-in-position (as a percentage of phonemes recalled in any position)

Having demonstrated that both phonological-lexical and semantic information benefited the total *number* of phonemes retained (in any order), we examined the destination of target phonemes in spoken responses to test our hypotheses about phonological coherence (see Table 3 for a summary of the results).

3.4.2.1. Phonemes correct-in-position forming whole nonwords (Whole CIP)

On average, 28% of the target phonemes were recalled as part of whole correct nonwords, and this accounted for 35% of target phonemes recalled in any position. On both testing days, phonemes from the FAM nonwords were more frequently recalled as a whole correct items than phonemes from NEW nonwords [**main effect of phonological familiarisation, *F (*1, 35) = 29.99, *p* < .001, η_p^2 = .46; Table 3, Figure 6**], replicating the familiarity effect observed in previous studies (Benetello et al., 2015; Savill, Ashton, et al., 2015).

Phonemes from SEM nonwords were also correctly recalled together as whole items more frequently than FAM nonword phonemes [**main effect of semantic training, F (1, 35) = 6.08, p < .05, η_p^2 = .15; Table 3, Figure 6**].

For both comparisons the recall improvement from Day 1 to Day 2 did not significantly interact with training condition.

3.4.2.2. Phonemes correct-in-position forming partially correct nonwords (Individual Phoneme CIP)

Partially correct responses accounted for approximately 26% of target phonemes and 34% of phonemes recalled in any position. The percentage of phonemes recalled as part of partially-correct nonwords was lower for FAM nonwords than NEW nonwords [**main effect of phonological familiarisation, F (1, 35) = 14.69, p < .001, η_p^2 = .30; Table 3, Figure 6**].

The **main effect of semantic training on partially correct responses was not significant** [**p = .11, η_p^2 = .07**].

For both comparisons the reduction in these responses from Day 1 to Day 2 did not significantly interact with training condition.

FIGURE 6 ABOUT HERE

3.4.3. Phonemes out of position

3.4.3.1. Phonemes recalled out of position forming whole item order errors (Whole ORD)

These errors occurred extremely rarely (under 1% of target phonemes and 1% of all target phonemes recalled). However, while the percentage of phonemes recalled as whole item order errors was marginally influenced by simple familiarisation [**F (1, 35) = 3.99, p = .053, η_p^2 =**

.10], these errors were affected by the type of training [**main effect of semantic training, *F* (1, 35) = 16.58, *p* < .001, η_p^2 = .32**]. This main effect was driven by a higher percentage of whole item order errors for SEM nonwords than FAM nonwords on Day 2; as indicated by a training type by day interaction [***F* (1, 35) = 11.32, *p* < .01, η_p^2 = .24; Figure 6**][5]. This suggests that the phonemes of meaningful items were more strongly 'glued together' (Jefferies et al., 2009) after a period of consolidation.

3.4.3.2. Phonemes recalled out of position corresponding to incorrect nonword responses (Individual Phoneme ORD)

Isolated phonemes out of position accounted for approximately 22% of target phonemes and 30% of recalled target phonemes. The rate of phonemes out of position was influenced by prior familiarisation on both days [**main effect of phonological familiarisation, *F* (1, 35) = 23.85, *p* < .001, η_p^2 = .41; Figure 6**].

We observed a main effect of semantic training on the rate of phonemes out of position with fewer positional errors for SEM nonwords [**main effect of semantic training, *F* (1, 35) = 9.22, *p* < .01, η_p^2 = .21; Figure 6**].

For both comparisons, there was no interaction between training condition and day.

TABLE 3 ABOUT HERE

[5] Due to the low frequency of Whole ORD errors, we also ran Wilcoxon signed rank tests to confirm the effects. These followed the same pattern: no familiarisation effect **[both days, *p* > .14]**, but an effect of the type of training on Day 2 [**effect of semantic training: Day 1 *Z* = -0.41, *p* = .683, *d* = -0.06; Day 2 *Z* = -3.46, *p* = .001, *d* = -0.74].**

3.4.4. Non-target phoneme responses

3.4.4.1. Target phonemes replaced by phoneme intrusions (INT) (as a percentage of total phoneme targets).

The intrusions of non-target phonemes accounted for 22% of target phonemes. These errors were influenced by phonological familiarity, with more intrusions for NEW nonwords than FAM on both days [**main effect of phonological familiarisation, F (1, 35) = 26.46, p < .001, η_p^2 = .43; see Figure 5**].

There was no main effect of semantic training [**p = .095, η_p^2 = .08**, **Figure 5; Table S1**].

For both comparisons the reduction in these errors from Day 1 to Day 2 did not significantly interact with training condition.

3.4.4.2 Phoneme omissions (OM) (as a percentage of total phoneme targets)

Phoneme omissions accounted for approximately 4% of target phonemes. There was a non-significant trend for omissions to be made less frequently for FAM nonwords than NEW nonwords [**main effect of phonological familiarisation, p = .074, η_p^2 = .09**]**;** and a significant difference between SEM and FAM conditions [**main effect of semantic training, F (1, 35) = 4.37, p < .05, η_p^2 = .11**]. However both effects corresponded to non-significant contrasts on both days [**all p > .17, all d < 0.19; Figure 5**].

The rate of phoneme omissions over the two days did not significantly interact with training condition.

3.5. Picture naming

These data showed that participants could independently generate very few correct names for the pictured objects, if any at all [**Day 1 correct: *M* 15%, *SD* = 12%**]. However, the capacity for naming considerably improved from Day 1 to Day 2, despite no additional exposure to the images [***t*(35) = -6.94, *p* < .001, *d* = -0.77, Day 2 correct: *M* 26%, *SD* = 16%**].

3.6. Phonological Discrimination task

Repeated measures ANOVAs were used to test the effect of exposure condition (SEM, FAM, NEW) on recognition of the nonwords and their discrimination from phonological neighbour distracters. Analyses of reaction times and accuracy were run separately for targets and phonological neighbour distracters, since we expected opposite effects for these items.

For targets, the accuracy and speed of responses was affected by training condition [**effect of condition on detection accuracy: F (2, 70) = 99.48, $p < .001$, $\eta_p^2 = .74$, and correct RT: F (2, 70) = 10.56, $p < .001$, $\eta_p^2 = .23$; Figure 7**]. Bonferroni-corrected pairwise comparisons showed that these effects were due to slower and less accurate detection of the NEW targets compared to both sets of trained targets (i.e., to FAM and SEM, which did not differ) [Accuracy: NEW vs. FAM, *t* (35) = -11.26, *p* <.001, *d* = -2.38; NEW vs. SEM, *t* (35) = -10.32, *p* < .001, *d* = -2.08; FAM vs. SEM, *t* (35) = -2.14, *p* = .117, *d* = -0.38; correct RT: NEW vs. FAM, *t* (35) = -3.76, *p* = .003, *d* = 0.51; NEW vs. SEM, *t* (35) = 3.36, *p* = .006, *d* = 0.55; FAM vs. SEM, *t* (35) < 1, *d* = 0.04].

The accuracy and speed with which phonological neighbour distracters were rejected was also influenced by training condition [**effect of condition on rejection accuracy: F (2, 70) = 6.70, $p < .01$, $\eta_p^2 = .16$, and correct RT: F (2, 70) = 3.34, $p < .05$, $\eta_p^2 = .09$; Figure 7**]. Bonferroni-corrected pairwise comparisons showed a complementary pattern in rejection accuracy: Rejection of the NEW neighbour distracters was *faster* and *less error-prone* than rejection of phonological neighbours of both sets of trained items (FAM and SEM, which did not differ) [Accuracy: NEW vs. FAM, t (35) = 3.20, p = .009, d = 0.51; NEW vs. SEM, t (35) = 3.21, p = .009, d = 0.52; FAM vs. SEM, t (35) < 1, d = -0.02]. Pairwise differences in neighbour rejection speed did not survive correction for multiple comparisons [correct RT: NEW vs. FAM, t (35) = -1.64, p = .333, d = -0.14; NEW vs. SEM, t (35) = -2.40, p = .066, d = -0.22; FAM vs. SEM, t (35) = -1.10, p = .843, d = -0.09].

FIGURE 7 ABOUT HERE

Sensitivity to the correct phonological form of the test items, as determined by d-prime accuracy for SEM [***M* = 2.35, SD = 0.68**], FAM [***M* = 2.17, SD = 0.65**] and NEW words [***M* = 1.54, SD = 0.58**]**,** differed only between NEW and both sets of trained items (FAM and SEM, again, did not differ) [**effect of condition on d-prime accuracy, F (2, 70) = 38.43, p < .001, η_p^2 = .52**]**.**

3.7. Follow-up analyses of nonwords reliably matched to images in the familiarisation phase

Recognition of SEM and FAM nonwords was similar by the end of testing in the phonological discrimination task, suggesting that the phonological forms of the nonwords in the two conditions had been similarly acquired. As this task was tested after ISR, however, we cannot completely discount the possibility that differences in ISR between SEM and FAM reflected the quality of phonological learning. To assess whether semantic training was likely to affect subsequent ISR even when equating the level of phonological learning across conditions, we compared the recall of SEM and FAM items that were *definitely learnt* (i.e., nonwords that were correctly paired with their images – and correctly rejected in incorrect pairings – on 100% of trials in the second half of training, i.e., achieved maximum d-prime). Multivariate ANOVAs using rank transformed data (Zwick, 1985) tested the effect of training condition on the Day 1 and Day 2 Whole CIP totals for these specific items. These showed that, whether all 100%

image-paired nonwords were assessed (352 SEM nonwords vs. 191 FAM nonwords) or whether a random subsample of the 100% SEM items that matched the number of 100% FAM items per participant were compared (191 SEM vs. 191 FAM nonwords, 4 items on average per condition per participant), the definitely-learnt SEM-trained items tended to be correctly recalled more frequently than the definitely-learnt FAM items, particularly on Day 2 [All definitely-learnt items: Pillai-Bartlett V = .01, F (2, 540) = 2.66, p = .071, η_p^2 = .01; Day 1, F (1, 490) = 2.51, p = .11, η_p^2 = .01, FAM M = 0.96, SD = 1.00, SEM M = 1.11, SD = 1.03; Day 2, F (1, 541) = 5.20, p < .05, η_p^2 = .01, FAM M = 1.20, SD = 1.12, SEM M = 1.42, SD = 1.13; Matched set: Pillai-Bartlett V = .02, F (2, 379) = 4.66, p < .05, η_p^2 = .02; Day 1 = F (1, 380) = 5.96, p < .05, η_p^2 = .02, FAM M = 0.96, SD = 1.00, SEM M = 1.21, SD = 1.06; Day 2, F (1, 380) = 8.32, p < .01, η_p^2 = .02, FAM M = 1.20, SD = 1.12, SEM M = 1.53, SD = 1.18].

4. Discussion

This study examined whether word meaning contributes to the phonological stability of words in verbal short-term memory (STM), independently of confounding effects of phonological-lexical familiarity. This work was motivated by research with semantic dementia patients showing that the availability of semantic knowledge can influence the stability of a word's phonological form; however, these effects might be explained by corresponding differences in the frequency of word use, since semantic dementia patients are likely to use well-understood words more frequently than degraded words in everyday conversation. Our approach was to examine STM at a phoneme level for newly trained nonwords with and without associated semantic information in healthy individuals, both immediately after training and after an overnight consolidation period. We demonstrate for the first time that having an associated meaning leads to more stable phonological output in the absence of differences in the amount of prior exposure to the phonological-lexical form, thus providing important converging support for the semantic binding hypothesis in healthy participants. Moreover we show that these effects start to take hold quickly, on the day of training. For the most part, effects of training did not interact with day of testing.

In summary, the main findings were (1) that immediate serial recall (ISR) was more phonologically coherent following phonological familiarisation (relative to new unfamiliar nonwords): more phonemes were recalled overall and more of these phonemes were produced in the correct place in the sequence as complete target items, and (2) nonwords associated with semantic information also showed a recall advantage relative to phonologically-familiar nonwords without such an association. While the availability of semantic information may have facilitated initial acquisition of the nonwords in the phonological familiarisation phase, two lines of evidence suggest that the short-term recall advantage for meaningful items reflected a semantic effect on phoneme binding in ISR beyond effects on phonological learning. First, the nonwords with and without semantic referents produced equivalent performance in a phonological discrimination task at the end of testing, suggesting that the phonological-lexical forms were specified to a similar degree in the two conditions. This perhaps reflected the fact that participants' exposure to the phonological forms of the nonwords was matched for the familiar and meaningful items. Secondly, the semantic association advantage for items in ISR remained even when performance in the training task was matched across conditions by selecting items for analysis that showed 100% accuracy in nonword-image pairings by the second half of the phonological familiarisation phase.

Our results are consistent with the 'semantic binding' account, in suggesting that both phonological-lexical and semantic knowledge maintain the integrity of the phonological trace. First, at a *phonological-lexical level*, familiarised nonwords showed benefits in ISR which reflected the opportunity to learn the phoneme sequence comprising each item, relative to unfamiliar nonwords. The increase in whole target items recalled, plus the lower number of phonological errors, reflects the successful binding of phonemes into target responses by the phonological system during STM. This pattern is readily explained by theoretical accounts in which phonological-lexical long-term knowledge comprises patterns of phonological activation learnt from previous buffering of the information between speech perception and production architectures (Jacquemot & Scott, 2006); by this view, an item's phonological trace in STM should be more stable when this sequence of phonemes corresponds with a pattern previously encountered and learned by the phonological system. In the case of maintaining a sequence of items, such as nonword lists in ISR, the more familiar the phonological activation patterns of the individual items are, the lower the risk that their phonemes will migrate and recombine with those from different items. In addition to these effects of phonological-lexical knowledge on the stability of a phoneme sequence in short-term memory, the semantic binding hypothesis also predicts a second source of constraint which is provided by the interaction between semantic representations and the phonological system.

The primary question motivating this study was whether semantic information supports the phonological maintenance of whole word traces in verbal short-term memory, even when phonological-lexical experience is equated. Familiarised items that were associated with semantic information showed better recall than familiarised forms without an associated meaning: this recall advantage was apparent immediately after familiarisation in terms of fewer target phonemes recalled out of position and, by the next day, corresponded to more whole items successfully recalled and fewer phonological errors. These data show that the stabilising influence of associated semantic information on phonological-lexical forms is not accounted for by relative phonological familiarity alone. The effects of semantic training on phoneme level accuracy can be accommodated by accounts that acknowledge that phonological processing occurs in a semantic context and therefore its stability can be influenced by semantic information (Dell et al., 1997; Plaut & Kello, 1999). While the present results are broadly compatible with different language-based frameworks (which envisage that temporary activation of linguistic representations is the basis for verbal STM), it is a specific prediction of the semantic binding hypothesis that conceptual representations provide a stabilising influence on the phonological system that reduces phoneme migration errors (Patterson et al., 1994). This effect occurs because phonological sequences are acquired and processed in the context of meaning: therefore activation of semantic representations can contribute to the pattern completion properties within the phonological system. These effects are likely to be particularly important for the maintenance of multiple items in STM, since in this situation, the phonemes of different items can be recombined. We (and others) have previously shown that when STM is challenged, and it is necessary to hold on to several fragile phonological representations, the task will strongly tap the ability of people to keep the phonemes from the same item together (Jefferies, Frankish, et al., 2006a; Jefferies, Frankish, & Noble, 2011; Jefferies, Grogan, Mapelli, & Isella, 2012; Jefferies et al., 2008; Knott et al., 1997; Majerus et al., 2007; Patterson et al., 1994; Savill, Ashton, et al., 2015; Savill, Metcalfe, Ellis & Jefferies, 2015). By this token, the semantic binding account would also anticipate that the availability of semantic representations at the familiarisation phase would influence acquisition of the phonological forms, since phonological learning requires phonological maintenance, and this would be more veridical when there is an associated meaning. More broadly, these observations fit with the view that the main purpose of STM is to allow the production and comprehension of multi-word utterances (Acheson & MacDonald, 2009; Hickok & Poeppel, 2007; Jacquemot & Scott, 2006).

In an alternative perspective, short-term phonological maintenance operates independently of long-term linguistic representations (including semantic knowledge). According to this view, the relative availability of these long-term memory representations at a late stage, during recall, determines the likelihood of successful reconstruction of the phonological trace ('redintegration'). Long-term phonological-lexical knowledge of word forms could allow missing phonological elements to be generated, while semantic information might influence the availability of relevant phonological-lexical representations used in this process. It follows that redintegration would have its key effects on serial recall at the whole item level (Gathercole, Pickering, Hall, & Peaker, 2001) while semantic binding should influence serial recall accuracy at the intra-item phoneme level (Jefferies, Frankish, et al., 2006a). For example, a redintegration-based explanation for the semantic-training advantage in whole item recall could reflect the availability of clear visual referents in identifying lexical representations for the reconstruction process and/or some capacity to generate a spoken response from primed semantic representations in LTM. Compared to nonwords that were associated with blurred images, nonwords with clear image referents and pre-trained meanings were at an advantage during the familiarisation task and also in subsequent free recall. These differences were probably driven by the relative visual distinctiveness of the clear object referents compared to more confusable blurred images, which were a less potent retrieval cue (e.g., Nelson, Reed, & Walling, 1973; Paivio & Csapo, 1973). The extent to which the distinctiveness of the images may have then contributed to the condition effects in ISR remains an open question. It seems unlikely, however, that the availability of a clear visual referent alone could account for all aspects of the semantic ISR advantage. Savill, Ashton, et al. (2015) used the images and nonwords in the present study but without the semantic training step, and did not find differences in serial recall performance between nonwords with clear and blurred image referents. Thus, while the semantic information available in the images may comprise the main component of the semantic advantage, differences in visual distinctiveness alone are unlikely to explain the better recall for the semantically-trained nonwords. In addition, the availability of meaningful referents to constrain late-stage redintegration should affect the recall of *whole* items in and out of sequence: it is harder to account for effects at the level of single phonemes within this framework without modification. For example, ISR for semantically-trained nonwords was characterised by fewer individual phoneme order errors and fewer non-target phonemes (intrusions) – yet since these phoneme-level responses were not part of correct target items, these effects apparently occurred in situations in which redintegration had failed.

Our findings here and in previous investigations (Jefferies, Frankish, et al., 2006a; Jefferies et al., 2009; Savill, Metcalfe, et al., 2015; Savill, Ashton, et al., 2015) indicate that theoretical frameworks that seek to explain the effects of long-term language knowledge on STM need to incorporate phoneme-level representations, allowing phonemes to migrate as the phonological trace decays. *If* the 'redintegration framework' were to adopt this architecture, the implications for changes in phoneme order errors would largely overlap with the predictions of the semantic binding account (since increases in redintegration of complete target items would lead to decreases in phoneme order errors when migrated phonemes are reinstated). As the current redintegration view stands, however, the semantic binding hypothesis offers a more straightforward account of our data. An implication common to these theories is that phonological-lexical and semantic influences on STM are unlikely to be comparable in magnitude. This is important to consider in light of previous null semantic effects in STM (e.g., Benetello, Cecchetto, & Papagno, 2015; Papagno et al., 2013). The impact of brief prior phonological-lexical experience, relative to no experience, is striking and robust (as found here and in previous reports, e.g., Savill, Ashton, et al., 2015); in comparison, semantic effects on STM are relatively subtle, presumably because these representations are not central to phonological processing.

Following previous research on word learning that has shown that trained novel words start to behave more like real words after a period of consolidation, particularly after sleep (e.g., when lexical competition effects emerge, Davis & Gaskell, 2009), we compared the effects of phonological familiarity and semantic training both immediately and after a period of consolidation. We demonstrated rapid effects of semantic training, in terms of STM performance, within a few minutes of training. On the whole, the effects of training did not interact with day, suggesting that overnight consolidation is not critical to these effects. Nevertheless, semantically-trained nonwords did show a trend towards more word-like behaviour the next day, in that they were more likely to be recalled as a whole item out of sequence (Jefferies, Frankish, et al., 2006a). Thus, the timing of the recall test does not explain the null result of semantic effects in Benetello et al. (2015). Instead, it might be that the nature of our semantic training task, which involved the acquisition of new multidimensional object representations that were subsequently associated with a new lexical form that facilitated the semantic effect. Differences in the sensitivity and structure of the ISR task between that study and ours, and differences in the studies' respective phonological familiarisation phases may have also been factors (i.e., regarding our use of feedback in the context of correct and incorrect nonword-image pairings, which were intended to encourage robust nonword-image pairings, and the control over the number of nonword presentations between conditions).

Although neuroimaging studies of word learning have not always found a differential neural effect of semantic information (Cornelissen et al., 2004; Grönholm, Rinne, Vorobyev, & Laine, 2005; Hultén, Vihla, Laine, & Salmelin, 2009), a recent study found that novel words given richer meanings more strongly engaged regions associated with semantic processing, such as left anterior MTG and left angular gyrus (Ferreira, Göbel, Hymers, & Ellis, 2015). Accordingly, increased semantic binding of phonological representations may be explained in terms of relatively stronger co-activation of networks supporting semantic representation (including anterior temporal areas atrophied in patients with semantic dementia) with networks specialised for phonological storage. At the functional level, interactivity between these networks might be observed in interfacing regions: for example, posterior MTG has been proposed as a 'ventral lexical' store that binds phonological information stored in STG to amodal conceptual representations in the temporal pole (Gow, 2012; Hickok & Poeppel, 2007).

Our findings also fit well with a recent ERP study that examined the mismatch negativity (MMN) response to spoken nonword oddballs that had been associated with either a stable semantic referent or that had no consistent association (Hawkins, Astle, & Rastle, 2015). The stable referent item elicited a stronger MMN effect in an auditory stream, suggesting that semantic training had facilitated acoustic-phonological processing for these items. On the day of training, the size of the MMN effect was related to explicit memory for the semantic associations, while the next day the MMN effect was maintained but not linked to explicit semantic knowledge. This study raises the interesting possibility that the ISR advantage we observed on Days 1 and 2 might be underpinned by different mechanisms, possibly relating to strategic redintegration strengthened by the availability of explicit semantic referents on Day 1 (e.g., benefitting largely from support from traces in episodic LTM of perceptual associations), and a more implicit effect on phonological stability on Day 2 (e.g., with greater support from developing amodal representations in semantic LTM). Moreover, the absence of strong interactions between training condition and day in this study does not preclude the possibility that future work using longer periods of consolidation would reveal such an effect. In any case, these two studies taken together leave little doubt that meaning supports the acquisition of stable phonological knowledge and that these phonological forms are subsequently retained more accurately in STM – consequently, we conclude that, contrary to recent claims, meaning is not “useless”.

5. Acknowledgments

This research was supported by a European Research Council Grant (283530 – SEMBIND). We thank Eleanor Whittaker and Emily Ball for their assistance with data collection.

6. References

Acheson, D. J., & MacDonald, M. C. (2009). Verbal working memory and language production: Common approaches to the serial ordering of verbal information. *Psychological Bulletin*, *135*, 50–68. doi:10.1037/a0014411.

Acheson, D. J., Postle, B. R., & MacDonald, M. C. (2010). The interaction of concreteness and phonological similarity in verbal working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*, 17–36. doi:10.1037/a0017679.The

Allen, R. J., & Hulme, C. (2006). Speech and language processing mechanisms in verbal serial recall. *Journal of Memory and Language*, *55*, 64–88. doi:10.1016/j.jml.2006.02.002

Baddeley, A. D. (1986). *Working Memory*. Oxford, UK: Clarendon Press.

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, *4*, 417–423. doi:10.1016/S1364-6613(00)01538-2

Baddeley, A. D., & Hitch, G. J. (1974). Working Memory. In G. H. Bower (Ed.), *Psychology of learning and motivation: Advances in research and theory* (Volume 8., pp. 47–89). New York, NY: Academic Press.

Baddeley, A. D., Hitch, G. J., & Allen, R. J. (2009). Working memory and binding in sentence recall. *Journal of Memory and Language*, *61*, 438–456. doi:10.1016/j.jml.2009.05.004

Benetello, A., Cecchetto, C., & Papagno, C. When meaning is useless. *Memory*. doi:10.1080/09658211.2014.945939

Campoy, G., & Baddeley, A. D. (2008). Phonological and semantic strategies in immediate serial recall. *Memory*, *16*, 329–340. doi:10.1080/09658210701867302

Caza, N., & Belleville, S. (1999). Semantic contribution to immediate serial recall using an unlimited set of items: evidence for a multi-level capacity view of short-term memory. *International Journal of Psychology*, *34*, 334–338. doi:10.1080/002075999399657

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. doi:10.1037/0278-7393.33.5.970

Cornelissen, K., Laine, M., Renvall, K., Saarinen, T., Martin, N., & Salmelin, R. (2004). Learning new names for new objects : Cortical effects as measured by magnetoencephalography q, *89*, 617–622. doi:10.1016/j.bandl.2003.12.007

Davis, M. H., & Gaskell, M. G. (2009). A complementary systems account of word learning: neural and behavioural evidence. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*, 3773–3800. doi:10.1098/rstb.2009.0111

Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. a. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review*, *104*, 801–838. doi:10.1037/0033-295X.104.4.801

Ellis, A. W. (1980). Errors in speech and short-term memory: The effects of phonemic similarity and syllable position. *Journal of Verbal Learning and Verbal Behavior*, *19*, 624–634. doi:10.1016/S0022-5371(80)90672-6

Ferreira, R. a., Göbel, S. M., Hymers, M., & Ellis, A. W. (2015). The neural correlates of semantic richness: Evidence from an fMRI study of word learning. *Brain and Language*, *143*, 69–80. doi:10.1016/j.bandl.2015.02.005

Forde, E. M. E., & Humphreys, G. W. (2002). The role of semantic knowledge in short-term memory. *Neurocase*, *8*, 13–27. doi:10.1093/neucas/8.1.13

Foygel, D., & Dell, G. S. (2000). Models of Impaired Lexical Access in Speech Production. *Journal of Memory and Language*, *43*, 182–216. doi:10.1006/jmla.2000.2716

Gathercole, S. E., Pickering, S. J., Hall, M., & Peaker, S. M. (2001). Dissociable lexical and phonological in influences on serial recognition and serial recall. *The Quarterly Journal of Experimental Psychology*, *54A*, 1–30. doi:10.1080/02724980042000002

Gow, D. W. (2012). The cortical organization of lexical knowledge: A dual lexicon model of spoken language processing. *Brain and Language*, *121*, 273–288. doi: 10.1016/j.bandl.2012.03.005

Grönholm, P., Rinne, J. O., Vorobyev, V., & Laine, M. (2005). Naming of newly learned objects : A PET activation study. *Cognitive Brain Research*, *25*, 359–371. doi:10.1016/j.cogbrainres.2005.06.010

Hawkins, E., Astle, D. E., & Rastle, K. (2015). Semantic Advantage for Learning New Phonological Form Representations. *Journal of Cognitive Neuroscience*, *27*, 775–786. doi:10.1162/jocn

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews. Neuroscience*, *8*, 393–402. doi:10.1038/nrn2113

Hoffman, P., Jefferies, E., Ehsan, S., Jones, R. W., & Lambon Ralph, M. A. (2009). Semantic memory is key to binding phonology: converging evidence from immediate serial recall in semantic dementia and healthy participants. *Neuropsychologia*, *47*, 747–760. doi:10.1016/j.neuropsychologia.2008.12.001

Hultén, A., Vihla, M., Laine, M., & Salmelin, R. (2009). Accessing newly learned names and meanings in the native language. *Human Brain Mapping*, *30*, 976–989. doi:10.1002/hbm.20561

Jacquemot, C., & Scott, S. K. (2006). What is the relationship between phonological short-term memory and speech processing? *Trends in Cognitive Sciences*, *10*, 480–486. doi:10.1016/j.tics.2006.09.002

Jefferies, E., Crisp, J., & Lambon Ralph, M. A. (2006). The impact of phonological or semantic impairment on delayed auditory repetition: Evidence from stroke aphasia and semantic dementia. *Aphasiology*, *20*, 963–992. doi:10.1080/02687030600739398

Jefferies, E., Frankish, C., & Noble, K. (2011). Strong and long: effects of word length on phonological binding in verbal short-term memory. *Quarterly Journal of Experimental Psychology*, *64*, 241–60. doi:10.1080/17470218.2010.495159

Jefferies, E., Frankish, C. R., & Lambon Ralph, M. A. (2006a). Lexical and semantic binding in verbal short-term memory. *Journal of Memory and Language*, *54*, 81–98. doi:10.1016/j.jml.2005.08.001

Jefferies, E., Frankish, C. R., & Lambon Ralph, M. A. (2006b). Lexical and semantic influences on item and order memory in immediate serial recognition: evidence from a novel task. *Quarterly Journal of Experimental Psychology*, *59*, 949–964. doi:10.1080/02724980543000141

Jefferies, E., Frankish, C. R., & Noble, K. (2009). Lexical coherence in short-term memory: strategic reconstruction or "semantic glue"? *Quarterly Journal of Experimental Psychology*, *62*, 1967–1982. doi:10.1080/17470210802697672

Jefferies, E., Grogan, J., Mapelli, C., & Isella, V. (2012). Paced reading in semantic dementia: word knowledge contributes to phoneme binding in rapid speech production. *Neuropsychologia*, *50*, 723–32. doi:10.1016/j.neuropsychologia.2012.01.006

Jefferies, E., Hoffman, P., Jones, R., & Lambon Ralph, M. A. (2008). The impact of semantic impairment on verbal short-term memory in stroke aphasia and semantic dementia: A comparative study. *Journal of Memory and Language*, *58*, 66–87. doi:10.1016/j.jml.2007.06.004

Jefferies, E., Jones, R., Bateman, D., & Lambon Ralph, M. A. (2004). When does word meaning affect immediate serial recall in semantic dementia? *Cognitive, Affective & Behavioral Neuroscience*, *4*, 20–42. doi:10.3758/CABN.4.1.20

Knott, R., Patterson, K. E., & Hodges, J. (1997). Lexical and semantic binding effects in short-term memory: Evidence from semantic dementia. *Cognitive Neuropsychology*, *14*, 1165–1216. doi:10.1080/026432997381303

Knott, R., Patterson, K. E., & Hodges, J. R. (2000). The role of speech production in auditory-verbal short-term memory: evidence from progressive fluent aphasia. *Neuropsychologia*, *38*, 125–142. doi:10.1016/S0028-3932(99)00069-X

Lewis-Peacock, J. A., Drysdale, A. T., Oberauer, K., & Postle, B. R. (2012). Neural evidence for a distinction between short-term memory and the focus of attention. *Journal of Cognitive Neuroscience*, *24*, 61–79. doi:10.1162/jocn_a_00140

Lin, C.-C., & Yang, C.-M. (2014). Evidence of sleep-facilitating effect on formation of novel semantic associations: An event-related potential (ERP) study. *Neurobiology of Learning and Memory*, *116C*, 69–78. doi:10.1016/j.nlm.2014.08.011

Majerus, S. (2013). Language repetition and short-term memory: an integrative framework. *Frontiers in Human Neuroscience*, *7*, 357. doi:10.3389/fnhum.2013.00357

Majerus, S., Norris, D. G., & Patterson, K. E. (2007). What does a patient with semantic dementia remember in verbal short-term memory? Order and sound but not words. *Cognitive Neuropsychology*, *24*, 131–151. doi:10.1080/02643290600989376

Majerus, S., & van der Linden, M. (2003). Long-term memory effects on verbal short-term memory: A replication study. *British Journal of Developmental Psychology*, *21*, 303–310. doi:10.1348/026151003765264101

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. doi:10.1037//0033-295X.102.3.419

Miller, L. M., & Roodenrys, S. (2009). The interaction of word frequency and concreteness in immediate serial recall. *Memory & Cognition*, *37*, 850–65. doi:10.3758/MC.37.6.850

Murdock, B. B. (1976). Item and order information in short-term serial memory. *Journal of Experimental Psychology: General*, *105*, 191–216. doi:10.1037/0096-3445.105.2.191

Nelson, D. L., Reed, V. S., & Walling, J. R. (1973). Pictorial Superiority Effect. *Journal of Experimental Psychology: Human Learning & Memory*, *2*, 523–528. doi:10.1037/0278-7393.2.5.523

Paivio, A., & Csapo, K. (1973). Picture superiority in free recall: Imagery or dual coding? *Cognitive Psychology*, *5*, 176–206. doi:10.1016/0010-0285(73)90032-7

Papagno, C., Vernice, M., & Cecchetto, C. (2013). Phonology without semantics? Good enough for verbal short-term memory. Evidence from a patient with semantic dementia. *Cortex*, *49*, 626–636. doi:10.1016/j.cortex.2012.04.015

Patterson, K. E., Graham, N., & Hodges, J. R. (1994). The impact of semantic memory loss on phonological representations. *Journal of Cognitive Neuroscience*, *6*, 57–69. doi:10.1162/jocn.1994.6.1.57

Peters, F., Majerus, S., De Baerdemaeker, J., Salmon, E., & Collette, F. (2009). Impaired semantic knowledge underlies the reduced verbal short-term storage capacity in Alzheimer's disease. *Neuropsychologia*, *47*, 3067–73. doi:10.1016/j.neuropsychologia.2009.07.002

Plaut, D. C., & Kello, C. T. (1999). The emergence of phonology from the interplay of speech comprehension and production: a distributed connectionist approach. In B. MacWhinney (Ed.), *The emergence of language* (pp. 381–415). Mahweh, NJ: Erlbaum.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: computational principles in quasi-regular domains. *Psychological Review*, *103*, 56–115. doi:10.1037/0033-295X.103.1.56

Poirier, M., & Saint-Aubin, J. (1995). Memory for related and unrelated words: further evidence on the influence of semantic factors in immediate serial recall. *The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *48*, 384–404. doi:10.1080/14640749508401396

Poirier, M., & Saint-Aubin, J. (1996). Immediate serial recall, word frequency, item identity and item position. *Canadian Journal of Experimental Psychology = Revue Canadienne de Psychologie Experimentale*, *50*, 408–412. doi:10.1037/1196-1961.50.4.408

Roche, J., Tolan, G. A., & Tehan, G. (2011). Concreteness effects in short-term memory: a test of the item-order hypothesis. *Canadian Journal of Experimental Psychology*, *65*, 245–53. doi:10.1037/a0024693

Romani, C., McAlpine, S., & Martin, R. C. (2008). Concreteness effects in different tasks: implications for models of short-term memory. *Quarterly Journal of Experimental Psychology*, *61*, 292–323. doi:10.1080/17470210601147747

Roodenrys, S., & Quinlan, P. T. (2000). The effects of stimulus set size and word frequency on verbal serial recall. *Memory*, *8*, 71–8. doi:10.1080/096582100387623

Saint-Aubin, J., Guérard, K., Chamberland, C., & Malenfant, A. (2013). Delineating the contribution of long-term associations to immediate recall. *Memory*, 37–41. doi:10.1080/09658211.2013.794242

Saint-Aubin, J., & Poirier, M. (1999). Semantic similarity and immediate serial recall: is there a detrimental effect on order information? *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology*, *52*, 367–394. doi:10.1080/713755814

Savill, N., Ashton, J., Gugliuzza, J., Poole, C., Sim, Z., Ellis, A. W., & Jefferies, E. (2015). tDCS to temporoparietal cortex during familiarisation enhances the subsequent phonological coherence of nonwords in immediate serial recall. *Cortex*, *63*, 132–144. doi:10.1016/j.cortex.2014.08.018

Savill, N., Metcalfe, T., Ellis, A. W., & Jefferies, E. (2015). Semantic categorisation of a word supports its phonological integrity in verbal short-term memory. *Journal of Memory and Language*, *84*, 128–138. doi:10.1016/j.jml.2015.06.003

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, *96*, 523–68. doi:10.1037/0033-295X.96.4.523

Tamminen, J., & Gaskell, M. G. (2013). Novel word integration in the mental lexicon : Evidence from unmasked and masked semantic priming. *The Quarterly Journal of Experimental Psychology*, *66*, 1001–1025. doi:10.1080/17470218.2012.724694

Tse, C.-S., Li, Y., & Altarriba, J. (2011). The effect of semantic relatedness on immediate serial recall and serial recognition. *The Quarterly Journal of Experimental Psychology*, *64*, 2425–2437. doi: 10.1080/09658210701467186

Walker, I., & Hulme, C. (1999). Concrete words are easier to recall than abstract words: Evidence for a semantic contribution to short-term serial recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 1256–1271. doi:10.1037/0278-7393.25.5.1256

Wilshire, C. E. (1998). Serial order in phonological encoding: an exploration of the “word onset effect” using laboratory-induced errors. *Cognition*, *68*, 143–66. doi:10.1016/S0010-0277(98)00045-6

Wilshire, C. E., Keall, L. M., & O’Donnell, D. J. (2010). Semantic contributions to immediate serial recall: evidence from two contrasting aphasic individuals. *Neurocase*, *16*, 331–351. doi:10.1080/13554791003620256

Zwick, R. (1985). Nonparametric One-Way Multivariate Analysis of Variance: A Computational Approach Based on the Pillai-Bartlett Trace. *Psychological Bulletin*, *97*, 148–152. doi:10.1037/0033-2909.97.1.148

Figure Legends

Figure 1. Schedule of training and test phases (shown in blue and green respectively)

Figure 2. Overview of semantic training phase. (A) Two sample trials in the semantic training task: The first two screens show a feature training trial, in which participants identified which 'fact' belonged to the image, and the following two screens show a blurred image trial, which participants passively viewed. Over the course of the task participants saw facts pertaining to the object's function (e.g., "It rests cooking utensils between stirs of food"), its context (e.g., "When in use, these are found on kitchen saucepans") and its mechanism (e.g., "This is used by clamping it into position"). Over the course of training, each fact was presented three times. (B) Participants were able to learn the fact associations: they showed increasingly higher accuracy and faster RTs over the course of the semantic training task.

Figure 3. An illustrative ISR trial (taken from Savill, Ashton, et al., 2015)

Figure 4. Overview of phonological familiarisation phase. (A) An illustrative trial sequence within the phonological familiarisation task, adapted from Savill, Ashton, et al. (2015). (B) Improvements in accuracy and correct reaction times with each presentation of the nonword-image pairs.

Figure 5. ISR data showing the percentage of target phonemes correctly recalled in any position (CAP) in each training condition (NEW, FAM, SEM), on Day 1 and Day 2. Data are expressed as a percentage of total target phonemes. The CAP score comprises Whole CIP, Ind. Phon CIP, Whole ORD and Ind. Phon. ORD. The percentages of target phonemes not recalled in any position (i.e., phoneme intrusions and omissions) are shown faintly above each CAP percentage bar. Black connectors detail significant pairwise comparisons for CAP phonemes. NEW = untrained nonwords, FAM = familiarised-only nonwords; SEM = semantically-trained nonwords.

Figure 6. Target related responses in ISR for each training condition on Day 1 and Day 2. Data for each of the response categories are expressed as a percentage of total phonemes recalled in or out of position (CAP). (NEW = untrained nonwords, FAM = familiarised-only nonwords; SEM = semantically-trained nonwords; Whole CIP = phoneme correct in position in the context of a whole correct nonword; Ind. Phon. CIP = phoneme correct in position in the context of a partially correct nonword; Whole ORD = phoneme within whole item order error; Ind. Phon. ORD = phoneme order error as part of an incorrect nonword. Target-unrelated intrusion and omission errors, which were analysed as a percentage of total phoneme targets, are shown in Figure 5).

Figure 7. Accuracy and reaction times in the phonological discrimination task. Target detection was better for FAM and SEM trained items relative to untrained NEW items (that had previously only been presented in ISR); while phonological neighbours of both sets of trained items were rejected with greater difficulty than neighbours of the less familiar NEW items. These findings suggest similarly increased phonological competition effects for both sets of trained targets.

Table 1. *Matched phonotactic properties of the nonwords sets*

		Set A	**Set B**	**Set C**	***F***	***p***
Sum. Phon. prob.	*M*	.191	.185	.186	.041	.960
	SD	.070	.076	.069		
Sum. Biphone prob.	*M*	.008	.008	.007	.275	.760
	SD	.006	.007	.005		
Word-likeness ratings	*M*	2.981	2.983	2.975	0.002	0.998
	SD	0.487	0.503	0.469		

Note. Sets A, B and C refer to the three sets of nonwords that corresponded to NEW, FAM or SEM conditions, which were rotated between participants. Phonotactic probabilities were calculated according to Vitevitch & Luce (2004). Word-likeness ratings correspond to ratings given by an independent set of participants regarding how likely they felt each nonword could be an English word (Likert scale of 1 to 5; n = 5). Sum. Phon. Prob = summed position-specific phoneme probability. Sum. Biphone prob. = summed position-specific biphone probability.

Table 2. *Average number of nonwords produced in Free Recall*

		Day 1		Day 2	
		FAM	SEM	FAM	SEM
Whole nonwords free recalled	*M*	0.69	1.50	1.11	1.72
	SD	0.89	1.56	1.06	1.60
Near-miss nonwords free recalled	*M*	0.53	0.75	0.64	1.17
	SD	0.74	1.11	0.90	1.16
No correct or near-miss recall	*% of pts*	11.11	19.44	22.22	22.22

Note. Near-miss recall = responses that differed from a trained nonword by a single phoneme. pts = participants.

Table 3. *Full ANOVA results for target phonemes recalled (CAP phonemes)*

Note. NEW = untrained nonwords, FAM = familiarised-only nonwords; SEM = semantically-trained nonwords. All *df* = 1, 35.

		Whole CIP	Ind. Phon. CIP	Whole ORD	Ind. Phon. ORD
Familiarisation (NEW vs FAM)	FAM Effect	$F = 29.99$, $p < .001$ $\eta_p^2 = .46$	$F = 14.69$, $p < .001$ $\eta_p^2 = .30$	$F = 3.99$, $p = .053$ $\eta_p^2 = .10$	$F = 23.85$, $p < .001$ $\eta_p^2 = .41$
	DAY	$F = 26.94$, $p < .001$ $\eta_p^2 = .44$	$F = 5.30$, $p = .027$ $\eta_p^2 = .13$	$F < 1$, $p = .394$ $\eta_p^2 = .02$	$F = 26.72$, $p < .001$ $\eta_p^2 = .43$
	FAM × DAY	$F = 2.13$, $p = .150$ $\eta_p^2 = .06$	$F = 1.09$, $p = .303$ $\eta_p^2 = .03$	$F < 1$, $p = .914$ $\eta_p^2 = .00$	$F = 1.433$, $p = .239$ $\eta_p^2 = .04$
Semantic Training (FAM vs SEM)	SEM Effect	$F = 6.08$, $p = .019$ $\eta_p^2 = .15$	$F = 2.69$, $p = .110$ $\eta_p^2 = .07$	$F = 16.58$, $p < .001$ $\eta_p^2 = .32$	$F = 9.22$, $p = .004$ $\eta_p^2 = .21$
	DAY	$F = 39.81$, $p < .001$ $\eta_p^2 = .53$	$F = 20.44$, $p < .001$ $\eta_p^2 = .37$	$F = 10.11$, $p = .003$ $\eta_p^2 = .22$	$F = 48.44$, $p < .001$ $\eta_p^2 = .58$
	SEM × DAY	$F < 1$, $p = .600$ $\eta_p^2 = .01$	$F = 2.42$, $p = .129$ $\eta_p^2 = .07$	$F = 11.32$, $p = .002$ $\eta_p^2 = .24$	$F < 1$, $p = .580$ $\eta_p^2 = .01$

Fig1.

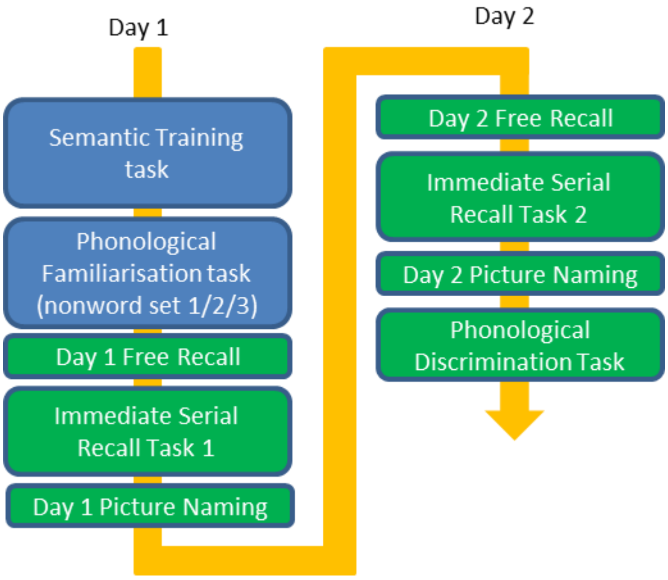

Day 1
Day 2
Semantic Training task
Phonological Familiarisation task (nonword set 1/2/3)
Day 1 Free Recall
Immediate Serial Recall Task 1
Day 1 Picture Naming
Day 2 Free Recall
Immediate Serial Recall Task 2
Day 2 Picture Naming
Phonological Discrimination Task

Fig2

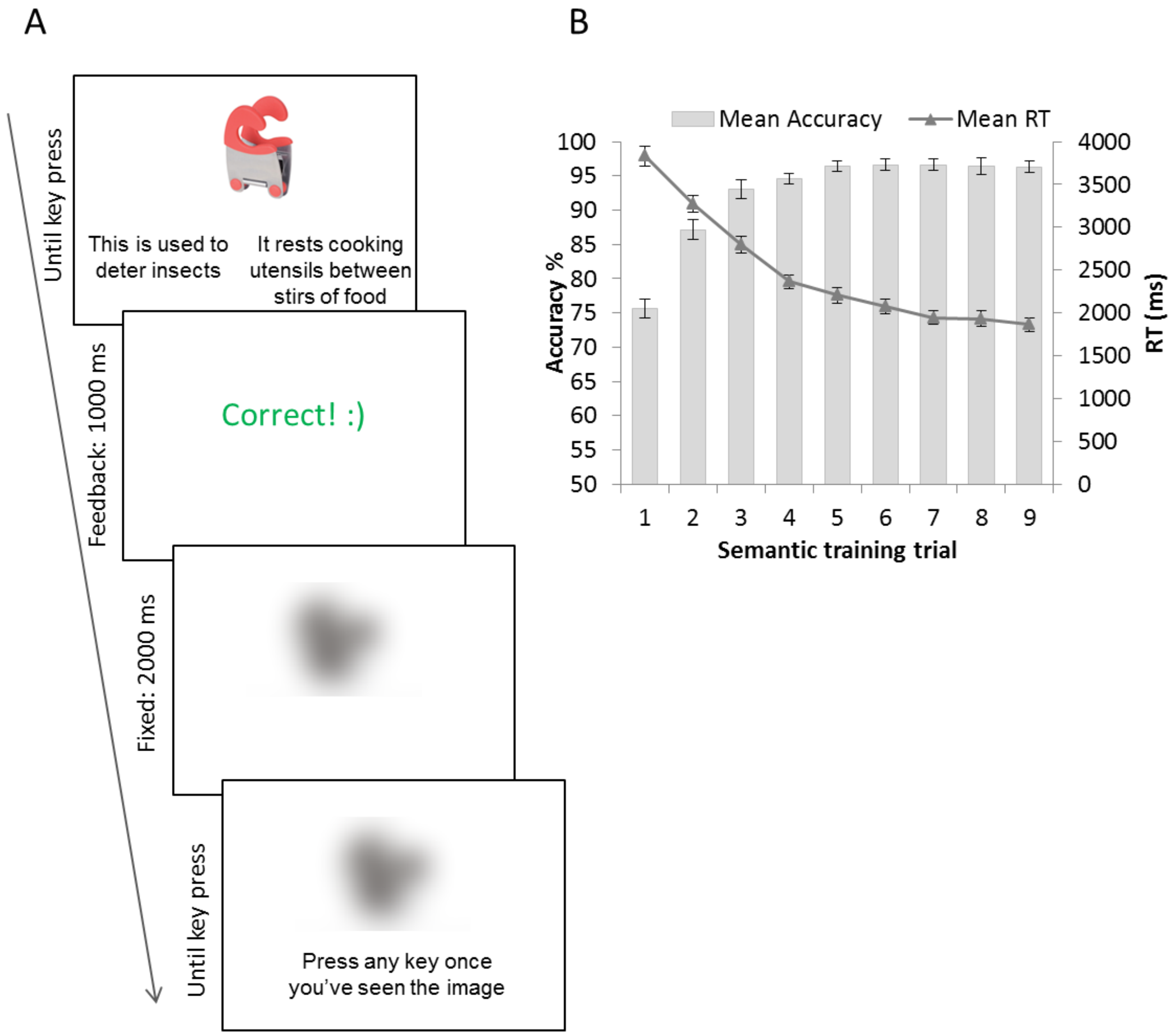

A
Until key press
This is used to deter insects
It rests cooking utensils between stirs of food
Feedback: 1000 ms
Correct! :)
Fixed: 2000 ms
Until key press
Press any key once you've seen the image
B
Mean Accuracy
Mean RT
Accuracy %
RT (ms)
Semantic training trial
100
95
90
85
80
75
70
65
60
55
50
4000
3500
3000
2500
2000
1500
1000
500
0
1 2 3 4 5 6 7 8 9

Fig3

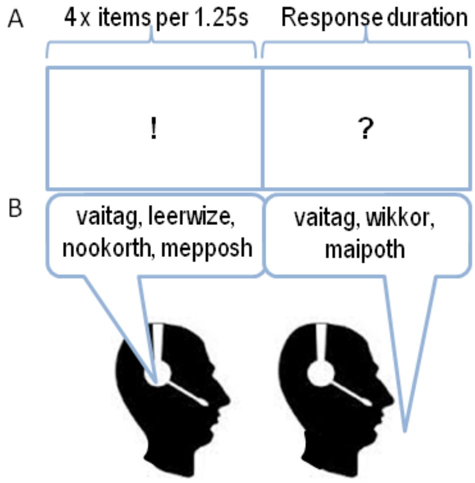

A
4 x items per 1.25s
Response duration
!
?
B
vaitag, leerwize, nookorth, mepposh
vaitag, wikkor, maipoth

Fig4

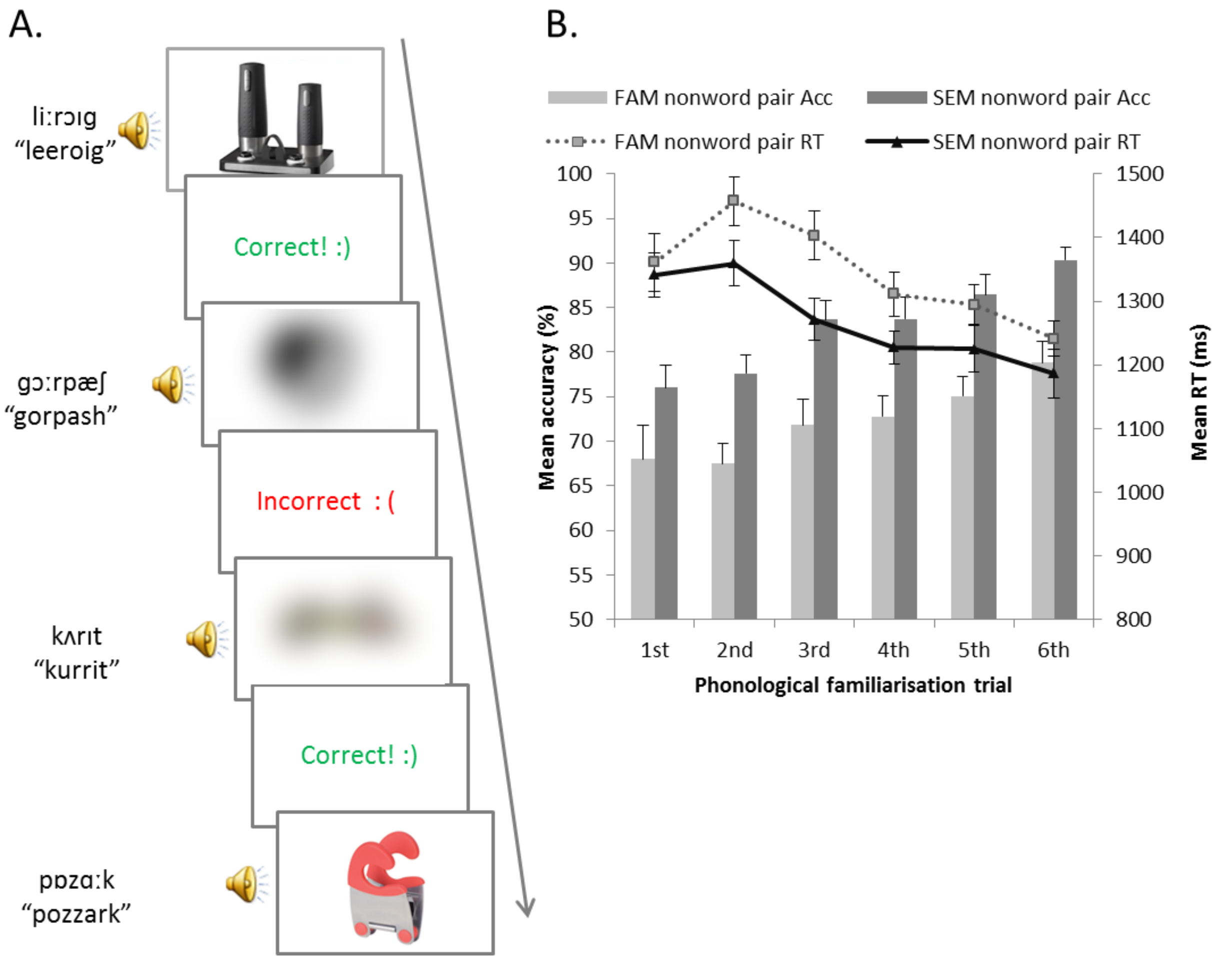
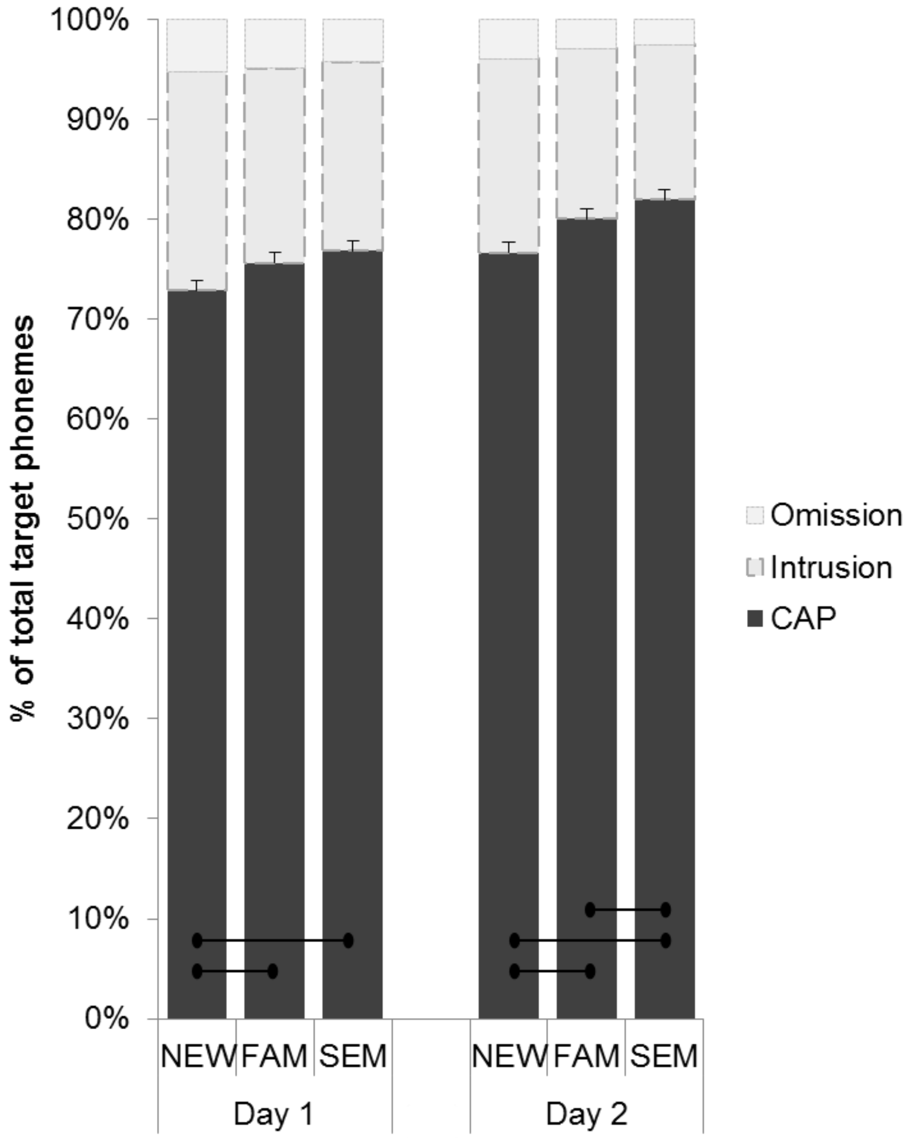

A.
liːrɔɪɡ
“leeroig”
Correct! :)
ɡɔːrpæʃ
“gorpash”
Incorrect : (
kʌrɪt
“kurrit”
Correct! :)
pɒzɑːk
“pozzark”
B.
FAM nonword pair Acc
SEM nonword pair Acc
FAM nonword pair RT
SEM nonword pair RT
Mean accuracy (%)
Mean RT (ms)
100
95
90
85
80
75
70
65
60
55
50
1500
1400
1300
1200
1100
1000
900
800
1st
2nd
3rd
4th
5th
6th
Phonological familiarisation trial

Fig5



100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
% of total target phonemes
NEW
FAM
SEM
NEW
FAM
SEM
Day 1
Day 2
Omission
Intrusion
CAP

Fig6

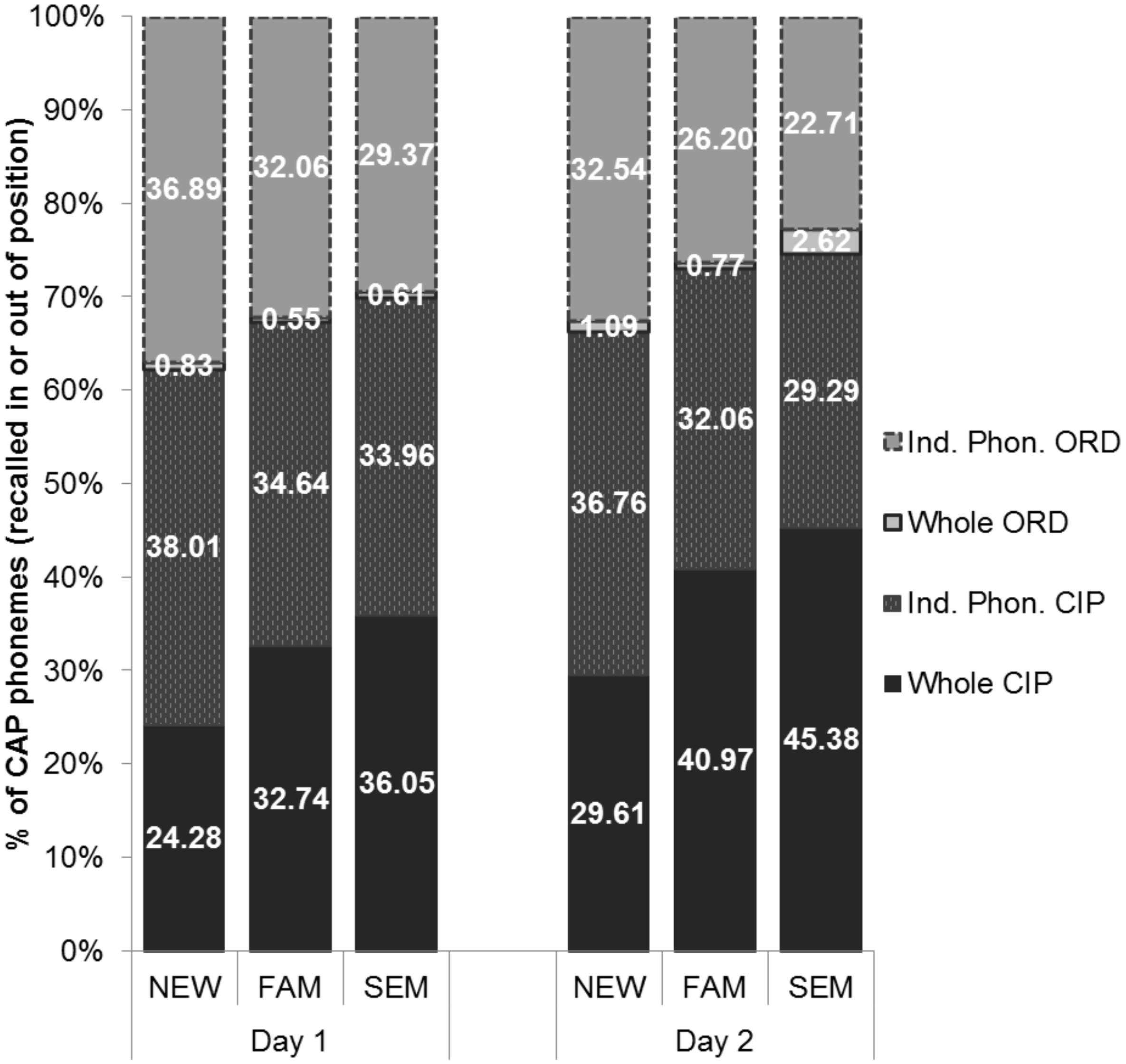
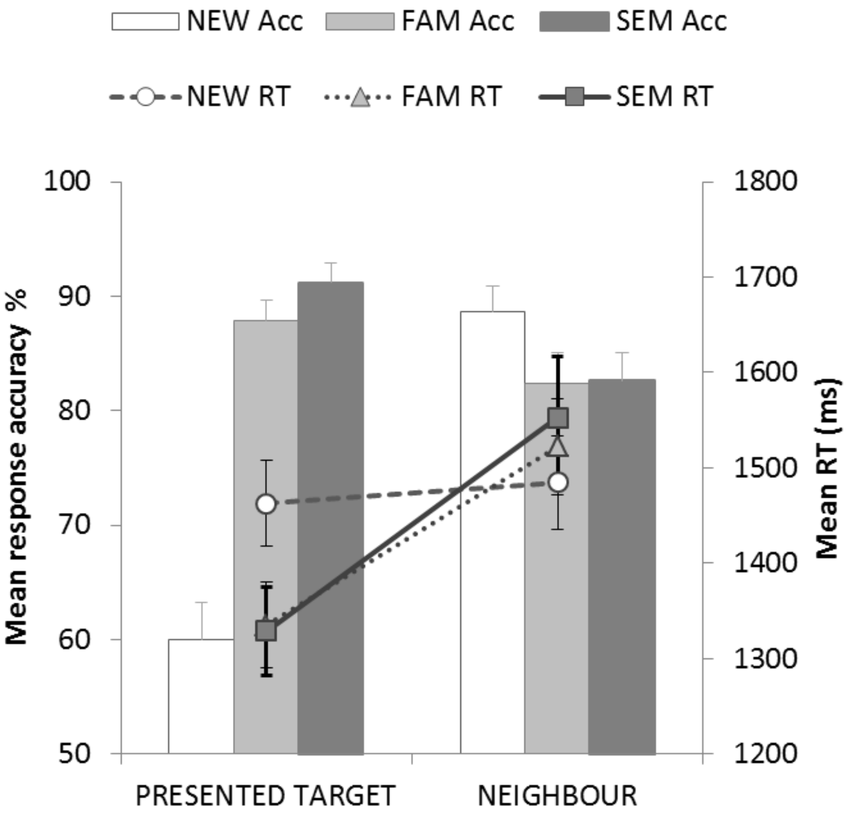

100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
% of CAP phonemes (recalled in or out of position)
36.89
0.83
38.01
24.28
32.06
0.55
34.64
32.74
29.37
0.61
33.96
36.05
32.54
1.09
36.76
29.61
26.20
0.77
32.06
40.97
22.71
2.62
29.29
45.38
NEW
FAM
SEM
Day 1
NEW
FAM
SEM
Day 2
Ind. Phon. ORD
Whole ORD
Ind. Phon. CIP
Whole CIP

Fig7


NEW Acc
FAM Acc
SEM Acc
NEW RT
FAM RT
SEM RT
Mean response accuracy %
100
90
80
70
60
50
Mean RT (ms)
1800
1700
1600
1500
1400
1300
1200
PRESENTED TARGET
NEIGHBOUR