***This article has been accepted for publication in Journal of Experimental Child Psychology, published by Elsevier***

Developmental Differences in the Timecourse of Word Learning: Greater Improvements for Children, Semantic Benefits for Adults

Running head: DEVELOPMENTAL DIFFERENCES IN THE TIMECOURSE OF WORD LEARNING

Noel Lam1, Marianna E. Hayiou-Thomas1, Matthew H. C. Mak2, and Lisa-Marie Henderson1

1Department of Psychology, University of York, York, YO10 5DD, United Kingdom

2Department of Psychology, University of Warwick, Coventry, CV4 7AL, United Kingdom

Correspondence to: Noel Lam (nl672@york.ac.uk)

For the purpose of open access, the author has applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising from this submission.

The data files and scripts for the pre-registered and exploratory analyses can be found at <https://osf.io/gruc4/>. Pre-registrations can be found at <https://osf.io/xkbnz> (Experiment 1) and <https://osf.io/w6mvp> (Experiment 2).

**Abstract**

Theories of memory consolidation claim that new word learning is bolstered by existing semantic knowledge. However, *when* semantic knowledge exerts its effects (i.e., at encoding and/or consolidation) and whether semantic benefits change across development remains unclear. Sixty-one children (Experiment 1) and 63 adults (Experiment 2) learned novel word forms paired with pictures of (i) real but rare animals akin to existing animals (Highly Linkable to existing knowledge), (ii) fictitious animals that were less clearly associated with familiar animals (Less Linkable), and (iii) “name tags” written with unfamiliar symbols (Unlinkable). Word form and meaning recall were tested immediately, 1 day and 1 week after learning. Children showed greater improvements across tests than adults despite comparable performance immediately after learning. Regardless of test sessions, semantic knowledge benefited adults’ recall of word form and meaning, with additional benefit from highly versus less linkable knowledge. Children only showed semantic benefits in meaning (and not word form) recall, with additional benefits from highly linkable knowledge. Instead, children’s word form recall was more globally associated with receptive vocabulary and nonword repetition. These results suggest that when present, the benefits of semantic information permeate across the timecourse of word learning; they also point to developmental differences in word learning mechanisms. Adults made clear use of associated semantic knowledge, whereas children showed more general associations between word learning and language abilities and greater benefit from offline consolidation. These results highlight the need for models of word learning and consolidation to incorporate developmental and individual differences.

## Introduction

Forming a robust long-term memory of a word form and its meaning so that it can be efficiently and flexibly retrieved for communication purposes is the ultimate goal of word learning. While new words can be encoded with minimal exposure (Mak et al., 2021), overwhelming evidence suggests that effective long-term memory formation, including the integration of a new word within existing lexical networks, is a gradual process that benefits from repeated presentation (Nation et al., 2007), retrieval practice (Antony et al., 2017) and sleep consolidation processes (James et al., 2017). However, word learning success is implicated by multiple factors, both intrinsic and extrinsic to the learner. For instance, word learning is influenced by prior knowledge, such as learning material that aligns closely with existing semantic information (Craik & Lockhart, 1972; McClelland, 2013). Here, we examine two critical questions regarding the role of semantics in word learning, key to advancing theory and applying this to optimising word learning in pedagogical contexts: *when* in the timecourse of word learning does existing semantic knowledge exert its effect, and *who* benefits the most from such knowledge?

### Word Learning Through the Lens of Memory Consolidation

Models of memory provide domain-general frameworks for understanding the word acquisition process across stages of encoding, consolidation, and retrieval. The Complementary Learning Systems (CLS) framework (McClelland et al., 1995) proposes dual memory systems underpinned by different neurobiological substrates, which work together to enable rapid encoding, long-term retention and integration with existing knowledge. Specifically, a hippocampal-based system supports rapid storage of information in a sparse, encapsulated manner. Information is then integrated gradually with prior knowledge in the neocortex to avoid interference with and disruption to existing knowledge. The CLS framework has been applied to language processing (Gaskell et al., 2019; Mak et al., 2023), and specifically to word learning. Novel word representations, similar to episodic memory, benefit from sleep-related consolidation during which new representations in the hippocampus are integrated with the neocortical system, thus facilitating their subsequent recall and recognition (Davis & Gaskell, 2009; Tamminen & Gaskell, 2013).

While the CLS framework was developed utilising adult data, children also showed similar sleep-associated improvements in word learning. For instance, word form recall and lexical integration improve after sleep, but not across equivalent time awake (Henderson et al., 2012). Moreover, a greater improvement in word recall across multiple days in children than in adults (James et al., 2019) suggests that children receive greater benefits from offline consolidation than adults. However, these developmental differences have mainly been demonstrated through learning new word forms in isolation (James et al., 2019; van Rijn et al., 2023) or through incidental exposure through stories (Henderson et al., 2015); it is less clear whether learning new word *meanings* follows comparable consolidation trajectories and shows similar developmental differences.

Emerging evidence suggests that larger offline consolidation benefits for children in word learning do not occur over repeat tests within the same day and only occur over several days (Olsson, 2022). One theory suggests that these effects may be specifically attributed to sleep (James et al., 2017). The proportion of slow wave sleep (SWS) decreases with age after reaching a peak in children aged 10-12 years, averaging roughly double the proportion of SWS relative to adults (Ohayon et al., 2004). Critically, slow wave oscillations and temporally aligned sleep spindles are associated with overnight improvements in word recall and lexical integration (Smith et al., 2018; Tamminen et al., 2013) in children and adults. For instance, in a polysomnography study, Smith et al (2018) showed that slow wave activity and spindle power predicted typically developing children’s post-sleep word recall, while spindle power was also predictive of the gains in lexical competition. As these neural activities represent reactivations of newly learned information during sleep, this provides evidence for the active consolidation of newly learned words and their incorporation into children’s lexicon.

There may also be alternative, not necessarily mutually exclusive, explanations for increased offline benefits in children. For instance, sleep provides a period free from interference, thus preserving memory and preventing forgetting, which may be more beneficial for children (see van Rijn et al., 2023 for further discussion). Furthermore, experimental designs with repeated testing elements provide additional consolidation benefits through retrieval, even in the absence of feedback (Antony et al., 2017; Roediger & Karpicke, 2006), and it remains plausible that children may benefit more from repeated retrieval opportunities than adults.

### Consistency with Prior Semantic Knowledge

Providing explicit semantic information alongside new word forms benefits word learning, as it enables more in-depth processing than when word forms are presented without meaning (Craik & Lockhart, 1972). For example, presenting new spoken words with picture referents and definitions led to better word recall one week after learning, compared to presenting new spoken words with only their orthographic forms in children aged 5-9 years (Henderson et al., 2013; Ouellette & Fraser, 2009). This suggests that the availability of semantic information, whether it be an object referent or a definition, can facilitate the memory of new lexical representations.

Theories of learning and memory diverge on *when* the influence of semantics might emerge in the word learning process. In an extension to the original CLS framework, McClelland (2013) proposed that consistency with schema, or existing knowledge, allows new information to be integrated into long-term storage sites more rapidly and is less reliant on hippocampal temporary storage (Tse et al., 2007). Thus, while CLS does not explicitly address the role of sleep in the timecourse of memory integration, it aligns with the prediction that learning new words associated with consistent existing semantic knowledge could show early benefits, even before sleep. Alternatively, the Information Overlap to Abstract (iOtA) model (Lewis & Durrant, 2011) proposes a sleep-based mechanism for forming new schemas and for assimilating new memories into existing schemas during offline reactivation in sleep. This perspective predicts a greater benefit of learning information consistent with prior knowledge *after* a period of sleep. Therefore, one can derive specific predictions about the timecourse of semantic benefits for word learning from these two models of memory, and the present study aims to shed light on these theories by examining the timescale of consolidation of novel words with different amounts of semantic information.

In one of the few studies to examine when consolidation may be supported by familiar semantics, Havas et al. (2018) examined word learning with familiar objects, novel objects, and without pictures. Better recognition of word forms and word-picture correspondence of familiar semantics was found after a 12-hour delay, regardless of whether it included sleep or wakefulness. However, the lack of an immediate test following learning makes it impossible to determine whether the advantages of familiar semantics were immediate, or whether they only emerged after a delayed period.

Using existing knowledge to support word learning may be partly dependent on the size and richness of the existing knowledge network, leading us to predict both developmental and individual differences in the benefit of semantics on word learning. For instance, adults - with more extensive lexical and semantic networks - may draw on top-down support from their richer networks to support word learning (Bermúdez-Margaretto et al., 2020). Conversely, children - receiving less support from a developing lexicon - may rely more on alternative mechanisms, such as memory consolidation processes (James et al., 2017). This would be consistent with the evidence discussed above that children demonstrate greater consolidation across time than adults (Wilhelm et al., 2013). Moreover, James et al (2019) showed larger gains in children's new word recall over a week than adults’, while adults showed lasting benefits from learning new words with more orthographic and phonological neighbours - compared to children who only showed such benefits at immediate test, with consolidation gains allowing no-neighbour words to catch up. However, given this neighbourhood manipulation did not specifically target semantic knowledge, it remains unclear if and when children show less support from semantic knowledge in word learning.

Individual differences in prior knowledge among children, specifically vocabulary knowledge, are associated with word acquisition. This is referred to as the “Matthew Effect” in vocabulary development, whereby “the rich get richer” as children with better vocabulary skills are better equipped to learn and retain new words, resulting in a widening vocabulary gap (Cain & Oakhill, 2011; James et al., 2017; Stanovich, 1986). Consolidation mechanisms could underlie the Matthew Effect, such that children with better vocabulary knowledge are superior at overnight consolidation, as evident by positive correlations between children’s vocabulary knowledge and overnight gains in new word recall and lexical integration (Henderson & James, 2018; James et al., 2017). Among younger children, a larger vocabulary is beneficial for learning novel labels and objects in familiar categories (Borovsky et al., 2016), suggesting that individual differences in vocabulary knowledge impact the ability to capitalise on connections with semantic knowledge to support learning and consolidation (Diakidoy, 1998). Alternatively, the Matthew Effect may manifest as a general language learning benefit, as children with better vocabulary knowledge on standardised tests perform more accurately on measures of new word learning across test sessions (James et al., 2021). This could be a reflection of variabilities in general language ability, as phonological skills such as phonological memory capacity (as measured by nonword repetition) also impact the construction of new word forms in both children and adults with and without language impairments (Bishop et al., 2012). Consistent with this, prior research suggests that vocabulary and nonword repetition predict different aspects of word learning, as vocabulary size was a predictor of semantic recall while nonword repetition predicted word recall and recognition (Adlof & Patten, 2017). Taken together, developmental differences may be present in the learning trajectory of new words between adults and children as well as within-age-group individual differences: adults may show better learning of new words that receive greater support from existing semantic knowledge, possibly earlier in the learning process, whereas consolidation across multiple days may be more beneficial to children.

### The Present Study

In two pre-registered experiments, we manipulated the extent to which new word meanings may form links with, and receive support from, existing knowledge during word learning. Using a within-subjects design, children aged 9 to 11 (Experiment 1) and adults (Experiment 2) learned new spoken word forms that were paired with picture referents from three semantic conditions (Highly Linkable, Less Linkable & Unlinkable) and their recall of word forms and meanings were assessed immediately following learning, after 1 day and after 1 week. This allowed for the examination of whether children and adults follow different learning trajectories for words receiving different levels of semantic knowledge support. The three semantic conditions manipulated the information available from picture referents presented with novel words: (i) rare animals that were selected to be similar to already-familiar animals (i.e., Highly Linkable), (ii) fictitious animals that were created to be dissimilar to known animals (i.e., Less Linkable) and (iii) “animal name tags” that comprised coloured geometric backgrounds with superimposed unfamiliar Taiwanese *zhuyin* symbols, which provide a visual cue but contain no animal-related content (i.e., Unlinkable). Crucially, these conditions were designed to manipulate the extent to which connections could be formed with existing knowledge. Similar to Havas et al (2018), the three semantic conditions allow us to contrast the absence of meaningful semantics in the Unlinkable condition with the presence of meaningful semantic information in the Less and Highly Linkable conditions. Any difference between these conditions will highlight the contribution of prior knowledge in supporting word learning. We expected the easiest connection with existing semantic knowledge in the Highly Linkable condition, harder to form existing semantic connections with the Less Linkable condition, and most difficult to form semantic connections in the Unlinkable condition.

The present study examined the effect of semantic information on the recall of new word forms and meanings. Based on prior studies (Henderson et al., 2013; James et al., 2019), we anticipated that recall would improve across test sessions, benefiting from both offline consolidation and repeated testing opportunities. It should be noted that studies adopting designs that can isolate the effects of sleep from repeated testing (Dumay & Gaskell, 2007; Havas et al., 2018; Henderson et al., 2012) have demonstrated additional sleep-consolidation benefits which cannot be explained by repeated testing effects alone. To allow room for improvements across one week, we aimed for a lower performance at the initial test, consistent with previous studies with a similar design (e.g. James et al., 2019; 2023).

To facilitate developmental comparison, we matched the initial performance in word form recall by reducing the amount of exposure to the word forms during training in adults (five exposures for adults, eight for children). This approach directly compares the effects of existing knowledge support and offline consolidation between adults and children across test sessions without the confounding effects of different levels of initial encoding; this is important because lower levels of initial encoding can lead to larger consolidation effects (Drosopoulos et al., 2007). We expected improvements in both adults and children but children would benefit more from offline consolidation processes and show greater improvements across the three test sessions compared to adults. At the same time, since adults have accumulated richer existing semantic knowledge and a mature language processing system, we predicted that this would enable them to benefit more from semantic information, particularly when it is highly consistent with existing knowledge.

Finally, we investigated individual differences in children’s word learning. While manipulating semantic information of the stimuli allows us to see if children and adults equally benefit from these local semantic associations, the incorporation of standardised measures informs whether the consolidation of word form and meaning recall is specifically associated with vocabulary knowledge (James et al., 2017) or more globally associated with language skills, such as phonological memory.

# Experiment 1

We tested the following pre-registered hypotheses with typically developing children aged 9 to 11 (Pre-registration available at: <https://osf.io/xkbnz/>):

H1: Newly learned word forms would be better recalled at the 1-day and the 1-week delayed test sessions than immediately following training, as the later test sessions can benefit from overnight consolidation and repeated testing. As there were further opportunities for offline consolidation and retrieval practice by the 1-week test, we hypothesised continued improvements in memory of new words at the 1-week delay test compared to the 1-day delay test.

H2: We predicted that items paired with semantic information (i.e. Highly Linkable and Less Linkable conditions) would be better learned than items paired with patterns (i.e. Unlinkable condition).

H3: We predicted that the better learning of the words paired with pictures from the Highly Linkable than the Less Linkable condition could either manifest at the immediate test and show reduced improvement from immediate to 1-day delay tests (following the CLS framework) or the benefit could manifest through increased improvement from immediate to 1-day delay tests (according to the iOtA model).

H4: Children with better receptive vocabulary would demonstrate greater improvements, both overnight and across one week, in the recall of word forms and form-meaning correspondence. As an exploratory analysis, we also considered language skills beyond vocabulary, using nonword repetition as a measure of phonological memory predicting word form and meaning recall.

## Method

### Participants

Our pre-registration specified the sample size as 60 children, based on an *a priori* power analysis with G\*Power (Version 3.1.9.6, Faul et al., 2007) following the interaction between semantic condition and test session (*η*p2 = .22) in Henderson et al. (2013), at a significance criterion of α = .05 and power = .90.

Sixty-one children (31 boys and 30 girls) aged 9-11 years (*M* = 10.08 years) were recruited from a mainstream primary school in North Yorkshire, United Kingdom. Fifty-nine of the participants were native monolingual English speakers, with two bilingual speakers of English with Thai and Hungarian, respectively. Both children reported using English at school and in everyday communication and their performance in standardised language tests (the British Picture Vocabulary Scale, 2nd Ed; Dunn et al., 1997) was within the normal range. Therefore, their data were retained in the analysis.

The study was approved by the Research Ethics Committee of the Department of Psychology. Consent was obtained from the school headteacher, and parents were fully informed and provided opt-out consent.

### Stimuli

Twenty-four novel words were paired with a picture from one of three categories: Highly Linkable animals, Less Linkable animals and Unlinkable patterns. All stimuli were placed on a 1080 x 1080 pixels square canvas with a plain white background (Figure 1). Animals for the Highly Linkable and Less Linkable conditions were digitally drawn in coloured using Procreate (version 5.2.6, Savage Interactive, 2022).



Figure 1. Examples of picture stimuli for each semantic condition (Left to Right: Highly Linkable, Less Linkable & Unlinkable).

#### Highly Linkable animals

The Highly Linkable stimuli were real breeds or species of animals that were deemed by the researchers to be unfamiliar to children aged 9 to 12 years. They were characterised by being easily associated with a familiar animal while having at least one unique physical feature that set it apart from that familiar animal (e.g. a *pipa* is a frog with a very flat body). Two pilot studies with 31 adults and 10 children (aged 8 to 12) were used to verify the selected animals were unfamiliar to adults and children but could be consistently related to similar familiar animals (e.g. pilot participants consistently deemed a *pipa* as similar to a frog, see Appendix A).

#### 

#### Less Linkable animals

The Less Linkable stimuli were fictitious animals consisting of at least one salient feature (e.g., big, chubby cheeks), created by combining features and textures from different animals such that they did not consistently resemble any single familiar animal. The aforementioned pilot studies verified that these animals were unfamiliar to adults and children and enabled the selection of 8 stimuli that could not be associated consistently with familiar animals.

#### Unlinkable patterns

Each picture in the Unlinkable condition consisted of a patterned background and two symbols adopted from the Taiwan *zhuyin* phonetic system. These symbols were used so that they provided a similar number of distinctive but meaningless features for participants to identify and recall. A different colour combination was used for each background and symbol combinations as global features. The symbols on each stimulus (albeit unfamiliar to the participants) did not match the sounds of the nonwords they were paired with.

#### Novel Word Stimuli

Twenty-four bi- or tri-syllabic words were selected, with 8 items (*jerboa, kochi, loaghtan, mata, okapi, pinchaque, pipa, zebu*) being the names or part of the common or scientific names of the animals selected for the Highly Linkable condition. The remaining items consisted of two lists of 8 nonwords adopted from the items in the Phonemic Decoding Efficiency subtest of the Test of Word Reading Efficiency (2nd Ed., Torgesen et al. 2012) and unusual names from the Novel Object and Unusual Name Database (Horst & Hout, 2016). Each list of the nonwords was matched in the number of syllables, bigram frequency and initial syllable, and items in each list were either paired with stimuli from the Less Linkable or Unlinkable conditions. They were recorded by a female native British English speaker.

### Procedure

All participants were tested at three time points: immediately after learning (T1), after a one-day delay (T2), and a one-week delay (T3). T3 took place 6 to 9 days (*M* = 7.8 days) after the first session, except for 3 children, who completed the final test 12 to 14 days after the first session due to absences from school. Data collection took place individually either in an empty classroom or in a quiet place in school during school hours. All experimental tasks were presented using Gorilla Experiment Builder (Anwyl-Irvine et al., 2020) on a 14”-Macintosh laptop. The words were presented via a pair of headphones and children responded either by pressing buttons on a keyboard or verbally to the attached microphone for later offline transcription and coding.

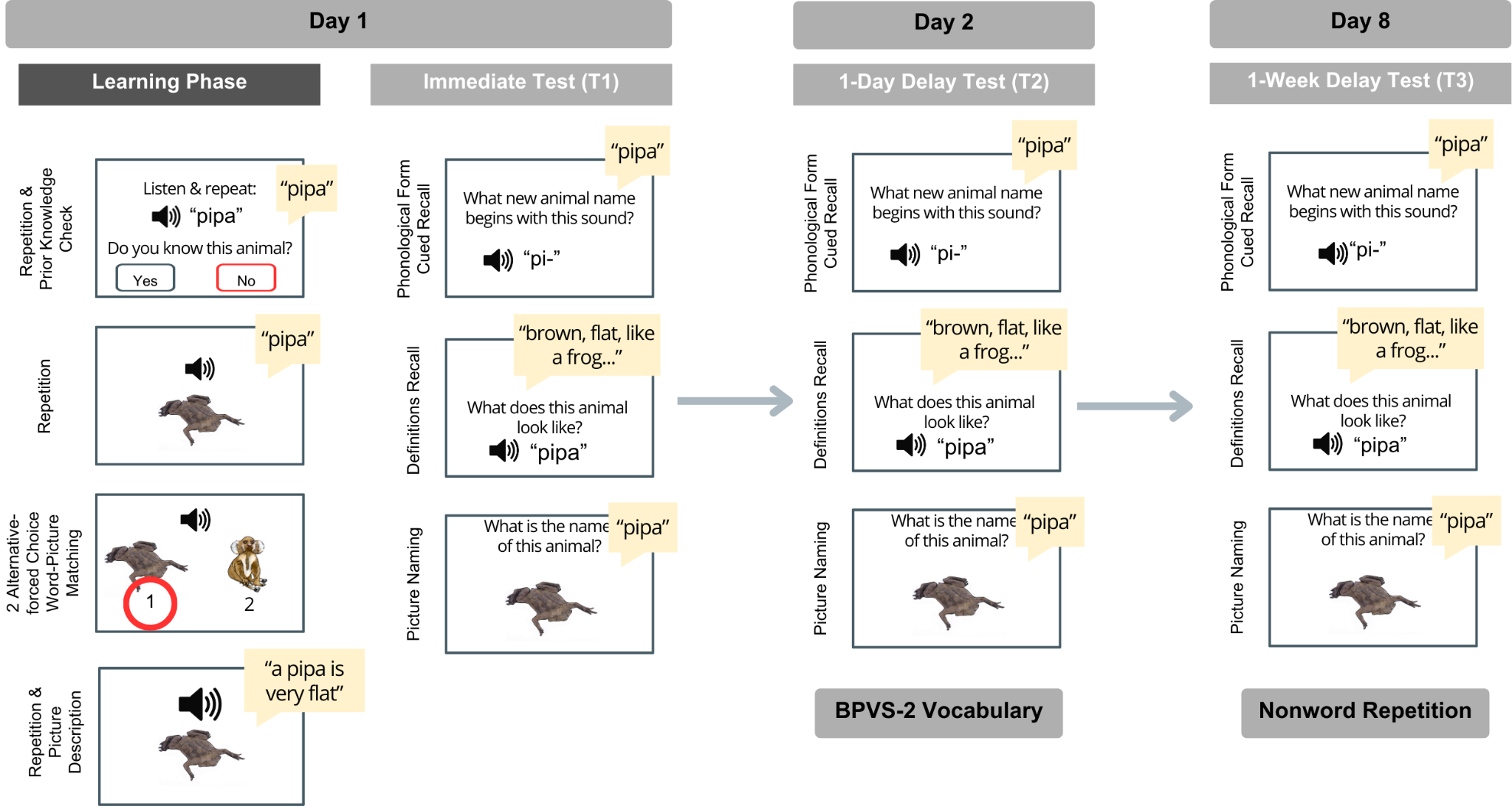


Figure 2. Overview of the procedure for Experiments 1 & 2. BPVS-2 Vocabulary and CTOPP-2 nonword repetition were only conducted with children in Experiment 1.

#### Learning Phase

Four training tasks were used to familiarise children with phonological forms and pictures (see Figure 2). In total, children heard each new word eight times, saw each picture six times and received three repetition opportunities.

##### Repetition and prior knowledge check

Children heard each new word and repeated it aloud. They were then asked to indicate whether they knew the meaning of the words and, if they did, to briefly describe the animal to demonstrate their knowledge. If their descriptions demonstrated any knowledge of the target animal (even limited e.g., alluding to a *pipa* as an amphibian), they were removed from analysis.

##### Word repetition

Children heard each new word with the corresponding picture stimuli presented on the screen. They repeated the words aloud and then heard the word again through the headphones.

##### Two-alternative forced-choice (2AFC) word picture-matching

In each trial, children heard a word and were presented on-screen with a target picture that matched the word and a distractor picture corresponding to another word belonging to one of the three semantic conditions. The numbers “1” and “2” appeared beneath each picture. Children were asked to select the picture that matched the word they heard by pressing the corresponding number key. Feedback was given through the presentation of the word with the correct picture on the screen for 3000ms. Children completed two blocks of the picture-matching task, with a different distractor picture paired with the target in each block.

##### Repetition and Picture description

Each new word was presented via headphones with the respective picture on the screen, children were asked to repeat the word aloud and to provide one physical description of the picture (e.g. “A *pipa* is very flat”). This task aimed to orient the participants towards the features embedded in the target picture.

#### Testing Phase

Three tasks were used to examine children’s knowledge of the word forms and meanings. The tasks were presented in each test session in a fixed order, with the items within each test presented in a randomised order.

##### Phonological cued recall

Children heard the first consonant(s) and vowel of the target word and were asked to complete the rest of the word. They were encouraged to give partial responses and make guesses if they were not sure of the answer. Responses were transcribed and scored based on accurate completion of the target word (0 or 1).

##### Definition recall

Children were presented with a word and were asked to describe the corresponding animal or pattern with as many details as they could recall. Responses were transcribed and a lenient scoring system was adopted (Hulme & Rodd, 2021). That is, a response received a score of “1” if at least one feature was recalled that allowed for correct identification of the target picture within the stimulus set, as confirmed by two independent scorers. Inter-rater reliability was calculated using Cohen’s kappa, *κ = .80*, suggesting strong agreement between scorers (McHugh, 2012). Ambiguous and conflicting responses were resolved on a case-by-case basis via discussion between the scorers.

##### Picture naming

A picture of an animal or pattern was presented on the screen and children were asked to recall the corresponding word. Partial responses and guesses were encouraged if they were unsure. Responses were transcribed and were scored based on whole-word accuracy (0, 1).

#### Standardised tests

The British Picture Vocabulary Scale, 2nd edition (BPVS-2, Dunn et al., 1997) and the nonword repetition subtest of the Comprehensive Test of Phonological Processing, 2nd edition (CTOPP-2; Wagner et al., 2013), were administered at T2 and T3 to measure children’s vocabulary knowledge and phonological working memory. Responses for the two standardised tests were scored and converted to standard scores as per the test manuals.

### Data Analysis

All data were analysed in RStudio (version 2021.09, RStudio Team) for R version 3.6.1 (R core team, 2019) with lme4 (Bates et al., 2015) and emmeans (Lenth, 2023) packages.

Items for which participants indicated prior knowledge during training were excluded from analysis, amounting to 30 out of 1464 total words or 2% of all data points removed from analyses. For each task in the testing phase, a model with fixed effects of test session, semantic training condition, BPVS standard score centred at pre-processing, and their interactions were fitted. As a three-level factor, two repeated contrasts were created for test sessions, 1) T1vsT2: for differences in performance immediately after learning and after a 1-day delay; 2) T2vsT3: for subsequent changes across a longer period of delay, between the 1-day delay and 1-week delay tests. For semantic training conditions, Helmert contrasts with 2 levels were set to evaluate performance differences, 1) Semantics1: between words paired with patterns in the Unlinkable condition and animals (both Highly and Less Linkable), and 2) Semantics2: between words paired with animals from the Highly and Less Linkable conditions.

After establishing the best-fitting mixed effects model, standardised dfbetas were calculated to identify influential cases using the package influence.ME (Nieuwenhuis et al., 2012). Participants with a standardised dfbeta beyond ±3.29 were then removed from the dataset and the model parameters were re-estimated with the trimmed dataset.

## Results

### Phonological cued recall



Figure 3. Average phonological cued recall accuracy for each semantic condition for children and adults. Error bars represent 95% confidence intervals and coloured dots represent participant-level performance

The accuracy for the phonological cued recall task was low overall, particularly at T1 (*M =* .09, *SD =* .28) as illustrated in Figure 3. However, children’s performance improved significantly at subsequent test sessions (T1vsT2: *β* = 1.01, *SE =* .15, *z* = 6.68, *p* <.001; T2vsT3: *β* = .86, *SE* = .11, *z* = 7.54, *p* <.001, see Table 1), with the mean proportion of cued recall accuracy being .17 (*SD =* .37) at T2 and .28 (*SD =* .45) at T3. There was no influence of semantic conditions, alone or interacting with the test session (*p*s >.05). Finally, children with higher vocabulary scores demonstrated higher cued recall accuracy (*β* = .02, *SE* = .01, *z =* 2.24, *p* = .025, see Figure 4A), suggesting a general association between existing vocabulary knowledge and cued recall accuracy for new word-forms.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *Table 1. Predictors of Children’s Phonological Cued Recall* | | | | | |
| Fixed effects | *b* | SE | *z* | *p* |  |
| (Intercept) | -2.30 | 0.25 | -9.24 | <.001 | \*\*\* |
| T1vsT2 | 1.01 | 0.15 | 6.68 | <.001 | \*\*\* |
| T2vsT3 | 0.86 | 0.11 | 7.54 | <.001 | \*\*\* |
| Semantics1 | -0.14 | 0.10 | -1.48 | 0.140 |  |
| Semantics2 | -0.09 | 0.26 | -0.33 | 0.740 |  |
| BPVSCentered | 0.02 | 0.01 | 2.24 | 0.025 | \* |
| T1vsT2:Semantics1 | 0.15 | 0.11 | 1.28 | 0.199 |  |
| T2vsT3:Semantics1 | -0.08 | 0.08 | -1.02 | 0.310 |  |
| T1vsT2:Semantics2 | 0.31 | 0.17 | 1.76 | 0.079 |  |
| T2vsT3:Semantics2 | 0.03 | 0.14 | 0.21 | 0.831 |  |
| T1vsT2:BPVSCentered | 0.02 | 0.01 | 1.74 | 0.083 |  |
| T2vsT3:BPVSCentered | 0.01 | 0.01 | 0.74 | 0.459 |  |
| Semantics1:BPVSCentered | -0.01 | 0.00 | -1.20 | 0.229 |  |
| Semantics2:BPVSCentered | -0.01 | 0.01 | -1.80 | 0.072 |  |
| T1vsT2:Semantics1:BPVSCentered | 0.01 | 0.01 | 0.67 | 0.505 |  |
| T2vsT3:Semantics1:BPVSCentered | 0.00 | 0.01 | -0.29 | 0.775 |  |
| T1vsT2:Semantics2:BPVSCentered | 0.00 | 0.02 | 0.23 | 0.816 |  |
| T2vsT3:Semantics2:BPVSCentered | 0.01 | 0.01 | 0.53 | 0.598 |  |
|  |  |  |  |  |  |
| Random effects | Variance | SD |  |  |  |
| Participant: (intercept) | 0.29 | 0.54 |  |  |  |
| Participant: Semantics1 (slope) | 0.03 | 0.16 |  |  |  |
| Participant: Semantics2 (slope) | 0.20 | 0.44 |  |  |  |
| Target: (intercept) | 1.22 | 1.11 |  |  |  |
| \* p<.05, \*\* p<.005, \*\*\* p<.001 |  |  |  |  |  |
| *Note*: model from 3738 observations collected from 53 participants across 24 items, after removing 8 influential participants | | | | | |

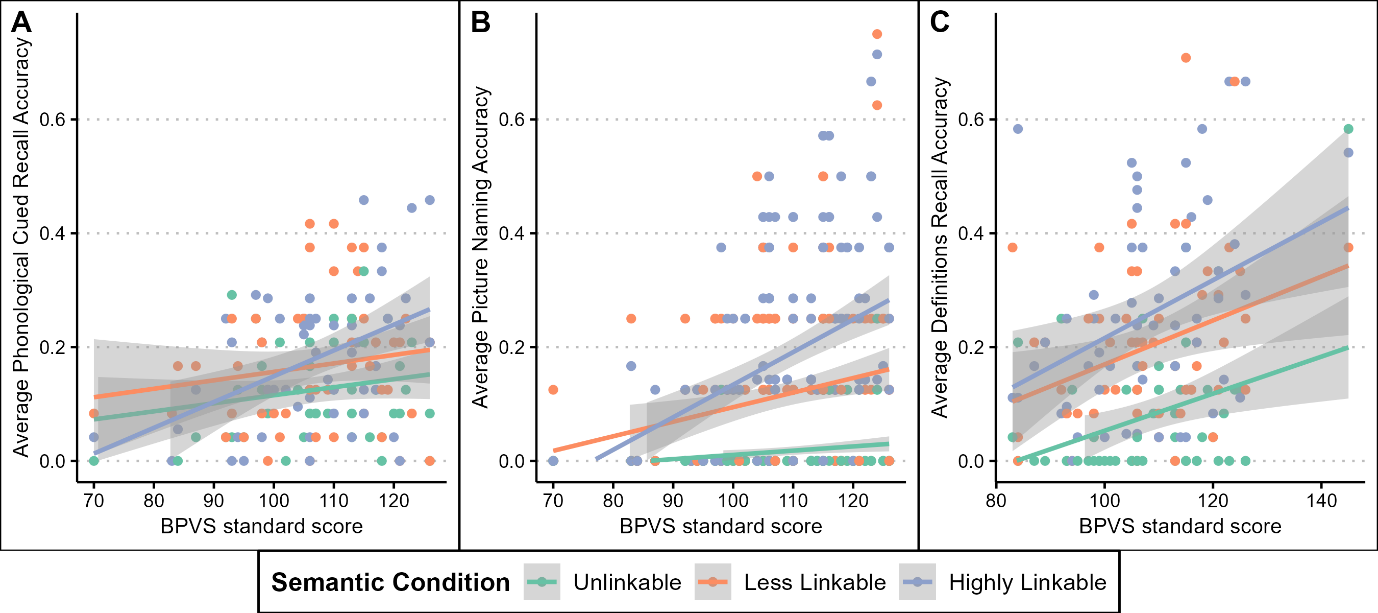


Figure 4. Children's average performance per semantic condition against the British Picture Vocabulary Scale (BPVS) standard scores in the A) Phonological cued recall, B) Picture naming, and C) Definitions recall tasks in Experiment 1. Gray bands

represent 95% confidence intervals.

### Picture naming

##### 

Figure 5. Average picture naming accuracy for each semantic condition of children and adults. Error bars represent 95% confidence intervals and coloured dots represent participant-level performance.

Similar to phonological cued recall, picture naming accuracy at T1 was low (*M =* .08; *SD =* .27, Figure 5). However, performance improved significantly at the 1-day delay test (*M =* .12; *SD =* .33; *β*= 0.53, *SE* = .18, *z* = 3.02, *p* = .003), with further improvements at the 1-week delay test (*M =* .16; *SD =* .36; *β* = .52, *SE* = .15, *z* = 3.51, *p* < .001). Again, there was no significant influence from the semantic conditions in picture naming accuracy, either alone or in interaction with test sessions (*p*s >.13, see Table 2). Receptive vocabulary knowledge was a significant predictor of picture naming accuracy (*β* = .05, *SE* = .01, *z* = 3.76, *p* < .001). As shown in Figure 4B, children with better vocabulary knowledge showed higher naming accuracy. However, the follow-up analysis suggested that there was no significant interaction between receptive vocabulary and changes across test sessions or with semantic conditions (*p*s > .18).

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *Table 2. Predictors of Children’s Picture Naming* | | | | | |
| Fixed effects: | *b* | SE | *z* | *p* |  |
| (Intercept) | -3.03 | 0.29 | -10.58 | <.001 | \*\*\* |
| T1vsT2 | 0.53 | 0.18 | 3.02 | 0.003 | \*\* |
| T2vsT3 | 0.52 | 0.15 | 3.51 | <.001 | \*\*\* |
| Semantics1 | 0.00 | 0.05 | 0.04 | 0.971 |  |
| Semantics2 | -0.07 | 0.08 | -0.87 | 0.382 |  |
| BPVSCentered | 0.05 | 0.01 | 3.76 | <.001 | \*\*\* |
| T1vsT2:Semantics1 | 0.08 | 0.13 | 0.61 | 0.543 |  |
| T2vsT3:Semantics1 | -0.03 | 0.11 | -0.24 | 0.809 |  |
| T1vsT2:Semantics2 | -0.04 | 0.21 | -0.20 | 0.839 |  |
| T2vsT3:Semantics2 | 0.28 | 0.18 | 1.53 | 0.126 |  |
| T1vsT2:BPVSCentered | 0.02 | 0.02 | 1.14 | 0.253 |  |
| T2vsT3:BPVSCentered | -0.02 | 0.01 | -1.34 | 0.181 |  |
| Semantics1:BPVSCentered | 0.00 | 0.00 | 0.98 | 0.327 |  |
| Semantics2:BPVSCentered | 0.00 | 0.01 | 0.15 | 0.881 |  |
| T1vsT2:Semantics1:BPVSCentered | -0.01 | 0.01 | -0.88 | 0.379 |  |
| T2vsT3:Semantics1:BPVSCentered | 0.01 | 0.01 | 0.50 | 0.614 |  |
| T1vsT2:Semantics2:BPVSCentered | -0.02 | 0.02 | -1.02 | 0.308 |  |
| T2vsT3:Semantics2:BPVSCentered | -0.01 | 0.02 | -0.48 | 0.628 |  |
|  |  |  |  |  |  |
| Random effects | Variance | SD |  |  |  |
| Participant: (intercept) | 0.89 | 0.94 |  |  |  |
| Target: (intercept) | 1.34 | 1.16 |  |  |  |
| \* p<.05, \*\* p<.005, \*\*\* p<.001 |  |  |  |  |  |
| *Note*: model from 3738 observations collected from 53 participants across 24 items, after removing 6 influential participants | | | | | |

### Definition recall

# Average definition recall accuracy for each semantic condition in children and adults.

Figure 6. Average definition recall accuracy for each semantic condition in children and adults. Error bars represent 95% confidence intervals and coloured dots represent participant-level performance.

Mean performance in definition recall is shown in Figure 6, summarised across the learning conditions and sessions. Children recalled an average proportion of .17 (*SD =* .37) of word meanings at the immediate test. Performance remained stable across test sessions (*MT2* = .18, *SD* = .38; *MT3* = .19, *SD* = .39) and there were no significant improvements across test sessions (*p*s > .4; see Table 3). Definition recall was significantly better for both types of animals than Unlinkable patterns (*β*= -.68, *SE* = .13, *z* = -5.18, *p* < .001). However, children showed comparable performance in recalling the meaning of words for Less Linkable (*M* = .20, *SD* = .40) and Highly Linkable animals (*M* = .27, *SD* = .44) and there was no significant interaction between test sessions and conditions (*p*s > .15). For individual differences, children with better receptive vocabulary again demonstrated better definition recall overall (*β* = .04, *SE* = .01, *z* = 4.03, *p* < .001, Figure 4C).

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *Table 3. Predictors of Children’s Definition Recall* | | | | | |
| Fixed effects: | b | SE | z | p |  |
| (Intercept) | -2.22 | 0.21 | -10.78 | <.001 | \*\*\* |
| T1vsT2 | 0.01 | 0.13 | 0.11 | 0.912 |  |
| T2vsT3 | 0.09 | 0.13 | 0.70 | 0.485 |  |
| Semantics1 | -0.68 | 0.13 | -5.18 | <.001 | \*\*\* |
| Semantics2 | -0.21 | 0.19 | -1.10 | 0.272 |  |
| BPVSCentered | 0.04 | 0.01 | 4.03 | <.001 | \*\*\* |
| T1vsT2:Semantics1 | -0.15 | 0.10 | -1.44 | 0.151 |  |
| T2vsT3:Semantics1 | 0.00 | 0.11 | -0.02 | 0.983 |  |
| T1vsT2:Semantics2 | 0.16 | 0.13 | 1.28 | 0.202 |  |
| T2vsT3:Semantics2 | -0.09 | 0.13 | -0.71 | 0.480 |  |
| T1vsT2:BPVSCentered | 0.01 | 0.01 | 0.91 | 0.363 |  |
| T2vsT3:BPVSCentered | 0.00 | 0.01 | 0.36 | 0.719 |  |
| Semantics1:BPVSCentered | 0.01 | 0.01 | 0.90 | 0.371 |  |
| Semantics2:BPVSCentered | 0.00 | 0.01 | -0.42 | 0.677 |  |
| T1vsT2:Semantics1:BPVSCentered | 0.00 | 0.01 | 0.56 | 0.578 |  |
| T2vsT3:Semantics1:BPVSCentered | 0.00 | 0.01 | -0.34 | 0.738 |  |
| T1vsT2:Semantics2:BPVSCentered | -0.01 | 0.01 | -0.92 | 0.360 |  |
| T2vsT3:Semantics2:BPVSCentered | 0.00 | 0.01 | 0.22 | 0.825 |  |
|  |  |  |  |  |  |
| Random effects | Variance | SD |  |  |  |
| Participant: (intercept) | 0.46 | 0.68 |  |  |  |
| Participant: Semantics1 (slope) | 0.40 | 0.63 |  |  |  |
| Participant: Semantics2 (slope) | 0.20 | 0.44 |  |  |  |
| Target: (intercept) | 0.62 | 0.79 |  |  |  |
| \* p<.05, \*\* p<.005, \*\*\* p<.001 |  |  |  |  |  |
| *Note*: model from 4086 observations collected from 58 participants across 24 items, after removing 3 influential participants. | | | | | |

## Discussion

Experiment 1 examined children’s learning of new spoken word forms and meanings across three test sessions. Consistent with Hypothesis 1, children showed improvements across repeated tests over the week in recalling word forms when given the first sounds or corresponding pictures as cues. Conversely, meaning recall remained stable across test sessions. This divergent pattern of improvement in word form memory but maintenance of word meaning memory over repeated tests is consistent with previous studies manipulating other dimensions of semantic knowledge. For instance, studies examining the influence of the number of semantic neighbours also showed no consolidation or repeated testing benefits in word meaning recall (James et al., 2023; Tamminen et al., 2013).

The present experiment considered the role of semantic knowledge by examining the impact of individual differences in vocabulary size and manipulating the ease of new word meaning forming links with existing knowledge on new word learning. Across three tasks, we found that existing receptive vocabulary knowledge had a general effect on word learning, such that children with better existing vocabulary recalled more word forms and meanings overall. Contrary to Hypothesis 2, children’s performance on the word form recall and picture naming tasks did not benefit from the inclusion of semantic information during learning, regardless of prior knowledge support. One potential explanation for the lack of a difference in learning between Less and Highly Linkable words may be that children lacked the relevant semantic knowledge to support new learning. However, this is unlikely as the Highly Linkable animals were specifically selected to resemble animals that children are familiar with from pilot studies. Furthermore, the definition recall performance showed that children’s recall of word meaning *was* more accurate with both Less and Highly Linkable animals, compared to Unlinkable patterns (but with no difference between these two conditions). Thus, rather than not having the existing knowledge to form links with the new words, children’s semantic knowledge may not be rich enough to effectively support the learning and/or retrieval of new phonological forms (Mak & Twitchell, 2020). To further explore this possibility, we turn to adults, who have a richer and more complex semantic network and examine the hypothesis that when existing semantic networks are more developed, semantics can make a greater contribution to word learning in Experiment 2.

# Experiment 2

Experiment 2 examined semantic benefits in the timecourse of adults’ word learning, testing the following pre-registered hypotheses (<https://osf.io/w6mvp/>):

1. Consistent with Experiment 1, memory for new word forms should improve across test sessions.
2. If a general benefit for semantic conditions was found, then items paired with semantic information (i.e. Highly and Less Linkable animals) would be better recalled than items paired with no semantic information (Unlinkable patterns).
3. There would be better learning of the words paired with semantic information that can be easily linked with existing knowledge (Highly Linkable) than Less Linkable animals. This could manifest either at the immediate test and show less relative improvement from T1 to T2 (following the CLS framework, McClelland, 2013) or, in line with the iOtA model (Lewis & Durrant 2011), which posits that sleep consolidation mechanisms would have greater support the consolidation for schema-consistent items, then we expect an increase in performance from T1 to T2.

## Method

### Participants

We aimed to recruit a comparable sample size as in Experiment 1 (n = 61). The final sample comprised 63 participants (60 females), aged 18 to 23 (M = 19.61 years). They were recruited from the undergraduate and postgraduate student population. All participants were proficient in English with normal or corrected-to-normal vision and hearing. Participants were asked to complete each test session at a similar time on each testing day, but we retained data from all sessions completed on the correct day. Seventy-five participants were recruited, but twelve participants were excluded from the analysis: seven because they completed the 1-day delay test of the experiment beyond one calendar day after the immediate test, four due to technical errors with audio recording and one because of self-reported history of a developmental disorder.

### Design & Procedure

The stimuli and tasks from Experiment 1 were used with some procedural changes: (i) The number of exposures for each novel word was reduced to five from eight exposures in Experiment 1, enabling adults’ initial performance on the word form recall task to match children’s T1 performance. The number of exposures was determined based on a pilot study conducted with 10 adults. (ii) As opposed to school-based face-to-face testing in Experiment 1, the experiment was delivered online via Gorilla Experiment Builder (Anwyl-Irvine et al., 2020). Participants were instructed to complete the experiment individually in a quiet room using a pair of headphones connected to their computer.

#### Learning Phase

The same training tasks were used as in Experiment 1. Participants were instructed to only repeat each word once during one round of repetition training and only one round of picture-word matching was administered.

#### Testing Phase

In addition to the same word testing tasks used in Experiment 1, adults completed the Stanford Sleepiness Scale (Hoddes et al., 1972) at the start of each time point to provide a measure of alertness. They also completed the Pittsburgh Sleep Quality Index (Buysse et al., 1989) at the end of the 1-week delay test for the identification of particularly poor sleep quality. Participants’ PSQI total score ranged from 2 to 12 (*M* = 6.75) and this was not correlated with the performance in any of the testing tasks (all *r*s < .08).

### Data analysis

As in Experiment 1, all data were analysed in RStudio (version 2021.09) for R (version 3.6.1, (RStudio Team, 2020), using the same contrasts for test sessions and semantic conditions. Items for which participants reported prior knowledge were excluded from analyses, resulting in the removal of 6 out of 1536 words, or 0.3% of the data points.

## Results

### Phonological cued recall

Phonological cued recall accuracy was low overall and similar to children’s performance in Experiment 1 (Figure 3). Adults recalled a mean proportion of .11 (*SD* = .31) novel word forms at T1, with comparable performance at T2 at .12 (*SD* = .32), which did not significantly differ from each other (*p* = .25, Table 4). Performance improved significantly at the 1-week delay test with a mean cued recall accuracy of .16 (*SD* = .37, *β* = .50, *SE* = .13, *z* = 4.00, *p* < .001).

Adults recalled word forms paired with pictures of Less and Highly Linkable animals more accurately than word forms paired with Unlinkable Patterns (*β* = -.36, *SE* = .11, *z* = -3.15, *p* = .002). A benefit for Highly Linkable semantic information over Less Linkable semantic information in recalling word form was also found (*β* = -.60, *SE* = .27, *z* = -2.23, *p* = .025) but there was no interaction between test session and semantic conditions (*p*s > .25).

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *Table 4. Predictors of Adults’ Phonological Cued Recall* | | | | | |
| Fixed effects | *b* | SE | *z* | *p* |  |
| (Intercept) | -2.85 | 0.27 | -10.50 | <.001 | \*\*\* |
| T1vsT2 | 0.16 | 0.14 | 1.16 | 0.248 |  |
| T2vsT3 | 0.50 | 0.13 | 4.00 | <.001 | \*\*\* |
| Semantics1 | -0.36 | 0.11 | -3.15 | 0.002 | \*\* |
| Semantics2 | -0.60 | 0.27 | -2.23 | 0.025 | \* |
| T1vsT2:Semantics1 | -0.06 | 0.10 | -0.56 | 0.575 |  |
| T2vsT3:Semantics1 | 0.11 | 0.10 | 1.14 | 0.254 |  |
| T1vsT2:Semantics2 | -0.02 | 0.15 | -0.15 | 0.883 |  |
| T2vsT3:Semantics2 | 0.06 | 0.14 | 0.39 | 0.696 |  |
|  |  |  |  |  |  |
| Random effects | Variance | SD |  |  |  |
| Participant: (intercept) | 0.58 | 0.76 |  |  |  |
| Participant: Semantics1 (slope) | 0.16 | 0.41 |  |  |  |
| Participant: Semantics2 (slope) | 0.21 | 0.45 |  |  |  |
| Target: (intercept) | 1.34 | 1.16 |  |  |  |
| \* p<.05, \*\* p<.005, \*\*\* p<.001 |  |  |  |  |  |
| *Note*: model from 4374 observations collected from 61 participants across 24 items, after removing 2 influential participants | | | | | |

### Picture naming

Performance was again low, with adults recalling a mean proportion of .08 (*SD =* .27) of the words correctly at the immediate test (Figure 5). Picture naming accuracy was comparable at the 1-day delay test (*M* = .09, *SD* = .29; *p* = .41, see Table 5) but improved significantly at the 1-week delay test (*M =* .11, *SD* = .31; *β* = .69, *SE* = .30, *z* = 2.27, *p* = .023).

A significant main effect of semantics was found: adults were more accurate at naming pictures of Less and Highly Linkable animals than naming pictures of Unlinkable patterns (*β* = -1.60, *SE* = .38, *z* = -4.17, *p* < .001). Higher naming accuracy was also found for Highly than Less Linkable animals (*β* = -.78, *SE* = .33, *z* = -2.25, *p* = .019). Similar to the results of the phonological cued recall, there was no significant interaction between test sessions and semantic conditions in influencing picture naming performance.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *Table 5. Predictors of Adults’ Picture Naming* | | | | | |
| Fixed effects | *b* | SE | *z* | *p* |  |
| (Intercept) | -4.64 | 0.52 | -8.84 | <.001 | \*\*\* |
| T1vsT2 | 0.29 | 0.35 | 0.83 | 0.406 |  |
| T2vsT3 | 0.69 | 0.30 | 2.27 | 0.023 | \* |
| Semantics1 | -1.60 | 0.38 | -4.17 | <.001 | \*\*\* |
| Semantics2 | -0.78 | 0.33 | -2.35 | 0.019 | \* |
| T1vsT2:Semantics1 | -0.15 | 0.34 | -0.44 | 0.662 |  |
| T2vsT3:Semantics1 | 0.40 | 0.29 | 1.40 | 0.163 |  |
| T1vsT2:Semantics2 | 0.29 | 0.19 | 1.50 | 0.134 |  |
| T2vsT3:Semantics2 | 0.12 | 0.17 | 0.70 | 0.485 |  |
|  |  |  |  |  |  |
| Random effects | Variance | SD |  |  |  |
| Participant: (intercept) | 2.38 | 1.54 |  |  |  |
| Participant: Semantics1 (slope) | 0.18 | 0.42 |  |  |  |
| Participant: Semantics2 (slope) | 0.36 | 0.60 |  |  |  |
| Target: (intercept) | 1.83 | 1.35 |  |  |  |
| \* p<.05, \*\* p<.005, \*\*\* p<.001 |  |  |  |  |  |
| *Note*: model from 4089 observations collected from 57 participants across 24 items, after removing 5 influential participants | | | | | |

### Definition recall

As shown in Figure 6, adults recalled an average of .15 (*SD* = .35) of the word meanings at the immediate test and their performance was maintained at the 1-day delay test (*M* = .15, *SD* = .36; *p* = .76; see Table 6). Furthermore, there were no significant changes at the 1-week delay test (*M* = .14, *SD* = .35; *p* = .81). Semantic condition predicted word meaning recall with the meaning recall accuracy for words paired with pictures of Less & Highly Linkable animals being higher than the recall of Unlinkable patterns (*β* = -2.95, *SE* = .76, *z* = -3.88, *p* < .001). Furthermore, Highly Linkable animals were recalled more accurately than Less Linkable (*β* = -.56, *SE* = .23, *z* = -2.42, *p* = .016). Semantics did not interact with either test session contrast, suggesting that a general benefit from prior knowledge in recalling word meaning was available across all test sessions.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *Table 6. Predictors of Adults’ Definition Recall* | | | | | |
| Fixed effects | b | SE | z | p |  |
| (Intercept) | -4.80 | 0.80 | -5.97 | <.001 | \*\*\* |
| T1vsT2 | -0.08 | 0.26 | -0.31 | 0.756 |  |
| T2vsT3 | -0.07 | 0.28 | -0.24 | 0.813 |  |
| Semantics1 | -2.95 | 0.76 | -3.88 | <.001 | \*\*\* |
| Semantics2 | -0.56 | 0.23 | -2.42 | 0.016 | \* |
| T1vsT2:Semantics1 | -0.22 | 0.25 | -0.89 | 0.375 |  |
| T2vsT3:Semantics1 | 0.03 | 0.27 | 0.12 | 0.902 |  |
| T1vsT2:Semantics2 | 0.18 | 0.14 | 1.33 | 0.184 |  |
| T2vsT3:Semantics2 | -0.01 | 0.14 | -0.09 | 0.930 |  |
|  |  |  |  |  |  |
| Random effects | Variance | SD |  |  |  |
| Participant: (intercept) | 4.56 | 2.14 |  |  |  |
| Participant: Semantics1 (slope) | 2.80 | 1.67 |  |  |  |
| Participant: Semantics2 (slope) | 0.26 | 0.51 |  |  |  |
| Target: (intercept) | 0.91 | 0.95 |  |  |  |
| \* p<.05, \*\* p<.005, \*\*\* p<.001 |  |  |  |  |  |
| Note: model from 4085 observations collected from 57 participants across 24 items, after removing 5 influential participants | | | | | |

# 

## Discussion

Experiment 2 examined the influence of semantics on word learning in adults. Inconsistent with Hypothesis 1, which predicted improvements in performance across each test session, adults demonstrated maintenance of performance between immediate and 1-day delay tests. Improvements in accuracy were observed only after a longer period of consolidation at the 1-week delay test for the tasks requiring new word form retrieval (phonological cued recall and picture naming). Potentially, adults profited from the additional test retrieval opportunity after a 1-day delay to stabilise their memory traces for further consolidation over sleep. Nonetheless, a maintenance effect was found for the definitions task across the immediate to the 1-day delay tests and also after a 1-week delay at T3.

Consistent with Hypothesis 2, semantic benefits were demonstrated in adults’ learning of novel word forms, such that both phonological cued recall and picture naming accuracy was higher in the two semantic conditions than in the Unlinkable condition. Similarly, adults’ recall of word meaning was more accurate for words in the two semantic conditions (compared to the Unlinkable). For Hypothesis 3, we found a benefit of recalling word forms that were consistent with existing knowledge, such that there was a general benefit for highly linkable over less linkable semantic conditions. However, contrary to the prediction based on the iOtA framework, this benefit was observed at the immediate test.

# Child-Adult Comparison

Experiments 1 and 2 revealed different patterns of results in children and adults despite comparable levels of initial performance. While children showed greater improvements across test sessions, both after a 1-day delay and across the week, greater semantic benefits were found in adults. Children’s greater offline consolidation capacity has been demonstrated in previous form recall tasks following new word learning (James et al., 2019; 2023). Specifically, while adults retain greater prior knowledge-dependence across test sessions; children have shown initial prior knowledge benefits that are then superseded by offline consolidation for words with less prior knowledge support (James et al., 2019) or they have not benefited from semantics in word form learning (James et al., 2023). Following these previous studies, we tested two pre-registered hypotheses for cross-experiment comparisons in the word form recall task[[1]](#footnote-1) between adults and children: (i) Children would demonstrate greater improvements from day 1 to day 2 than adults in their phonological cued recall performance. (ii) Adults would receive greater benefit from semantic information, such that their recall of words paired with more semantic information would be better than that of children.

## Cross-Experiment Analysis

The same test session and semantic condition contrasts were set as in previous analyses, with an additional factor of Group, which compared children in Experiment 1 and adults in Experiment 2. Items excluded in previous analyses due to prior knowledge were also removed from current analysis, which accounted for 1.2% of all data points. The full model is presented in Table 7.

The main focus of this analysis was developmental differences between adults and children. Overall, children performed better than adults (*β =* -.24, *SE =*.07, *z =* 3.27*, p =* .001). This is due to children showing greater improvements than adults at the 1-day delay test (*β =* -.42, *SE =*.10, *z =* -4.40*, p <* .001), while adults’ performance showed improvement after the 1-week delay (*β =* -.16*, SE =*.08*, z =* -1.95*, p =* .051)*.* Post-hoc comparisons showed that, as per our manipulation of reduced exposure, adults and children had comparable performance at the immediate test *(p =* .36), but children showed greater improvements across consolidation opportunities than adults (T1 vs T2: *β =* -.66, *SE =*.18, *z* =-3.70*, p <* .001;T2 vs T3: *β =* -.97*, SE =*.17*,* *z* = -5.87*, p <* .001).

Furthermore, there was a significant interaction between group and Semantics2 (Less vs Highly Linkable: *β =* -.13, *SE =*.06*,* *z* = -2.19*, p =* .028). Post-hoc analysis suggests children outperformed adults in both the Unlinkable *(β =* -.56*, SE =*.22*,* *z* =-2.51*, p =* .012)and the Less Linkable conditions (*β =* -.71*, SE =*.21*,* *z* =-3.45*, p <* .001), but the groups showed comparable performance in the Highly Linkable condition (*p=* .33).

To summarise, the cross-experiment comparison showed a greater improvement across test sessions in children than in adults despite matching their initial levels of learning, consistent with existing literature showing enhanced consolidation in children (e.g. James et al., 2019; Wilhelm et al., 2013). While the two age groups showed comparable performance in the Highly Linkable condition, children’s memory for words with Less Linkable semantics surpassed adults’.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Table 7. Predictors of Developmental Differences in Phonological Cued Recall | | | | | |
| **Fixed effects** | *b* | SE | *z* | *p* |  |
| (Intercept) | -2.42 | 0.22 | -10.81 | <.001 | \*\*\* |
| T1vsT2 | 0.57 | 0.10 | 6.03 | <.001 | \*\*\* |
| T2vsT3 | 0.63 | 0.08 | 7.82 | <.001 | \*\*\* |
| Semantics1 | -0.24 | 0.08 | -2.83 | 0.005 | \*\* |
| Semantics2 | -0.35 | 0.23 | -1.50 | 0.133 |  |
| Group | -0.24 | 0.07 | -3.27 | 0.001 | \*\* |
| T1vsT2:Semantics1 | 0.02 | 0.07 | 0.33 | 0.743 |  |
| T2vsT3:Semantics1 | 0.03 | 0.06 | 0.55 | 0.581 |  |
| T1vsT2:Semantics2 | 0.13 | 0.11 | 1.22 | 0.221 |  |
| T2vsT3:Semantics2 | 0.08 | 0.09 | 0.88 | 0.379 |  |
| T1vsT2:Group | -0.42 | 0.10 | -4.40 | <.001 | \*\*\* |
| T2vsT3:Group | -0.16 | 0.08 | -1.95 | 0.051 | . |
| Semantics1:Group | -0.02 | 0.04 | -0.48 | 0.628 |  |
| Semantics2:Group | -0.13 | 0.06 | -2.19 | 0.028 | \* |
| T1vsT2:Semantics1:Group | -0.08 | 0.07 | -1.15 | 0.252 |  |
| T2vsT3:Semantics1:Group | 0.06 | 0.06 | 0.97 | 0.332 |  |
| T1vsT2:Semantics2:Group | -0.16 | 0.11 | -1.51 | 0.132 |  |
| T2vsT3:Semantics2:Group | -0.04 | 0.09 | -0.45 | 0.651 |  |
|  |  |  |  |  |  |
| Random effects | Variance | SD |  |  |  |
| Participant: (intercept) | 0.48 | 0.69 |  |  |  |
| Participant: Semantics1 (slope) | 0.07 | 0.26 |  |  |  |
| Participant: Semantics2 (slope) | 0.20 | 0.45 |  |  |  |
| Target: (intercept) | 0.48 | 0.69 |  |  |  |
| \* p<.05, \*\* p<.005, \*\*\* p<.001 |  |  |  |  |  |
| Note: model from 8536 observations collected from 120 participants across 24 items, after removing 4 influential participants | | | | | |

# 

# General Discussion

This study showed that adults and children benefit from different learning mechanisms during word learning: While both adults and children showed offline gains in newly learned word forms across one week, children’s improvement exceeded that of adults after both a 1-day and a 1-week delay. Conversely, adults’ recall of word form and meaning benefited from the inclusion of semantic information during learning, especially when direct links with prior knowledge could be formed. Despite the general benefit of learning new words with semantic information compared to no semantics in meaning recall, children did not benefit more strongly from semantic associations when recalling phonological forms. Instead, children with better receptive vocabulary showed overall form-learning benefits. Contrary to theoretical predictions, words paired with varying levels of semantic links received comparable offline support across test sessions, suggesting that semantic and consolidation processes work as largely separate mechanisms to support word learning, and that both are subject to developmental differences.

## Offline Gains of New Words Across One Week

Children’s word form recall improved continuously from immediate to the first delayed test after a 1-day delay, with further improvement at the one-week test, consistent with previous experiments (e.g. Henderson et al., 2013). Adults showed smaller improvements, as their word form recall only showed improvements at the one-week test. The behavioural improvements in both adults and children are consistent with the CLS framework and previous sleep research which suggests that offline hippocampal reactivation opportunities during sleep between test sessions contribute to the strengthening of word form memory (Davis & Gaskell, 2009; James et al., 2017). However, these improvements are also likely to be partly due to repeated testing. Notwithstanding, previous studies showed that significant improvements in word form recall are only found after repeated testing if sleep occurs between tests, but not over an equivalent length of daytime wakefulness (Dumay & Gaskell, 2007; Henderson et al., 2012). Nevertheless, the design of the present experiment did not allow us to dissociate the relative contribution from repeated testing and offline reactivation during sleep in enhancing the representations of new words.

When directly examining developmental differences in the changes across test sessions, we found that children’s offline gains in word form recall were more robust than adults’. Developmental difference was found even when both groups showed improvements at the end of the 1-week delay, consistent with previous research adopting a similar testing schedule (James et al., 2021, 2023). The differences in memory improvement following delays encompassing sleep consolidation opportunities are consistent with developmental changes in associated sleep parameters (e.g., slow wave oscillations) which contribute to active systems consolidation (Ohayon et al., 2004; van Rijn et al., 2023). As these SWS parameters are relevant to the transfer of new information from hippocampal to long-term neocortical storage, these developmental changes have been proposed to underlie the differences found in behavioural measures (James et al., 2017). Therefore, greater improvements in word form recall seen in children could be supported by more robust cortical consolidation taking place during interleaving sleep. Aligning with this, a recent study has shown that whilst adults can show steeper gains in new word form recall over repeated tests within a single day, children show steeper gains over sleep (Olsson, 2022). Thus, we deem repeated testing to be a less likely explanation of our findings. Furthermore, another recent study also showed greater memory benefits following a nap in children than adults, but the size of this benefit was not correlated with the magnitude of sleep parameters (van Rijn et al., 2023). Potentially then, children may benefit more than adults from the protection against interference that sleep offers. Further research is needed to fully understand the mechanisms that account for greater consolidation benefits in children and to examine if the size of the developmental differences in consolidation will hold at different levels of learning. Nevertheless, the continued improvement of word-form memory across the week following learning in both age groups also adds to a growing literature showing that word learning is a gradual process (Gordon et al., 2022), with potential roles for sleep-related consolidation processes and repeated testing opportunities that differ across development.

Contrary to hypothesis and word form recall task results, definition recall performance showed a maintenance effect rather than improvement across test sessions. Similar results were demonstrated in a recent study with adults and children, whose memory for new word forms improved after a 1-day delay, but memory for new word meanings remained stable across test sessions (James et al., 2023). Therefore, performance gains across multiple retrieval opportunities are more robust for word forms than for word meanings, perhaps due to different mechanisms underlying the encoding and storage of lexical forms and semantic knowledge. Presently, it is unclear whether lexical and semantic knowledge undergoes the same processes of consolidation and incorporation into long-term storage systems, and the present results underscore the need to address this.

## Semantic Influences in Word Learning

Children’s and adults’ word learning differed in terms of the role of semantic influence. Adults showed word learning benefits in the presence of semantic information over non-semantic patterns, with further benefits for semantic information that is highly linkable to prior knowledge. This is consistent with prior research showing that the presence of familiar semantic information benefits word learning (Havas et al., 2018).

On the contrary, no strong evidence for semantic influence was found in children’s learning of word forms. Past research investigating children’s use of semantics in word learning and retention yields mixed findings. For instance, children who learned new words with meaning showed better form recall than children who only learned the word forms after a 1-week delay, but not in earlier test sessions (Henderson et al., 2013). Yet, children also show initial lexical benefits from orthographic neighbours, which could reflect broader lexical, phonological and semantic prior knowledge (James et al., 2019). However, when learning words with different numbers of overlapping connections in existing semantic networks, children did not show word form benefits for either dense or sparse semantic networks (James et al., 2023). As the current study compared the ease of linking new and existing semantic information, our design was more analogous to that of James et al. (2023), whose results are consistent with the current findings. Potentially, children may be less likely to use semantics to support word form acquisition, relative to their capacity to draw on other forms of prior knowledge such as orthography and phonology. The present results also extend the scope of these prior findings by manipulating the availability of prior knowledge, from access to the density of existing semantic network connections (James et al., 2023) or variations within the lexicon (James et al., 2019) to the addition of new semantic knowledge with different amounts of prior semantic support.

Contrary to our initial predictions, developmental comparisons suggest that while adults and children showed comparable recall of word forms that are consistent with prior semantic knowledge, children have a better memory than adults for words that have less semantic support. Adults showed a higher dependence on prior knowledge specifically relevant to the items and a smaller role in learning with semantic information in general compared to no semantic support. Semantic benefits from existing knowledge were observed immediately after learning and maintained across a week, consistent with the CLS framework (McClelland, 2013). This contrasts with children’s word form recall, which showed comparable performance across the three semantic conditions, perhaps indicating less reliance on semantic support from existing knowledge and greater reliance on declarative memory consolidation processes. When accounting for the effects of age on children, a further exploratory analysis revealed a similar trend of increasing reliance on semantics (Appendix Table D1). While children across 9 to 11 years of age showed comparable recall for word forms in the Unlinkable semantics conditions, older children showed a greater general semantic benefit than younger children, consistent with the semantic benefits found in adults. This raises a further question of *when* in the language development trajectory learners transition from relying more heavily on memory consolidation processes to supporting their learning with prior knowledge, and whether educational practices can leverage this shift in learning mechanisms when designing learning materials and teaching strategies. Future studies with a broader age range of children, potentially including adolescents as well, can provide a clearer picture of the developmental trajectory, considering the interaction of age, changes in consolidation systems and variabilities in vocabulary and broader language skills. Furthermore, as children benefit from semantics in recalling word meanings, exploring strategies that can allow this benefit to transfer to word form learning presents an important avenue for research and teaching practices.

## The Timecourse of Semantic Influence in Word Learning

This study aimed to examine *when* semantics exerts influence on word learning, with two specific potential theoretical predictions: prior knowledge support emerges immediately after learning, without needing an extended period of consolidation (McClelland, 2013) or prior knowledge benefit emerges through reactivation processes during sleep (Lewis & Durrant, 2011). Adult results are consistent with predictions of the CLS framework, as new word forms paired with semantic information were recalled more accurately immediately after learning and this benefit was maintained across test sessions. Combined with the absence of interaction with changes across test sessions, we showed that adults use prior knowledge to support new word form acquisition without the need for a period of prolonged consolidation.

The pattern from the children's results is less clear. There were no strong significant effects of semantic conditions, either alone or in combination with changes across test sessions in the pre-registered analysis with receptive vocabulary as a predictor of performance. However, the exploratory analyses with nonword repetition and age show a greater overnight improvement for words in the Less Linkable condition than the Highly Linkable condition (Appendix Tables B1; D1). This is consistent with the CLS account with more rapid acquisition of prior knowledge supporting the rapid acquisition of novel words paired with Highly Linkable animals, while words paired with Less Linkable animals rely more on the gradual transferral of new memories from the hippocampus to the neocortical long-term storage and benefit from sleep. As this is a small effect found in exploratory analyses, and close to inference level in the analysis with receptive vocabulary, our data is unable to provide definitive evidence to thoroughly evaluate this finding concerning the predictions from the CLS framework and the iOtA model. Future studies will benefit from further investigating which aspects of new words are more or less likely to receive support from existing knowledge.

## Individual differences in children’s language abilities in supporting new learning

Theories and models of word encoding and consolidation emphasise the consistency with existing knowledge or schema as supporting the effective incorporation of new information (McClelland, 2013). Logically, the breadth of existing knowledge in a developing lexicon is a key factor in influencing the success of forming new connections, as broader vocabulary knowledge can provide more support for the establishment of new lexical entries. In the present study, receptive vocabulary, measuring vocabulary breadth, was a significant predictor of children’s performance in all three testing tasks, such that children with better receptive vocabulary showed higher accuracy across all test sessions. Consistent with previous studies, children with better vocabulary knowledge showed greater recall of new word meanings (Adlof & Patten, 2017) and word forms (James et al., 2019; 2021).

Our exploratory analyses also sought to understand the contribution of general language processing and phonological memory capacity to word learning by measuring nonword repetition performance. Akin to the results for receptive vocabulary, nonword repetition scores were also a significant predictor of overall performance in all three tasks (Appendix Tables B1-B3). Thus, a better capacity to maintain phonemes in short-term memory may foster a more efficient establishment of new representations. Furthermore, children with better nonword repetition performance show greater overnight improvements in phonological cued recall. Past studies into word learning in individuals with phonological memory weaknesses point towards poor nonword repetition performance as an indicator of less effective phonological encoding, thus limiting new words that were subsequently stored and consolidated (Bishop et al., 2012). Here, with typically developing children, we show that having good phonological processing skills provides a general benefit for learning new words and allows for more efficient overnight consolidation. This highlights the importance of not only understanding the most beneficial timescale for word learning but also understanding to whom it may be the most beneficial.

## Conclusions

This study builds upon previous research and theoretical frameworks of learning and memory, examining developmental differences in the benefits of consolidation mechanisms and prior semantic knowledge in supporting vocabulary acquisition. Critically, we demonstrate a developmental shift in the reliance on underlying mechanisms and processes that support word learning, whilst controlling for initial levels of learning. Directly comparing the memory of new words, children’s greater consolidation resulted in better overall learning than adults. Children showed continuous improvements in the recall for newly acquired word forms across a week regardless of the type of semantic information the words were paired with, with individual differences in global vocabulary and language abilities among children also contributing to the effectiveness of new word form recall. In contrast, the benefits of locally available semantic information, especially semantic information that forms easy association with existing knowledge, were found consistently across test sessions in adults, suggesting that they were more able to utilise semantic information available at encoding. This pattern of results is partially consistent with the CLS framework that prior knowledge-consistent information can be integrated directly, such that a graded semantic benefit was observed in adults’ word learning across semantic conditions. However, the present results did not show greater consolidation across time for less consistent items with existing knowledge, as suggested by the need for repeated hippocampal reactivations and a prolonged consolidation period. Therefore, the present evidence suggests that not only is it important to establish the timecourse and mechanisms of semantic contributions to word learning across the lifespan, but it is also crucial to identify individual differences to understand *whom* this semantic information benefits.

# Reference

Adlof, S. M., & Patten, H. (2017). Nonword Repetition and Vocabulary Knowledge as Predictors of Children’s Phonological and Semantic Word Learning. *Journal of Speech, Language, and Hearing Research: JSLHR*, *60*(3), 682–693. https://doi.org/10.1044/2016\_JSLHR-L-15-0441

Antony, J. W., Ferreira, C. S., Norman, K. A., & Wimber, M. (2017). Retrieval as a Fast Route to Memory Consolidation. *Trends in Cognitive Sciences*, *21*(8), 573–576. https://doi.org/10.1016/j.tics.2017.05.001

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*, *52*(1), 388–407. https://doi.org/10.3758/s13428-019-01237-x

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*, 1–48. https://doi.org/10.18637/jss.v067.i01

Bermúdez-Margaretto, B., Beltrán, D., Shtyrov, Y., Dominguez, A., & Cuetos, F. (2020). Neurophysiological Correlates of Top-Down Phonological and Semantic Influence during the Orthographic Processing of Novel Visual Word-Forms. *Brain Sciences*, *10*(10). https://doi.org/10.3390/brainsci10100717

Bishop, D. V. M., Barry, J. G., & Hardiman, M. J. (2012). Delayed retention of new word-forms is better in children than adults regardless of language ability: a factorial two-way study. *PloS One*, *7*(5), e37326. https://doi.org/10.1371/journal.pone.0037326

Borovsky, A., Ellis, E. M., Evans, J. L., & Elman, J. L. (2016). Lexical leverage: category knowledge boosts real-time novel word recognition in 2-year-olds. *Developmental Science*, *19*(6), 918–932. https://doi.org/10.1111/desc.12343

Buysse, D. J., Reynolds, C. F., 3rd, Monk, T. H., Berman, S. R., & Kupfer, D. J. (1989). The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry Research*, *28*(2), 193–213. https://doi.org/10.1016/0165-1781(89)90047-4

Cain, K., & Oakhill, J. (2011). Matthew effects in young readers: reading comprehension and reading experience aid vocabulary development. *Journal of Learning Disabilities*, *44*(5), 431–443. https://doi.org/10.1177/0022219411410042

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*(6), 671–684. https://doi.org/10.1016/S0022-5371(72)80001-X

Davis, M. H., & Gaskell, M. G. (2009). A complementary systems account of word learning: Neural and behavioural evidence. *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, *364*(1536), 3773–3800. <https://doi.org/10.1098/rstb.2009.0111>

Diakidoy, I.-A. N. (1998). The role of reading comprehension in word meaning acquisition during reading. *European Journal of Psychology of Education*, *13*(2), 131–154. https://doi.org/10.1007/bf03173086

Drosopoulos, S., Schulze, C., Fischer, S., & Born, J. (2007). Sleep’s function in the spontaneous recovery and consolidation of memories. *Journal of Experimental Psychology. General*, *136*(2), 169–183. https://doi.org/10.1037/0096-3445.136.2.169

Dumay, N., & Gaskell, M. G. (2007). Sleep-associated changes in the mental representation of spoken words. *Psychological Science*, *18*(1), 35–39. https://doi.org/10.1111/j.1467-9280.2007.01845.x

Dunn, L. M., Dunn, L. M., Whetton, C., & Burley, J. (1997). *British Picture Vocabulary Scale (2nd ed.)*. NFER-Nelson.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G\*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/bf03193146

Gaskell, M. G., Cairney, S. A., & Rodd, J. M. (2019). Contextual priming of word meanings is stabilized over sleep. *Cognition*, *182*, 109–126. https://doi.org/10.1016/j.cognition.2018.09.007

Gordon, K. R., Lowry, S. L., Ohlmann, N. B., & Fitzpatrick, D. (2022). Word Learning by Preschool-Age Children: Differences in Encoding, Re-Encoding, and Consolidation Across Learners During Slow Mapping. *Journal of Speech, Language, and Hearing Research: JSLHR*, 1–22. https://doi.org/10.1044/2022\_JSLHR-21-00530

Havas, V., Taylor, J., Vaquero, L., de Diego-Balaguer, R., Rodríguez-Fornells, A., & Davis, M. H. (2018). Semantic and phonological schema influence spoken word learning and overnight consolidation. *Quarterly Journal of Experimental Psychology* , *71*(6), 1469–1481. https://doi.org/10.1080/17470218.2017.1329325

Henderson, L. M., Devine, K., Weighall, A., & Gaskell, G. (2015). When the daffodat flew to the intergalactic zoo: Off-line consolidation is critical for word learning from stories. *Dev. Psychol.*, *51*(3), 406–417.

Henderson, L. M., & James, E. (2018). Consolidating new words from repetitive versus multiple stories: Prior knowledge matters. *Journal of Experimental Child Psychology*, *166*, 465–484. https://doi.org/10.1016/j.jecp.2017.09.017

Henderson, L. M., Weighall, A., & Gaskell, G. (2013). Learning new vocabulary during childhood: effects of semantic training on lexical consolidation and integration. *Journal of Experimental Child Psychology*, *116*(3), 572–592. https://doi.org/10.1016/j.jecp.2013.07.004

Henderson, L. M., Weighall, A. R., Brown, H., & Gareth Gaskell, M. (2012). Consolidation of vocabulary is associated with sleep in children. *Developmental Science*, *15*(5), 674–687. https://doi.org/10.1111/j.1467-7687.2012.01172.x

Hoddes, E., Zarcone, V., & Dement, W. (1972). Stanford sleepiness scale. *Psychophysiology*, *9*, 150. https://doi.org/10.1007/978-3-540-28840-4.pdf#page=1212

Horst & Hout. (2016). *The Novel Object and Unusual Name (NOUN) Database*. http://www.sussex.ac.uk/wordlab/noun

Hulme, R. C., & Rodd, J. M. (2021). Learning new word meanings from story reading: the benefit of immediate testing. *PeerJ*, *9*, e11693. https://doi.org/10.7717/peerj.11693

James, E., Gaskell, M. G., & Henderson, L. M. (2019). Offline consolidation supersedes prior knowledge benefits in children’s (but not adults') word learning. *Developmental Science*, *22*(3), e12776. https://doi.org/10.1111/desc.12776

James, E., Gaskell, M. G., Murphy, G., Tulip, J., & Henderson, L. M. (2023). Word learning in the context of semantic prior knowledge: Evidence of interference from feature-based neighbours in children and adults. *Language, Cognition and Neuroscience*, *38*(2), 157–174. https://doi.org/10.1080/23273798.2022.2102198

James, E., Gaskell, M. G., Pearce, R., Korell, C., Dean, C., & Henderson, L. M. (2021). The role of prior lexical knowledge in children’s and adults' incidental word learning from illustrated stories. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *47*(11), 1856–1869. https://doi.org/10.1037/xlm0001080

James, E., Gaskell, M. G., Weighall, A., & Henderson, L. M. (2017). Consolidation of vocabulary during sleep: The rich get richer? *Neuroscience and Biobehavioral Reviews*, *77*, 1–13. https://doi.org/10.1016/j.neubiorev.2017.01.054

Lenth, R. V. (2023). *Estimated Marginal Means, aka Least-Squares Means [R package emmeans version 1.8.8]*. https://CRAN.R-project.org/package=emmeans

Lewis, P. A., & Durrant, S. J. (2011). Overlapping memory replay during sleep builds cognitive schemata. *Trends in Cognitive Sciences*, *15*(8), 343–351. https://doi.org/10.1016/j.tics.2011.06.004

Mak, M. H. C., Curtis, A. J., Rodd, J. M., & Gaskell, M. G. (2023). Episodic memory and sleep are involved in the maintenance of context-specific lexical information. *Journal of Experimental Psychology. General*, *152*(11), 3087–3115. https://doi.org/10.1037/xge0001435

Mak, M. H. C., Hsiao, Y., & Nation, K. (2021). Anchoring and contextual variation in the early stages of incidental word learning during reading. *Journal of Memory and Language*, *118*, 104203. https://doi.org/10.1016/j.jml.2020.104203

Mak, M. H. C., & Twitchell, H. (2020). Evidence for preferential attachment: Words that are more well connected in semantic networks are better at acquiring new links in paired-associate learning. *Psychonomic Bulletin & Review*, *27*(5), 1059–1069. https://doi.org/10.3758/s13423-020-01773-0

McClelland, J. L. (2013). Incorporating rapid neocortical learning of new schema-consistent information into complementary learning systems theory. *Journal of Experimental Psychology. General*, *142*(4), 1190–1210. https://doi.org/10.1037/a0033812

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*(3), 419–457. https://doi.org/10.1037/0033-295X.102.3.419

McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia Medica: Casopis Hrvatskoga Drustva Medicinskih Biokemicara / HDMB*, *22*(3), 276–282. https://doi.org/10.1016/j.jocd.2012.03.005

Nation, K., Angell, P., & Castles, A. (2007). Orthographic learning via self-teaching in children learning to read English: effects of exposure, durability, and context. *Journal of Experimental Child Psychology*, *96*(1), 71–84. https://doi.org/10.1016/j.jecp.2006.06.004

Nieuwenhuis, R., Te Grotenhuis, M., & Pelzer, B. (2012). influence.ME: Tools for Detecting Influential Data in Mixed Effects Models. *The R Journal*, *4*(2), 38–47. https://doi.org/10.32614/RJ-2012-011

Ohayon, M. M., Carskadon, M. A., Guilleminault, C., & Vitiello, M. V. (2004). Meta-analysis of quantitative sleep parameters from childhood to old age in healthy individuals: developing normative sleep values across the human lifespan. *Sleep*, *27*(7), 1255–1273. https://doi.org/10.1093/sleep/27.7.1255

Olsson, M. A. (2022). *Does spacing retrieval practice lead to a benefit in word learning before and after sleep? A developmental perspective* [Phd, University of York]. https://etheses.whiterose.ac.uk/31875/1/OLSSON\_109031197\_THESIS.pdf

Ouellette, G., & Fraser, J. R. (2009). What exactly is a yait anyway: the role of semantics in orthographic learning. *Journal of Experimental Child Psychology*, *104*(2), 239–251. https://doi.org/10.1016/j.jecp.2009.05.001

Roediger, H. L., 3rd, & Karpicke, J. D. (2006). The Power of Testing Memory: Basic Research and Implications for Educational Practice. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *1*(3), 181–210. https://doi.org/10.1111/j.1745-6916.2006.00012.x

RStudio Team. (2020). *RStudio: Integrated Development Environment for R*. RStudio, PBC. http://www.rstudio.com/

Savage Interactive. (2022). *Procreate* (Version 5.2.6).

Smith, F. R. H., Gaskell, M. G., Weighall, A. R., Warmington, M., Reid, A. M., & Henderson, L. M. (2018). Consolidation of vocabulary is associated with sleep in typically developing children, but not in children with dyslexia. *Developmental Science*, *21*(5), e12639. https://doi.org/10.1111/desc.12639

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, *21*(4), 360–407. https://doi.org/10.1598/RRQ.21.4.1

Tamminen, J., & Gaskell, M. G. (2013). Novel word integration in the mental lexicon: Evidence from unmasked and masked semantic priming. *Q. J. Exp. Psychol.*, *66*(5), 1001–1025. <https://doi.org/10.1080/17470218.2012.724694>

Tamminen, J., Lambon Ralph, M. A., & Lewis, P. A. (2013). The role of sleep spindles and slow-wave activity in integrating new information in semantic memory. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *33*(39), 15376–15381. https://doi.org/10.1523/JNEUROSCI.5093-12.2013

Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (2012). *TOWRE-2 Test of Word Reading Efficiency*. New York Pearson.

Tse, D., Langston, R. F., Kakeyama, M., Bethus, I., Spooner, P. A., Wood, E. R., Witter, M. P., & Morris, R. G. M. (2007). Schemas and memory consolidation. *Science (New York, N.Y.)*, *316*(5821), 76–82. https://doi.org/10.1126/science.1135935

van Rijn, E., Gouws, A., Walker, S. A., Knowland, V. C. P., Cairney, S. A., Gaskell, M. G., & Henderson, L. M. (2023). Do naps benefit novel word learning? Developmental differences and white matter correlates. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *158*, 37–60. https://doi.org/10.1016/j.cortex.2022.09.016

Wagner, R. K., Torgesen, J. K., Rashotte, C. A., & Pearson, N. A. (2016). *Comprehensive Test of Phonological Processing–Second Edition* [Dataset]. <https://doi.org/10.1037/t52630-000>

Wilhelm, I., Rose, M., Imhof, K. I., Rasch, B., Büchel, C., & Born, J. (2013). The sleeping child outplays the adult’s capacity to convert implicit into explicit knowledge. *Nature Neuroscience*, *16*(4), 391–393. https://doi.org/10.1038/nn.3343

Acknowledgements: The authors wish to thank Veniamin Shiron for creating the animal illustrations.

1. For consistency, cross-experimental comparisons of the picture naming and definition recall were also conducted, which were not pre-registered. The comparison for these two tasks showed similar developmental trends as that of the phonological recall task with children performing better than adults. Full results of these comparisons can be found in Appendix C. [↑](#footnote-ref-1)