Paired-associate versus cross-situational: How do verbal working memory and word familiarity affect word learning?

Anne Neveu[1] · Margarita Kaushanskaya[2]

Abstract

Word learning is one of the first steps into language, and vocabulary knowledge predicts reading, speaking, and writing ability. There are several pathways to word learning and little is known about how they differ. Previous research has investigated paired-associate (PAL) and cross-situational word learning (CSWL) separately, limiting the understanding of how the learning process compares across the two. In PAL, the roles of word familiarity and working memory have been thoroughly examined, but these same factors have received very little attention in CSWL. We randomly assigned 126 monolingual adults to PAL or CSWL. In each task, names of 12 novel objects were learned (six familiar words, six unfamiliar words). Logistic mixed-effects models examined whether word-learning paradigm, word type and working memory (measured with a backward digit-span task) predicted learning. Results suggest better learning performance in PAL and on familiar words. Working memory predicted word learning across paradigms, but no interactions were found between any of the predictors. This suggests that PAL is easier than CSWL, likely because of reduced ambiguity between the word and the referent, but that learning across both paradigms is equally enhanced by word familiarity, and similarly supported by working memory.

Keywords Psycholinguistics · Working memory · Paired-associate word learning · Cross-situational word learning

Introduction

Word learning is an important aspect of discovering a new language in adulthood, and can also take place in the native language, for example with rare words (e.g., *recondite, quotidian*) or new concepts (e.g., *teraflop, BEV*). To examine the cognitive processes underlying word learning, researchers have developed laboratory-based word-learning paradigms, with paired-associate word learning (PAL) and cross-situational word learning (CSWL) tasks representing two of the most common tasks tested in laboratory settings.

In PAL, a novel word has a single referent, a translation (e.g., Krepel et al., 2021) or picture representing the word (e.g., Gupta et al., 2004). There is no ambiguity as to the label-referent pairing. In that paradigm, the focus has been on the ability to create a phonological representation of the novel

word and to store it in long-term memory (Kazanas et al., 2020; Litt et al., 2019; Rothkopf, 1957; Steinel et al., 2007; Ylinen et al., 2020). In contrast, in CSWL, there is ambiguity in the possible word-referent mappings, and a novel word is learned by aggregating word and object associations across trials. In an experimental setting, this ambiguity is created by presenting two objects on screen, while two spoken words are presented sequentially. Only after several trials where the same pool of objects is presented in different pairings can the learner infer which word corresponds to which object. Although learning might be predicted to be impeded by the multitude of possible pairings in the environment, evidence in infants (Smith & Yu, 2008, 2013) and adults (Kachergis et al., 2012; Smith et al., 2010; Suanda & Namy, 2012; Yu & Smith, 2007) shows that learning in this context can and does take place successfully.

The two paradigms have been examined under different theoretical umbrellas, limiting our understanding of the common processes underlying word learning. For instance, one of the most common manipulations in PAL is related to word familiarity, and the degree to which long-term knowledge contributes to short-term encoding and retrieval processes (Ellis & Beaton, 1993; Papagno et al., 1991;

✉ Anne Neveu
aneveu@health.ucsd.edu

1 Department of Psychiatry, University of California, San Diego, La Jolla, CA, USA

2 Department of Communication Sciences & Disorders, University of Wisconsin - Madison, Madison, WI, USA

Papagno & Vallar, 1992; Service & Craik, 1993), but this topic has barely been broached in CSWL (see Escudero et al., 2013, for an exception). Similarly, while many studies have examined the role of verbal working memory as the system underlying PAL (e.g., Baddeley et al., 1998, 2017; Ellis & Beaton, 1993; Freedman & Martin, 2001; Gupta, 2003; Kazanas et al., 2020; Litt et al., 2019; Papagno et al., 1991; Papagno & Vallar, 1992; Rothkopf, 1957; Steinel et al., 2007; Ylinen et al., 2020), studies on CSWL have developed from statistical learning theories applied to language (e.g., Aslin, 2017; Conway et al., 2010; Frost et al., 2015; Graf Estes et al., 2007; Mirman et al., 2008; Saffran et al., 1996a, b), and thus have paid very little attention to the memory systems that might underlie CSWL. Yet, although PAL and CSWL differ in the level of ambiguity at word-referent exposure, both paradigms involve associating a word form with a referent, and maintaining this mapping over time.

Therefore, the purpose of this study is twofold: First, to examine the degree to which PAL and CSWL are sensitive to word familiarity effects, and second, to examine the extent to which PAL and CSWL performance is related to verbal working memory in adults. The overarching goal of this study is to examine the degree to which mechanisms that underlie PAL and CSWL are similar or different, keeping PAL and CSWL designs and methods as comparable as possible.

Paired-associate learning and verbal working memory

Extensive research has found an association between verbal memory and PAL in various populations, such as children, adults, and patients with brain damage (Baddeley et al., 1998, 2017; Freedman & Martin, 2001; Gupta, 2003; Papagno et al., 1991; Papagno & Vallar, 1992). The verbal memory system is part of Baddeley and Hitch's model of working memory (Baddeley, 2003; Baddeley & Hitch, 1974), which comprises a central executive, a phonological loop, a visual sketchpad, and an episodic buffer. The component relevant to verbal working memory is the phonological loop, which is a temporary memory store that supports the processing of verbal information through articulatory rehearsal (subvocal speech). According to the model, the central executive is the locus of control for directing attention to the phonological loop while learning novel words. The role of the phonological loop is to temporarily store verbal representations, to generate longer-lasting ones in memory. Beyond supporting memory for sequences of familiar words, this process has been suggested to support novel word learning (Baddeley et al., 1998, 2017).

The role of the phonological loop has been established both in children and in adult word learning, where phonological memory capacity has been found to be positively associated with word-learning performance. For example, in studies with adults and children, participants with better ability to repeat nonwords and/or series of digits (two measures that typically index phonological working memory) tend to perform better on PAL tasks (e.g., Atkins & Baddeley, 1998; Baddeley et al., 1988, 2017; Gathercole et al., 1992, 1999; Service, 1992; Speciale et al., 2004). In addition, when verbal memory load is directly tapped, for example with phonologically similar, or long stimuli to learn, or in conditions of articulatory suppression (e.g., having to learn while repeating the word "the"), unfamiliar word learning is disrupted (Papagno et al., 1991; Papagno & Vallar, 1992).

While abundant evidence suggests that verbal working memory supports word learning, long-term memory has been found to contribute to word learning as well, particularly when learning involves phonologically familiar items in both children (Gathercole et al., 1991, 1999; Magro et al., 2018) and adults (e.g., Ellis & Beaton, 1993; Majerus et al., 2004; Papagno et al., 1991; Papagno & Vallar, 1992; Service & Craik, 1993). This conclusion has been drawn from studies demonstrating that familiar stimuli (such as real words, and phonologically viable nonwords) tend to be recalled with higher accuracy than unfamiliar stimuli (such as nonwords, and phonologically less viable nonwords). For nonwords, familiarity effects emerge when their phonotactics conform to that of the known language (Ellis & Beaton, 1993; Papagno et al., 1991; Papagno & Vallar, 1992; Service & Craik, 1993). For example, the findings in Papagno and Vallar (1992) showed that only the learning of novel words was affected by phonological similarity and item length, but not the learning of associations between words already known. Evidence for this effect was also found in Ellis and Beaton (1993), where monolingual speakers of English had to learn German translations of English words. The authors showed that words that were easier to pronounce, and so closer to English phonology, were better remembered.

This familiarity effect suggests that prior knowledge of lexical and phonological features, stored in long-term memory, is solicited at learning. In other words, the phonological loop relies on long-term memory during PAL (e.g., Ellis & Beaton, 1993; Majerus et al., 2004; Papagno et al., 1991; Papagno & Vallar, 1992; Service & Craik, 1993). Because verbal working memory has limited capacity (e.g., Cowan et al., 2007), less familiar or unfamiliar novel words (which elicit less support from representations in long-term memory) require more involvement of the phonological loop to be encoded into long-term memory.

Evidence from neurophysiological studies also corroborates the word familiarity effect (e.g., Hanten & Martin, 2001; Martin & Saffran, 1999). For example, it has been shown that a highly educated patient (Ph.D. in molecular biology) who had had a verbal working memory

impairment was able to rely on lexical, semantic, and visual representations to process and retain phonologically familiar words. However, this patient had reduced capacity in processing and retaining nonwords and foreign language novel vocabulary.

To summarize, PAL has been examined through the framework of Baddeley's working memory model, with verbal working memory as the primary mechanism underlying it. The word familiarity effect has been central to establishing how verbal working memory works in conjunction with long-term memory to support word-learning processes. However, very little of this research has crossed into CSWL literature.

Cross-situational word learning and verbal working memory

The CSWL paradigm arose from the statistical learning literature and thus has not examined what role working memory might play in novel word learning. Two approaches to learning are posited in CSWL: associative learning (e.g., Yu & Smith, 2007, 2011) and hypothesis testing (Trueswell et al., 2013, see Yurovsky & Frank, 2015, for an integrative account). Associative learning relies on aggregating information on sets of word-referent co-occurrences to eventually infer word meanings. Hypothesis testing stipulates that hypotheses are formulated on word-referent mappings, evidence related to hypotheses is encoded, and word meanings are inferred from these hypotheses. Neither approach allocates a central role to working memory in CSWL. However, novel words have to be processed and retained in CSWL, therefore paired-associate and cross-situational word learning could be supported by verbal working memory similarly. Moreover, there is some evidence for the role of working memory in CSWL, although this association has generally been established via a different approach from the one used in PAL (Mulak et al., 2019; Vlach & Sandhofer, 2014; Yu & Smith, 2007).

The effect of verbal working memory on CSWL performance has been tested indirectly by varying the level of ambiguity during the teaching phase of the experiment (Mulak et al., 2019; Vlach & Sandhofer, 2014; Yu & Smith, 2007). For instance, in Mulak et al.'s study (2019), adults were tested in one of four conditions, differing in the level of ambiguity at novel word-referent exposure, from a 1 × 1 pairing (as in PAL), where one word is auditorily played while one picture is displayed on the screen, to 2 × 2, 3 × 3, and 4 × 4 pairings. In each condition, working memory load increased such that one additional word and one additional object to learn were presented. For example, at the highest load in the 4 × 4 condition, phonological and visual encoding for four words and four objects had to be

associated, whereas half the number of stimuli had to be processed in the 2 × 2 condition. Findings showed that while participants learned above chance in all conditions, learning performance was best in the 1 × 1 condition, providing evidence for the first time that PAL is easier than CSWL. Findings were similar in Vlach and Sandhofer (2014) and Yu and Smith (2007), who contrasted word-learning performance in 2 × 2, 3 × 3, and 4 × 4 conditions. Better performance in the less ambiguous conditions was interpreted as reflecting a lighter load on working memory resources through fewer word-referent mappings to learn, although working memory was not directly tested in these studies (Mulak et al., 2019; Vlach & Sandhofer, 2014; Yu & Smith, 2007).

While working memory load has been indirectly tested in CSWL, the role of word familiarity, and thus of long-term memory, has received very little attention in the CSWL literature. Like Mulak et al. (2019), Escudero et al. (2013) manipulated phonological overlap within the pool of novel words in a CSWL task. In that study, participants were tested in a CSWL paradigm where stimuli were non-minimal pairs, near-minimal pairs, vowel-minimal pairs, or consonant-minimal pairs (*bon, pon, ton, don* and *dit, dut, deet, doot*). Participants were able to learn words above chance in all four conditions, but performance on the vowel minimal pairs was significantly worse than in the other three conditions. This suggests that the phonological loop was particularly taxed when encoding phonological detail for vowel minimal pairs in monosyllabic nonwords, which disrupted learning. While phonologically similar stimuli increased the difficulty of the learning task in that experiment, it is possible that familiar words, namely words that are already known, would facilitate novel word learning, in the absence of phonological overlap in the experimental stimuli. However, considering that CSWL has a higher level of ambiguity than PAL, which imparts a higher load on working memory and has led to decreased levels of learning performance (Mulak et al., 2019), it is possible that the phonological facilitation effect associated with familiarity is weaker in CSWL compared to PAL.

To summarize, robust findings in PAL suggest that verbal working memory is a core mechanism that supports novel word learning, and that familiar novel words enjoy a learning advantage, likely due to the involvement of long-term memory in verbal working memory. In CSWL, disruptions in learning due to phonologically similar stimuli suggest that the task is sensitive to phonological detail. Moreover, CSWL performance declines with increased ambiguity at learning, indirectly suggesting a role of working memory load. However, neither word familiarity nor the association between CSWL success and working memory has been directly studied in adults.

Current study

In the present study, we examined the extent to which word-learning performance might vary as a function of word-learning paradigm – PAL versus CSWL, and word familiarity – unfamiliar (nonwords) vs. familiar (existing English words) words. Additionally, we examined whether verbal working memory would predict word learning similarly or differently across paradigms and word types.

Our primary prediction was that word-learning performance would be higher in PAL than in CSWL, in line with the hypothesized higher verbal working memory load associated with CSWL (Mulak et al., 2019; Vlach & Sandhofer, 2014; Yu & Smith, 2007), and consistent with a single prior study that has examined this question in children (Vlach & DeBrock, 2019). In line with Baddeley's model of working memory (Baddeley, 2003; Baddeley & Hitch, 1974), and strong findings regarding the facilitative role played by phonological familiarity in novel word learning (Baddeley et al., 1998, 2017), we also hypothesized that word-learning performance would be higher for familiar versus unfamiliar words, across both paradigms. We additionally hypothesized that higher verbal working memory performance would be associated with higher word-learning performance across both paradigms and word types. Critically, we also anticipated a number of interactions among the variables of interest, testing the critical prediction that verbal working memory would be differentially involved in PAL versus CSWL, and in the learning of phonologically unfamiliar versus familiar novel words. Because CSWL is expected to be more taxing than PAL, it is possible that working memory would associate more strongly with word-learning performance in this paradigm. However, because CSWL relies on unconscious statistical learning more than PAL, it may not tax working memory to the same extent as PAL, and demands on working memory may be neutralized for this task. Additionally, it is possible that verbal working memory performance would associate more strongly with word-learning performance when words are less familiar, across paradigms. Finally, learning condition and word familiarity could compound such that working memory performance would be more strongly associated with word-learning performance on unfamiliar words in CSWL.

Method

Participants

We recruited 136 participants via the online platform Prolific (Palan & Schitter, 2018). This sample size was derived from a power analysis conducted using the function "modelPower" in the lmSupport package (Curtin, 2018) in R Studio (v.

4.0.0; R Core Team, 2020) to determine the sample sizes needed to detect moderate effects of subject-level variable and interaction terms on word-learning accuracy. A meta-analysis looking at second language word learning from spoken input found a large effect of vocabulary gains ($g = 1.05$, equivalent to d, with the exception that g corrects for small-sample bias) (de Vos et al., 2018). However, this figure conflates studies with different populations, ages, treatments, and testing procedures, and does not include all the variables included in the present study. Therefore, we chose a more conservative, medium effect size of $d = 0.7$.

Our filters on Prolific included participant location in the USA, US nationality, age between 18 and 40 years, English as a first language, no knowledge of another language, no language related-disorders, and no hearing difficulties. Participants were compensated at a rate of US$10 per hour. Participants provided informed consent to voluntarily take part in the study through a form approved by the Institutional Review Board of the University of Wisconsin – Madison. Participants were randomly assigned to a word learning condition (PAL or CSWL). Participant characteristics can be found in Table 1.

Participants completed the Language Experience and Proficiency Questionnaire (LEAP-Q, Marian et al., 2007), adapted to an online format, reporting languages known, percentage of exposure to the language(s), and self-rated proficiency in each language. Participants who indicated exposure to a language other than English more than 5% of the time were excluded (six participants).

Materials

Stimuli

Novel word stimuli consisted of six disyllabic unfamiliar words (nonwords) chosen from the Gupta et al. database (2004), half with first-syllable stress, half with second-syllable stress, and following English phonology. There were additionally six English words, matched to the unfamiliar words on stress pattern, and on biphone frequencies using the Clearpond database (Marian et al., 2012).

These 12 words were each paired with a novel object. The pictures for the novel objects were chosen from the Novel Object and Unusual Name (NOUN) Database (Horst & Hout, 2016). The objects were chosen such that they had average high novelty ratings and were comparable in saliency. We selected 12 novel object pictures that we paired with the 12 words to create the teaching list. To control for variability in word-object pairing due to any subjective saliency in the pictures, we changed the word-object pairings and created a second list, using the same pool of stimuli, such that the pictures that were paired with familiar words in the first list were paired with the unfamiliar words in the second list, and vice versa. Assignment of lists to

Table 1 Participant characteristics

	PAL	CSWL	Welch's t-test	Overall
	Mean (SD) N = 62	Mean (SD) N = 64	t[d]	Mean (SD) N = 126
Age (years)	29.73 (6.83)	30.23 (6.09)	-0.44	29.98 (6.44)
Gender[a]	33 (F) 24 (M) 3 (N-B) 2(NR)	31 (F) 30 (M) 3 (N-B)	0.60[d]	64 (F) 54 (M) 6 (N-B) 2 (NR)
Years of education[b]	4.65 (1.24)	4.50 (1.28)	0.64	4.57 (1.26)
Nonverbal IQ[c]	62.14 (14.49)	63.92 (17.74)	-0.62	63.04 (16.18)
English vocabulary[c]	52.12 (18.85)	50.98 (20.10)	0.31	51.54 (19.43)
Backward digit-span[c]	58.57 (16.68)	58.98 (17.88)	-0.14	58.78 (17.23)
Nonword repetition[c]	51.47 (22.04)	51.19 (18.88)	0.07	51.33 (20.38)

[a]Females (F), Males (M), Non-binary (N-B), no reply (NR)

[b]Coded 1 through 8, from "Less than High School" to "Ph.D./M.D./J.D."

[c]Percent correct

[d]For the variable of gender, Pearson's Chi-squared test was run instead as the data are categorical

[*]$p < .05$. [**]$p < .01$. [***]$p < .001$

participants was randomized across participants. Word and object stimuli are presented in Appendix B.

Learning conditions

Paired-associate word learning The PAL task contained three teaching cycles, in which the 12 novel word-object pairings were presented three times, across three blocks. Presentation was randomized within blocks and block order was randomized across participants. Participants first saw instructions (see Appendix A1-i) describing the word-learning task. Then, a black fixation cross appeared in the center of the white screen for 1,000 ms. The novel object picture appeared on the screen, and its name was spoken with a 400-ms delay. The object remained visible on the screen for 3,400 ms longer before the software automatically moved on to show the next object. The teaching cycle contained one attention check, placed after the first block for all participants. The check consisted of a sentence and solicited a response: "We're checking that you're still here. Press "next" to continue learning."

Once the teaching cycle was completed, the experiment moved to the testing cycle. Each novel word was tested three times in a four-alternative forced-choice task. Participants were presented with instructions to choose the picture that matched the word played, among the four options presented (see instructions in Appendix A1-ii). A fixation cross appeared for 1,500 ms, then the four pictures appeared on the screen but were not yet clickable as there was a 700-ms delay until the target word was spoken. Pictures became clickable (a thin black frame around them appeared) at the target word offset.

The next testing trial appeared after the participant clicked on the picture chosen. Order of presentation across trials was randomized within and across blocks for all participants. There were always four different objects presented on the screen, and pairings were pseudorandomized such that two words within the four never appeared together more than six times. Each picture appeared between 10 and 14 times on screen, in all four zones (top left/right, bottom left/right). Responses were scored as 1 or 0 depending on whether the participant's answer matched the correct answer.

Cross-situational word learning The design of the CSWL task followed the same structure as the PAL when applicable, to enhance comparability between paradigms. Participants first saw instructions that did not give away that they were going to learn novel words (see Appendix A2-i). Next, a black fixation cross appeared in the center of the white screen for 1,000 ms. The 12 novel word-object parings were presented three times, in blocks. Presentation was similarly randomized within and across blocks. Two novel objects were presented on-screen and two novel words were spoken one after the other, after a 400-ms delay from picture onset. To keep the same timing per trial as in PAL, one object was named between 400 ms and 1,700 ms, and the second object was named between 1,700 ms and 3,400 ms. The screen automatically moved to the next set of stimuli at 3,400 ms. As in PAL, a single attention check was placed after the first block. Presentation and naming of the novel objects was counterbalanced left and right, such that in each block, each of the 12 words appeared once on the left and once on the right. Each novel word was heard twice in one block and six times overall. The left-right pairings were always different.

Half of the unfamiliar words and half of the familiar words were named on their matching side four times out of the six (i.e., the picture presented on the left matched the first word uttered, and the picture on the right matched the second word uttered). The remaining half were named twice on their matching side and four times on their non-matching side. The six unfamiliar words were presented twice with another unfamiliar word and four times with a familiar word. The six familiar words were presented twice with another familiar word and four times with an unfamiliar word.

The testing phase was the same as in PAL, except for the instructions that asked to pick the picture that matched the sound played (see instructions in Appendix A2-ii), and the pseudorandomization in creating the three testing blocks additionally ensured that the pairings seen in the teaching cycle were not systematically reproduced at test.

Procedure

Participants completed the experiment independently and asynchronously on the Gorilla Experiment Builder platform (Anwyl-Irvine et al., 2020). The experiment began with the consent form. If participants did not consent to audio, they were redirected to a version of the experiment without the tasks requiring audio recording (the Woodcock Johnson test of English vocabulary and the non-word repetition task), such that all their other data could still be included in the dataset for analysis. A sound check was included to ensure participants adjusted their volume to a comfortable level, and that auto play worked. Participants were encouraged to wear headphones or to move to a quiet area. In both versions, the word-learning experiment was next, and set to randomly assign participants to either the PAL list A or B, or the CSWL list A or B, in equal ratios.

Participants then completed two working memory tasks: a backward digit-span task (van den Noort et al., 2006; Wechsler, 1997) and a nonword repetition task (Lado, 2017), in this order. Previous research on individual differences in verbal working memory suggests that the two tasks load onto separate factors within the construct of verbal working memory, with the nonword repetition task corresponding to its phonological component (Gathercole et al., 2004; Warmington et al., 2019), and the backward digit-span task corresponding to its executive component (Warmington et al., 2019). Both aspects of verbal working memory have been linked to vocabulary acquisition (Baddeley et al., 1998), with the more robust relationship observed for nonword repetition than for backward digit-span tasks (Baddeley et al., 1998). Therefore, both tasks were included to examine whether they would differentially predict learning across existing words and nonwords. The backward digit-span contained 16 trials, from 2 to 9 digits, with two trials per level. Digit sequences were recorded by a native American English speaker and stimuli were normalized at 70 dB. Participants were instructed to listen carefully to the digit sequences and retype them in reverse order (see Appendix A3 for full instructions). There was one practice trial. The box to type in the response only appeared after the audio finished playing, to limit the possibility of typing the numbers as they were being spoken. There was no time limit per trial. The task was designed to be automatically scored and individual sum scores were converted to percent correct.

For the nonword repetition task, only stimuli between three and five syllables (three pairs per level for a total of nine trials) were retained after piloting revealed ceiling effects beyond this threshold. The original recordings from Lado (2017) were used and normalized at 70 dB. Participants were instructed to listen to the pairs of nonsense words and then record themselves repeating them (see Appendix A4 for full instructions). Participant recordings were scored for accuracy by two research assistants (with each RA scoring about 50% of the data) and 10% of these data were double-scored for reliability by a third research assistant. Research assistants were instructed to score following an "all-or-none" principle, such that if there was any error in repeating at least one phoneme, a score of 0 was recorded for that nonword, otherwise a score of 1 was recorded to indicate an accurate production. An intraclass correlation coefficient using two-way random effects and a single-rater unit was computed between the main scorers and the second scorer for total score per item. Results showed good agreement between the double-scorer and the combined data of the initial two scorers (ICC = .86, p < .001). 110 cases had to be removed due to lack of response or poor audio quality that prevented scoring (9.86% of the data). Individual sum scores were converted to percent correct.

Participants were allowed to opt out of the nonword repetition task if they did not consent to audio recordings. For the participants who provided both the nonword repetition data and the backward digit-span data (n = 109), the two tasks positively correlated ($t(107) = 3.79$, $r = .34$, p < .0001). However, because only a subset of participants provided data for both measures, only the backward digit-span was used in models as it had no missing data.

Participants also completed the Woodcock Johnson Picture Vocabulary (English – Woodcock-Johnson III – Tests of Achievement; Mather & Woodcock, 2001), to test their vocabulary knowledge in English. They were also allowed to opt-out of this task if they did not consent to audio recordings. This test was adapted to an online format and shortened, so that participants started at the adult starting point as indicated in the test manual. Participants saw six pictures on the screen and had to record themselves naming the pictures (see Appendix A5 for full instructions). Recordings were scored for accuracy by two research assistants (with each RA scoring about 50% of the data), and

10% of this data was double-scored for reliability by a third research assistant. Research assistants were instructed to give 1 point for each accurate word, and 0 points for inaccurate productions. A list of accurate and inaccurate word options was provided to help scoring. An intraclass correlation coefficient using two-way random effects and a single-rater unit was computed between the main scorers and the second scorer for total score per item. Results showed excellent agreement between the double-scorer and the combined data of the initial two scorers (ICC = .99, $p < .001$). In the initial consent form, participants ticked a box to indicate whether they consented to audio recordings – five participants not previously excluded did not consent. Additionally, 45 items had to be removed due to lack of a response or poor audio quality that prevented scoring (8.82% of the data). Individual sum scores were converted to percent correct.

Finally, participants completed a measure of non-verbal intelligence, the Visual Matrices of the *Kaufman Brief Intelligence Test* (KBIT-2) (Kaufman & Kaufman, 2004), which was also designed to be automatically scored. Participants were instructed to choose the picture that would best complete a relationship or rule in a set of pictures or patterns (see Appendix A6 for full instructions). Time per trial was limited to 30 s. Feedback was provided on the first three trials, as in the test manual: a green tick mark was shown if the answer was correct, otherwise a red cross appeared. There was no opportunity for self-correction during trials with feedback, and trials did not drop back to an earlier level if an incorrect response was given. Individual sum scores were converted to percent correct.

Participants who did not consent to audio recordings completed the backward digit-span after the word-learning task, and then the KBIT-2. The experiment was set to take ~ 25 min.

Analyses

As the first step, we excluded participants who took longer than average to respond to an attention check inserted after the first teaching cycle of the word learning experiment. The varPlot function of the lmSupport package in R (Curtin, 2018) helped determine these data points based on reaction time and frequency data (four participants).

Similarly, we removed testing trials where reaction times were over three standard deviations above a participant's mean (83 cases excluded, 1.83% of the data), and below 150 ms, to discard any automatic responses (e.g., that would indicate clicking without looking at or considering the options for response) (85 cases excluded, 1.91% of the data).

We examined the extent to which learning condition and word type increased or decreased the likelihood of learning a word. The unit of analysis is a binary outcome variable corresponding to each response for each test item treated (1

for correct, or 0 for incorrect), for each of the three testing cycles. We constructed logistic mixed-effects models in R Studio, version 4.0.0 (lme4 package, Bates et al., 2015). Logistic regression (or the generalized linear model) is used when the assumptions of normality, constant variance, and linearity of the general linear model are violated, as is the case when the dependent variable is binary (Judd et al., 2017). Moreover, mixed-effects models allow researchers to analyze data that is nested (e.g., repeated measures on a participant within a condition) and crossed (e.g., items used in more than one condition) (Boisgontier & Cheval, 2016; Clark, 1973). Our dependent variable was item-level dichotomous accuracy data (0, 1) from the testing cycles, aggregated over the three testing blocks. Each model examined whether predictors increased or decreased the likelihood (log-odds) of making a correct response during testing. Model assumptions were tested using the DHARMa package (Hartig, 2022), and were satisfied.

We included Learning Condition (coded -0.5 for PAL, and 0.5 for CSWL), Word Type (coded -0.5 for nonwords, and 0.5 for existing words) and the backward digit-span score centered around each participant's mean as fixed effects, and their interaction. Singularity issues emerged when adding a slope for the backward digit-span in the by-item random effects structure, therefore we resolved this issue by simplifying the random effects structure following recommendations of Brauer and Curtin (2018), with by-subject and by-item random intercepts, a by-subject slope for Word Type, and a by-item slope for Learning Condition.

Results

Analyses were conducted on 126 participants (4,368 observations). Word learning performance was significantly above chance (at 25%, testing was a four-alternative forced-choice task) in PAL ($M = 92\%$, $SD = 27\%$; range: 39–100%; $t(2121) = 112.13$, $p < .0001$) and in CSWL ($M = 71\%$, $SD = 45\%$; range: 17–100%; $t(2245) = 48.26$, $p < .0001$). See Table 2 for mean learning proportions by Learning Condition and Word Type.

We interpret all main effects and interactions in terms of odds ratios, used in logistic regression where residuals are binomial, to describe a change in odds in the dependent

Table 2 Average accuracy per word type (familiar or unfamiliar) and learning condition (PAL or CSWL)

	PAL Mean (SD)	CSWL Mean (SD)
Familiar words	0.94 (0.24)	0.73 (0.44)
Unfamiliar words	0.90 (0.30)	0.69 (0.46)

variable as the result of the change in the independent variable of a magnitude of one. The odds ratio is the exponential of the parameter estimate.

The logistic mixed-effects model revealed a significant main effect of Learning Condition such that the odds of participants in the PAL condition accurately identifying the novel words increased by a factor of 6.05 compared to the CSWL condition ($b = -1.80$, $SE = 0.33$, $z = -5.41$, $OR = 6.05$, $p < .0001$). Results also revealed a significant main effect of Word Type such that the odds of identifying familiar words more accurately than unfamiliar words increased by a factor of 1.67 ($b = 0.51$, $SE = 0.22$, $z = 2.30$, $OR = 1.67$, $p < .05$). The main effect of the backward digit-span task was significant ($b = 0.03$, $SE = 0.01$, $z = 2.96$, $OR = 1.03$, $p < .01$), such that as backward digit-span score increased, the odds of accurately identifying novel words increased by a factor of 1.03, averaged across learning conditions and word types. None of the two-way and three-way interactions were significant. See Table 3 for full model results and Fig. 1 to view graphed results.

Additionally, an exploratory model on a smaller dataset, with the nonword repetition task scores used instead of the backward digit-span scores was run, and the findings were largely similar. However, in contrast to the analyses where the backward digit-span was used, these analyses yielded a significant three-way interaction of Learning Condition, Word Type, and nonword repetition ($b = -0.02$, $SE = 0.01$, $z = -2.06$, $OR = 0.98$, $p < .05$). We followed up on the three-way interaction by examining the simple effects of the nonword repetition task for familiar and unfamiliar words, and in the PAL and the CSWL conditions. Only one finding approached significance: higher nonword
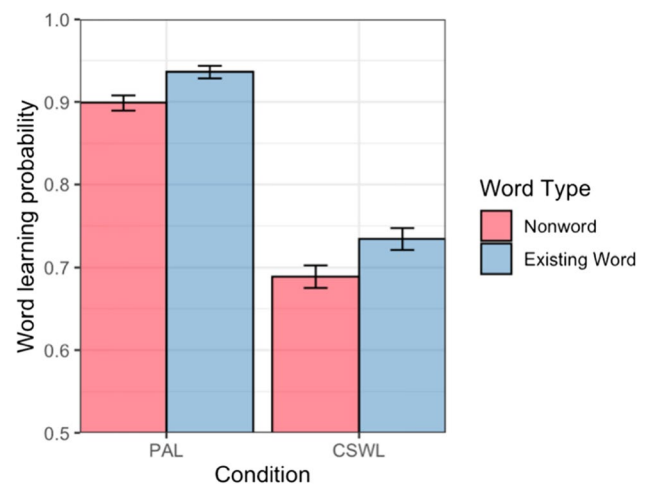
Fig. 1 Word learning probability as a function of Learning Condition and Word Type. Standard error bars represent error at the item level on word learning probability, across learning conditions and word types

repetition task scores for participants in the PAL condition increased the odds of learning familiar words by a factor of 1.02 compared to learning unfamiliar words ($b = 0.02$, $SE = 0.01$, $z = 1.77$, $OR = 1.02$, $p = 0.08$). However, given the marginal nature of this effect, and the smaller dataset that included nonword repetition data, we would not over-interpret this finding.

We also ran a logistic mixed-effects model on the first testing block data, to mitigate the possibility that ceiling effects in the testing data diminished the interactions in the model. The results were highly consistent with the results of the analyses that included all three testing blocks. The model revealed a significant main effect of Learning Condition such that the odds of participants in the PAL condition accurately identifying the novel words compared to those in the CSWL condition increased by a factor of 6.30 ($b = -1.84$, $SE = 0.31$, $z = -6.01$, $OR = 6.30$, $p < .0001$). Results also revealed a significant main effect of Word Type such that the odds of learning familiar words more accurately than unfamiliar words increased by a factor of 1.95 ($b = 0.67$, $SE = 0.19$, $z = 3.48$, $OR = 1.95$, $p < .001$). The main effect of the backward digit-span task was also significant, such that as backward digit-span score increased, the odds of novel word-learning performance increased by a factor of 1.03, averaged across learning conditions and word types ($b = 0.03$, $SE = 0.01$, $z = 3.13$, $OR = 1.03$, $p < .01$). The interaction of Learning Condition by Word Type was marginal ($b = -0.72$, $SE = 0.39$, $z = -1.88$, $OR = 0.49$, $p = 0.06$), such that the odds of learning familiar words more accurately than unfamiliar words in PAL significantly increased by a factor of 2.80 ($b = 1.03$, $SD = 0.34$, $z = 3.03$, $OR = 2.80$, $p < .01$). In

Table 3 Word learning accuracy by learning condition, word type, and backward digit-span performance

	Accuracy	
	b (SE)	z
Intercept	2.39 (0.17)	12.85[***]
Condition	-1.80 (0.33)	-5.41[***]
Word Type	0.51 (0.22)	2.30[*]
Digit Span	0.03 (0.01)	2.96[**]
Condition × Word Type	-0.29 (0.29)	-1.00
Condition × Digit Span	-0.002 (0.02)	-0.09
Word Type × Digit Span	0.004 (0.01)	0.59
Condition × Word Type × Digit Span	-0.01 (0.01)	-0.74

Model: Accuracy ~ Condition * Word Type * Digit Span + (1+ Word Type | Participant) + (1+ Condition |Target Word)

[***] $p < .001$, [**] $p < .01$, [*] $p < .05$

Table 4 Word learning accuracy by learning condition, word type, and backward digit-span performance on the first testing block

	Accuracy	
	b (SE)	z
Intercept	1.93 (0.15)	12.58***
Condition	-1.84 (0.31)	-6.01***
Word Type	0.67 (0.19)	3.48***
Digit Span	0.03 (0.01)	3.13**
Condition × Word Type	-0.72 (0.39)	-1.88 .
Condition × Digit Span	-0.01 (0.02)	-0.56
Word Type × Digit Span	0.01 (0.01)	0.49
Condition × Word Type × Digit Span	-0.01 (0.02)	-0.64

Model: Accuracy ~ Condition * Word Type * Digit Span + (1| Participant) + (1|Target Word)

Due to the removal of two-thirds of the data, the model converged but remained singular after maximizing the simplification of the random-effects structure and setting the adaptive Gauss-Hermite approximation to the log-likelihood to 0 (nAGQ = 0)

*** $p < .001$, ** $p < .01$, * $p < .05$

CSWL, these odds were only marginally better ($b = 0.31$, $SD = 0.18$, $z = 1.72$, $OR = 1.36$, $p = 0.09$). See Table 4 for full model results and Fig. 2 to view graphed results.

Discussion

Previous research has examined PAL and CSWL separately, limiting the ability to identify similarities or differences between the processes that support different types of word
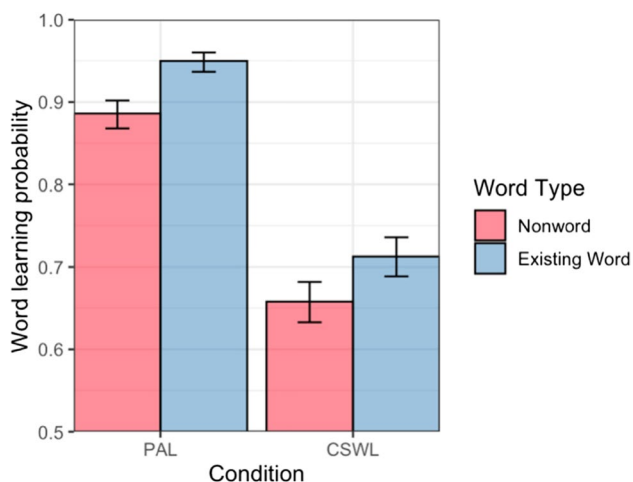
Fig. 2 Word learning probability as a function of Learning Condition and Word Type, on the first block of testing data. Standard error bars represent error at the item level on word learning probability, across learning conditions and word types

learning. While research suggests that more familiar words are learned better and that verbal working memory is positively associated with word learning in PAL (e.g., see Baddeley et al., 1998, 2017, for a review, but see Service & Craik, 1993), a limited number of studies has examined these effects in CSWL (Escudero et al., 2013; Mulak et al., 2019). Therefore, we examined novel word learning in PAL and CSWL, varying word familiarity and measuring participants' verbal working memory. Our findings confirm our hypotheses that word learning in PAL would be more successful than in CSWL, that familiar novel words would be easier to learn than unfamiliar nonwords, and that verbal working memory would be positively associated with word learning across both learning conditions. However, contrary to our hypotheses, no interaction between word type and condition was found, and no interaction was observed among verbal working memory, condition, and word type.

These findings suggest that PAL is an easier learning paradigm, likely due to its lack of ambiguity between word and object compared to CSWL, and in line with Mulak et al. (2019). However, while novel word learning was overall lower in CSWL, our hypothesis as to the compounded effect of learning unfamiliar words in CSWL resulting in worst learning performance was not verified. Instead, while learning of unfamiliar words was significantly worse than learning of familiar words in each paradigm taken separately, the size of the learning gap did not significantly differ across paradigms. Furthermore, this effect was independent of working memory performance, for both paradigms and both word types. The working memory literature and PAL studies suggest that the familiarity effect supporting novel word-object mappings of familiar words emerges from the interaction between long-term memory and working-memory (e.g., Ellis & Beaton, 1993; Majerus et al., 2004; Papagno et al., 1991; Papagno & Vallar, 1992; Service & Craik, 1993). Our findings are consistent with this literature, and extend it to CSWL. However, contrary to the evidence in this body of literature, we did not find a stronger association between verbal working memory and unfamiliar word learning in this study.

This lack of association could be explained by the fact that some interference may have emerged from the existing meanings of the familiar words, where the known meanings would have had to be inhibited in the novel word learning process, compared to that of the nonwords, possibly leveling the need to rely on verbal working memory across word types. In addition, it is also possible that some interference took place within the nonwords, which may have been more similar among each other than the familiar words, making it difficult to disentangle the role of verbal working memory in learning both types of words.

Another possibility, related to subjects-level effects, is that participants in this age group (18–40 years old) were

at peak cognitive performance, as suggested by Service and Craik (1993), which may have masked any extra-reliance on the working memory system for unfamiliar items in the present experiment. In Service and Craik's study (1993), correlational analyses revealed that while more phonologically familiar items were learned better across groups, only participants in the older group showed an association between verbal working memory and learning of phonologically unfamiliar novel words. However, the interpretation of this study should be made with caution as it may have been underpowered (Gupta, 2003). In our study, the spread in the backward digit-span performance ranged from 12.5 to 100, with a mean of 58.67, median at 56.25 and mode at 50, suggesting a relatively normal distribution of scores in this sample, and thus making us question this interpretation.

Alternatively, it could be that a role of verbal memory may not have emerged because the stimuli, whether familiar or unfamiliar, conformed to participants' native language phonotactics (Speciale et al., 2004). Unfamiliar words may not have been unfamiliar enough to trigger the "word unfamiliarity effect," such that the unfamiliar words may still have been supported by long-term phonological knowledge, similarly to the familiar words. To test this hypothesis, an additional condition with unfamiliar words conforming to the phonotactics of a language unknown to participants and different enough from English could be added to examine whether the phonological loop would be further taxed in this context, and whether the two learning paradigms might be differentially sensitive to the unfamiliar phonology.

The marginal three-way interaction among word type, learning condition, and nonword repetition in the nonword repetition analysis suggests that a higher nonword repetition score was associated with more successful learning of familiar words compared to nonwords. However, this finding should be interpreted with caution as its significance was marginal and derived from a smaller dataset. Future studies would need to examine nonword repetition as a predictor of novel word learning, ensuring that all data can be retained, and using a task that involves a broader range of difficulty, to better capture its association in learning familiar and unfamiliar words across word learning conditions. Another method to further examine the hypothesis that a higher working memory load might affect the performance of the phonological loop in word learning, especially on unfamiliar words, would be to increase working memory load during teaching cycles for both tasks. This could be achieved with a concurrent task such as including items between learning trials that must be kept in memory for the short-term (akin to the process of a reading span task; Daneman & Carpenter, 1980).

Regarding the observation that learning condition and word type did not interact, it is possible that the similar effect of word familiarity across paradigms emerges in this specific experimental design, where number of words and exposures, and word length and phonology led to learning at high to very high rates (ranging between 69% and 94% across paradigms and word type). The level of task difficulty was calibrated through piloting to avoid ceiling and floor effects that would have obscured effects of learning condition and word type. However, it is possible that if complexity of both tasks were increased, or if CSWL level of ambiguity at exposure was expanded, differences could emerge between learning familiar and unfamiliar words across word-learning paradigms. To examine whether a different pattern of results would be observed during the initial stages of learning, we took advantage of our design that involved three blocks of testing trials. We re-analyzed the data on the first testing block only, hypothesizing that should the interactive effect of learning condition and phonological familiarity be obscured by repeated testing, it would be more likely to be revealed in the first testing block. Indeed, unlike the analyses collapsing across the three testing blocks, analyses of the first testing block data revealed a marginal interaction of learning condition by word type, such that existing words were learned significantly better than nonwords in PAL but only marginally better in CSWL. However, these analyses should be interpreted cautiously, given their marginal nature.

A potential limitation of the present experimental design lies in the fact that we did not include a post-awareness test to examine whether participants became aware that the CSWL task was indeed a word-learning task. Given the repeated-testing design, participants completing the CSWL should have become aware of the nature of the task by the second testing block even if they were not during the first testing block; however, analyses of the first-block data were highly similar to analyses collapsing across the three testing blocks, revealing largely identical effects of phonological working memory across tasks. This may indeed be because participants treated the two tasks similarly, consciously learning the novel words in both PAL and CSWL, despite not being instructed to do so for the CSWL task. Future studies would benefit from attempting to reduce intentional learning during the CSWL task, although our findings indicate that even during the first testing block (when participants' intent to learn and the realization that they will be tested was minimized), phonological working memory was similarly involved in CSWL and PAL.

One additional aspect of our design may have influenced the pattern of findings, specifically, the difference between number of exposures to words and objects within teaching trials across the two paradigms. CSWL requires ambiguity by design: we presented two objects at a time on the screen at learning, which means that participants in this paradigm saw the pictures twice as many times as in PAL, as we kept the number of teaching trials per block at 12 across both paradigms. However, as mean learning performance scores show, performance difference was still clearly found between the paradigms in the expected direction.

Another aspect of the design that could have affected results is time per stimulus at teaching, although we do not know whether it did or not in the present study. Individuals were exposed to stimuli for half the time per trial in CSWL compared to PAL because in CSWL, two words and two objects were presented per trial, but the overall time per trial was the same as in PAL. Further research is warranted to understand how time per trial could affect learning in PAL versus CSWL. On the one hand, a longer time at exposure could support more detailed encoding of the information in memory such that recognition is enhanced at immediate testing. However, it could be that longer times during exposure trials lead to increased forgetting rates within trials. The effect could be different across PAL and CSWL, however, considering that in PAL, participants are explicitly instructed that they need to learn, which is not the case in CSWL, and any difference in "forgetting" rates might be obscured by different encoding rates. We note that in Mulak et al. (2019), while presentation of each pairing was kept constant across learning condition, time per trial increased as the ambiguity at learning increased. However, learning rates in the same condition as in the present study (2×2) were at an average of 50%, which is lower than in the present study, but not drastically so considering the stimuli in Mulak et al. (2019) had the additional difficulty of phonological overlap.

To conclude, paired-associate and cross-situational word-learning paradigms have rarely been examined in parallel, which limits the understanding of the processes underlying these paradigms. Notably, word familiarity and phonological working memory have been crucial to understanding the mechanisms that underlie word-learning performance in paired-associate paradigms, but have been rarely considered in the CSWL literature. In this study, we tested monolingual adults on PAL and CSWL, varying word familiarity and measuring the extent to which phonological working memory might associate with word-learning performance. Findings suggest that word learning is easier in PAL than in CSWL, and that the role of word familiarity and higher levels of phonological working memory are both associated with better learning performance across both paradigms. This means that the roles of working memory and long-term memory in supporting novel word learning is similar across the two paradigms. The overall pattern of results observed in this study indicates that despite drastically distinct theoretical approaches that have been applied to PAL and CSWL, the two paradigms are remarkably similar in the memory systems that they rely on.

References

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.

Aslin, R. N. (2017). Statistical learning: A powerful mechanism that operates by mere exposure. *Wiley Interdisciplinary Reviews: Cognitive Science, 8*(1–2), e1373.

Atkins, P. W., & Baddeley, A. D. (1998). Working memory and distributed vocabulary learning. *Applied Psycholinguistics, 19*(4), 537–552.

Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience, 4*(10), 829–839.

Baddeley, A. D., & Hitch, G. (1974). Working memory. The psychology of learning and motivation. In G. A. Bower (Ed.), *Recent advances in learning and motivation* (pp. 47–89). Academic.

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

Baddeley, A. D., Gathercole, S. E., & Papagno, C. (2017). The phonological loop as a language learning device. In A. Baddeley (Ed.), *Exploring working memory. Selected works of Alan Baddeley* (pp. 164–198). Taylor & Francis.

Baddeley, A., Papagno, C., & Vallar, G. (1988). When long-term learning depends on short-term storage. *Journal of Memory and Language, 27*(5), 586–595. https://doi.org/10.1016/0749-596X(88)90028-9

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Boisgontier, M. P., & Cheval, B. (2016). The anova to mixed model transition. *Neuroscience & Biobehavioral Reviews, 68*, 1004–1005.

Brauer, M., & Curtin, J. J. (2018). Linear mixed-effects models and the analysis of nonindependent data: A unified framework to analyze categorical and continuous independent variables that vary within-subjects and/or within-items. *Psychological Methods, 23*(3), 389.

Clark, H. H. (1973). The language-as-a-fixed effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior, 12*, 335–359.

Conway, C. M., Bauernschmidt, A., Huang, S. S., & Pisoni, D. B. (2010). Implicit statistical learning in language processing: Word predictability is the key. *Cognition, 114*(3), 356–371.

Cowan, N., Morey, C., & Chen, Z. (2007). The legend of the magical number seven. In S. Della Sala (Ed.), *Tall tales about the mind & brain: Separating fact from fiction* (pp. 45–59). OUP Oxford.

Curtin, J. (2018). lmSupport: Support for linear models. R package version 2.9.13. https://CRAN.R-project.org/package=lmSupport

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19*(4), 450–466.

de Vos, J. F., Schriefers, H., Nivard, M. G., & Lemhöfer, K. (2018). A meta-analysis and meta-regression of incidental second language word learning from spoken input. *Language Learning, 68*(4), 906–941.

Ellis, N. C., & Beaton, A. (1993). Psycholinguistic determinants of foreign language vocabulary learning. *Language Learning, 43*(4), 559–617.

Escudero, P., Mulak, K., & Vlach, H. (2013). Cross-situational statistical learning of phonologically overlapping words. In

Proceedings of the annual meeting of the cognitive science society (Vol. 35, No. 35).

Freedman, M. L., & Martin, R. C. (2001). Dissociable components of short-term memory and their relation to long-term learning. *Cognitive Neuropsychology, 18*(3), 193–226.

Frost, R., Armstrong, B. C., Siegelman, N., & Christiansen, M. H. (2015). Domain generality versus modality specificity: The paradox of statistical learning. *Trends in Cognitive Sciences, 19*(3), 117–125.

Gathercole, S. E., Willis, C., Emslie, H., & Baddeley, A. D. (1991). The influences of number of syllables and wordlikeness on children's repetition of nonwords. *Applied Psycholinguistics, 12*(3), 349–367.

Gathercole, S. E., Willis, C. S., Emslie, H., & Baddeley, A. D. (1992). Phonological memory and vocabulary development during the early school years: A longitudinal study. *Developmental Psychology, 28*(5), 887.

Gathercole, S. E., Service, E., Hitch, G. J., Adams, A. M., & Martin, A. J. (1999). Phonological short-term memory and vocabulary development: Further evidence on the nature of the relationship. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition, 13*(1), 65–77.

Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology, 40*, 177–190.

Graf Estes, K., Evans, J. L., Alibali, M. W., & Saffran, J. R. (2007). Can infants map meaning to newly segmented words? Statistical segmentation and word learning. *Psychological Science, 18*(3), 254–260.

Gupta, P. (2003). Examining the relationship between word learning, nonword repetition, and immediate serial recall in adults. *The Quarterly Journal of Experimental Psychology Section A, 56*(7), 1213–1236.

Gupta, P., Lipinski, J., Abbs, B., Lin, P. H., Aktunc, E., Ludden, D., ... & Newman, R. (2004). Space aliens and nonwords: Stimuli for investigating the learning of novel word-meaning pairs. *Behavior Research Methods, Instruments, & Computers*, *36*(4), 599–603.

Hanten, G., & Martin, R. C. (2001). A developmental phonological short-term memory deficit: A case study. *Brain and Cognition, 45*(2), 164–188.

Hartig, F. (2022). DHARMa: Residual diagnostics for hierarchical (multi-level/mixed) regression models. R package version 0.4.5. http://florianhartig.github.io/DHARMa/

Horst, J. S., & Hout, M. C. (2016). The Novel Object and Unusual Name (NOUN) Database: A collection of novel images for use in experimental research. *Behavior Research Methods, 48*(4), 1393–1409.

Judd, C. M., McClelland, G. H., & Ryan, C. S. (2017). *Data analysis: A model comparison approach to regression, ANOVA, and beyond*. Routledge.

Kachergis, G., Yu, C., & Shiffrin, R. M. (2012). An associative model of adaptive inference for learning word–referent mappings. *Psychonomic Bulletin & Review, 19*(2), 317–324.

Kaufman, A., & Kaufman, N. (2004). *Kaufman brief intelligence test* (2nd ed.). Pearson, Inc.

Kazanas, S. A., Altarriba, J., & O'Brien, E. G. (2020). Paired-associate learning, animacy, and imageability effects in the survival advantage. *Memory & Cognition, 48*(2), 244–255.

Krepel, A., de Bree, E. H., Mulder, E., van de Ven, M., Segers, E., Verhoeven, L., & de Jong, P. F. (2021). Predicting EFL vocabulary, reading, and spelling in English as a foreign language using paired-associate learning. *Learning and Individual Differences, 89*, 102021.

Lado, B. (2017). Aptitude and pedagogical conditions in the early development of a nonprimary language. *Applied Psycholinguistics, 38*(3), 679–701. https://doi.org/10.1017/S0142716416000394

Litt, R. A., Wang, H. C., Sailah, J., Badcock, N. A., & Castles, A. (2019). Paired associate learning deficits in poor readers: The contribution of phonological input and output processes. *Quarterly Journal of Experimental Psychology, 72*(3), 616–633.

Magro, L. O., Attout, L., Majerus, S., & Szmalec, A. (2018). Short- and long-term memory determinants of novel word form learning. *Cognitive Development, 47*, 146–157.

Majerus, S., Van der Linden, M., Mulder, L., Meulemans, T., & Peters, F. (2004). Verbal short-term memory reflects the sublexical organization of the phonological language network: Evidence from an incidental phonotactic learning paradigm. *Journal of Memory and Language, 51*(2), 297–306.

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research, 50*(4), 940–967.

Marian, V., Bartolotti, J., Chabal, S., & Shook, A. (2012). CLEARPOND: Cross-linguistic easy-access resource for phonological and orthographic neighborhood densities. *PLoS ONE, 7*(8), e43230. https://doi.org/10.1371/journal.pone.0043230

Martin, N., & Saffran, E. M. (1999). Effects of word processing and short-term memory deficits on verbal learning: Evidence from aphasia. *International Journal of Psychology, 34*(5–6), 339–346.

Mather, N., & Woodcock, R. W. (2001). *Examiner's manual: Woodcock-Johnson III tests of achievement*. Riverside.

Mirman, D., Magnuson, J. S., Estes, K. G., & Dixon, J. A. (2008). The link between statistical segmentation and word learning in adults. *Cognition, 108*(1), 271–280.

Mulak, K. E., Vlach, H. A., & Escudero, P. (2019). Cross-situational learning of phonologically overlapping words across degrees of ambiguity. *Cognitive Science, 43*(5), e12731.

Palan, S., & Schitter, C. (2018). Prolific.ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance, 17*, 22–27.

Papagno, C., & Vallar, G. (1992). Phonological short-term memory and the learning of novel words: The effect of phonological similarity and item length. *The Quarterly Journal of Experimental Psychology Section A, 44*(1), 47–67.

Papagno, C., Valentine, T., & Baddeley, A. (1991). Phonological short-term memory and foreign-language vocabulary learning. *Journal of Memory and Language, 30*(3), 331–347.

R Core Team. (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing. https://www.R-project.org/

Rothkopf, E. Z. (1957). A measure of stimulus similarity and errors in some paired-associate learning tasks. *Journal of Experimental Psychology, 53*(2), 94.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996a). Statistical learning by 8-month-old infants. *Science, 274*(5294), 1926–1928.

Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996b). Word segmentation: The role of distributional cues. *Journal of Memory and Language, 35*(4), 606–621.

Service, E. (1992). Phonology, working memory, and foreign-language learning. *The Quarterly Journal of Experimental Psychology Section A, 45*(1), 21–50.

Service, E., & Craik, F. I. M. (1993). Differences between young and older adults in learning a foreign vocabulary. *Journal of Memory and Language, 32*(5), 608–623.

Smith, L., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition, 106*(3), 1558–1568.

Smith, L. B., & Yu, C. (2013). Visual attention is not enough: Individual differences in statistical word-referent learning in infants. *Language Learning and Development, 9*(1), 25–49.

Smith, L. B., Colunga, E., & Yoshida, H. (2010). Knowledge as process: Contextually cued attention and early word learning. *Cognitive Science, 34*(7), 1287–1314.

Speciale, G., Ellis, N. C., & Bywater, T. (2004). Phonological sequence learning and short-term store capacity determine second language vocabulary acquisition. *Applied Psycholinguistics, 25*(2), 293–321.

Steinel, M. P., Hulstijn, J. H., & Steinel, W. (2007). Second language idiom learning in a paired-associate paradigm: Effects of direction of learning, direction of testing, idiom imageability, and idiom transparency. *Studies in Second Language Acquisition, 29*(3), 449–484.

Suanda, S. H., & Namy, L. L. (2012). Detailed behavioral analysis as a window into cross-situational word learning. *Cognitive Science, 36*(3), 545–559.

Trueswell, J. C., Medina, T. N., Hafri, A., & Gleitman, L. R. (2013). Propose but verify: Fast mapping meets cross-situational word learning. *Cognitive Psychology, 66*(1), 126–156.

van den Noort, M., Bosch, P., & Hugdahl, K. (2006). Foreign language proficiency and working memory capacity. *European Psychologist, 11*, 289–296.

Vlach, H. A., & Sandhofer, C. M. (2014). Retrieval dynamics and retention in cross-situational statistical word learning. *Cognitive Science, 38*(4), 757–774.

Vlach, H. A., & DeBrock, C. A. (2019). Statistics learned are statistics forgotten: Children's retention and retrieval of cross-situational word learning. Journal of Experimental Psychology: Learning, Memory, and Cognition, 45(4), 700

Warmington, M. A., Kandru-Pothineni, S., & Hitch, G. J. (2019). Novel-word learning, executive control and working memory: A bilingual advantage. *Bilingualism: Language and Cognition, 22*(4), 763–782.

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale—3rd Edition (WAIS-3 R)*. Harcourt Assessment.

Ylinen, S., Nora, A., & Service, E. (2020). Better phonological short-term memory is linked to improved cortical memory representations for word forms and better word learning. *Frontiers in Human Neuroscience, 14*, 209.

Yu, C., & Smith, L. B. (2007). Rapid word learning under uncertainty via cross-situational statistics. *Psychological Science, 18*(5), 414–420.

Yu, C., & Smith, L. B. (2011). What you learn is what you see: Using eye movements to study infant cross-situational word learning. *Developmental Science, 14*(2), 165–180.

Yurovsky, D., & Frank, M. C. (2015). An integrative account of constraints on cross-situational learning. *Cognition, 145*, 53–62.