Article:

Henderson, L, Devine, K, Weighall, A orcid.org/0000-0002-6736-287X et al. (1 more author) (2015) When the Daffodat Flew to the Intergalactic Zoo: Off-Line Consolidation Is Critical for Word Learning From Stories. Developmental Psychology, 51 (3). pp. 406-417. ISSN 0012-1649

https://doi.org/10.1037/a0038786

When the daffodat flew to the intergalactic zoo: Off-line consolidation is critical for word learning from stories.

Lisa Henderson[1], Katy Devine[1], Anna Weighall[2], Gareth Gaskell[1]

[1] University of York

[2] University of Leeds

Corresponding author: Lisa Henderson

Contact Details: Department of Psychology, University of York, York, YO 10 5DD

Abstract

Previous studies using direct forms of vocabulary instruction have shown that newly learnt words are integrated with existing lexical knowledge only *after* off-line consolidation (as measured by competition between new and existing words during spoken word recognition). However, the bulk of vocabulary acquisition during childhood occurs through incidental exposure to verbal material; hence, the role of consolidation may be different or limited when learning is less explicit. To address this, 40 children (aged 7-10 years) and 33 adults listened to a fictitious story that contained 12 novel words (e.g., "*daffodat"*). Lexical integration was measured by comparing pause detection latencies to existing competitors (e.g., "*daffo_dil"*) and control words for which no new competitor had been encountered. Pause detection latencies were slower for existing competitors than control words (signifying increased lexical competition) 24 hours after exposure to the novel words but not immediately. Both groups recalled significantly more novel words when tested 24 hours after hearing the story than immediately. Importantly, children with better expressive vocabulary knowledge showed larger consolidation effects for the novel words, both in terms of strengthening of explicit knowledge and their integration with existing knowledge. Off-line consolidation is therefore required for the integration of new and established knowledge when words are learned under relatively naturalistic conditions. Furthermore, a richer body of established vocabulary knowledge may facilitate (or benefit from) swift lexical integration of new vocabulary.

Keywords: Vocabulary knowledge, implicit learning, dual systems framework, lexical integration, vocabulary acquisition

Rapid vocabulary growth during childhood is crucial for academic success and is closely tied to literacy development (Cunningham & Stanovich, 1997; Hoff, 2003; Joshi, 2005; Keenan et al., 2006; Nation & Angell, 2006; Nation & Snowling, 2004). Hence, it is important that we understand how best to maximise word learning opportunities during the school years by examining the process by which children acquire new spoken words. Most estimates suggest that children aged 5-6 years have a working vocabulary of approximately 2500-5000 words (Beck & McKeown, 1991) and during early school years children learn about 3000 words per year – roughly 8 words per day (Baumann & Kameenui, 1991; Beck & McKeown, 1991; Graves, 1986). Estimates suggest that only ~10% of words acquired in a year are learned through direct instruction by adults, with the majority of words learned incidentally through more implicit means, including conversation, television, and, in particular, story exposure (Aktar, 2004; Alloway, Williams, Jones, & Cochrane, 2013; Elley, 1989; Biemiller, 2003; Nagy & Herman, 1987). Indeed, the most commonly observed effect of early story exposure is on children's vocabulary knowledge during the preschool (Hamilton, 2014; Justice, Meier & Walpole, 2005; Senechal & Cornell, 1993; Waisk & Bond, 2001; Walsh & Blewitt, 2006) and primary school years (Dickinson, 1984; Elley, 1989; Nagy, Anderson, & Herman, 1987; Penno, Wilkinson & Moore, 2002; Robbins & Ehri, 1994). This study examines how children and adults learn new spoken words that they encounter whilst listening to stories.

Spoken word recognition is often characterized in terms of a pre-lexical level, in which phonemes and/or lower level items such as phonetic features are processed, and a lexical level, the resulting representation that corresponds to the word form (McClelland & Elman, 1986; Norris, 1994). In addition, associated word meaning is represented at a semantic level (Gupta & MacWhinney, 1997) and the linkage between these levels of information permits rapid, flexible word recognition (Perfetti & Hart, 2002). Consequently, when an individual hears a familiar word, the speech signal maps onto a pre-lexical representation, and activation spreads to the lexical and semantic levels, culminating with word recognition. Upon hearing an unfamiliar word, pre-lexical activation occurs, but there is nothing at the lexical level that corresponds to the signal. Hence, for word learning to take place, a new lexical representation must be established. This study is primarily concerned with examining the time course by which new *lexical representations* emerge following spoken word learning through story encounters.

Previous findings suggest that school-aged children's ability to recognise new word forms learned through listening to stories persists (Dickinson, 1984; Elley, 1989) or increases (Senechal & Cornell, 1993; Wilkinson & Houston-Price, 2013) over time. Moreover, new word form knowledge gained from story exposure is enhanced during daytime naps in preschool children (Williams & Horst, 2014). These findings align with the view that word learning is a prolonged process that depends upon off-line consolidation, particularly during sleep (Dumay & Gaskell, 2007; Henderson, Weighall, Brown & Gaskell, 2012; Tamminen et al., 2010). According to a dual-systems account of vocabulary acquisition (Davis & Gaskell, 2009), sparse representations of new words are initially set up in the hippocampus, but over time, particularly during sleep, a long-term representation is strengthened in neocortical memory (for neuroimaging evidence supporting this account see Davis, Di Betta, MacDonald & Gaskell, 2009). The coordination of a short- and longer-term system is proposed to accommodate new memories quickly and protect existing memories from damage.

Using the dual-systems account as a theoretical framework, Gaskell and colleagues have examined the extent to which off-line consolidation is important for the *integration* of new and existing lexical knowledge. Arguably, typical tests of whether a new word form has been learnt (e.g., recall, recognition) can only provide an index of explicit episodic knowledge and do not directly address whether the new information has been incorporated within lexical networks and competes for recognition with existing lexical entries (Gaskell & Dumay, 2003; Leach & Samuel, 2007). Lexical competition is a key component of many models of word recognition that is proposed to allow for fast and efficient retrieval of stored lexical information (Gaskell & Marslen-Wilson, 2002; Grainger & Jacobs, 1996; Marslen-Wilson, 1989; McClelland & Elman, 1986; Norris, 1994). Hence, lexical integration has been measured by the strength of lexical competition between a novel word (e.g., "*cathedruke*") and an existing competitor (e.g., "*cathedral*") (Gaskell & Dumay, 2003; Dumay & Gaskell, 2007; Tamminen & Gaskell, 2008). Findings suggest that lexical competition effects (e.g., slowed responses to "*cathedral*" after learning "*cathedruke*") do not emerge immediately after ~30 minutes of training on a set of novel competitors, but instead typically emerge after a period of sleep, presumably via off-line consolidation.

A similar delayed emergence of lexical integration has been reported when children aged 7-12 years learn spoken novel competitors (Brown, Weighall, Henderson & Gaskell, 2012; Henderson, Weighall, Brown & Gaskell, 2012; Henderson, Weighall, Brown & Gaskell, 2013a). Furthermore, lexical competition effects between new real words associated with the science curriculum (e.g., "*hippocampus*") and existing competitors (e.g., "*hippopotamus*") emerged in 5-9-year-olds after a 24-hr delay even when the new words were paired with a meaningful picture and definition during training (e.g., "*A hippocampus* *is a part of your brain that helps you remember things*") (Henderson et al., 2013b). This suggests that the delay in lexical integration is not a consequence of learning meaningless or fictitious words (see also Takashima, Bakker, van Hell, Janzen & McQueen, 2014).

In the paradigm developed by Gaskell and colleagues participants are typically trained on novel words using phoneme monitoring and phoneme segmentation tasks in which they are provided with numerous, explicit exposures to the word forms. This has been the case even in studies that have trained new words in meaningful contexts (e.g., Henderson et al., 2013b). Szmalec et al (2012) adopted a more implicit "Hebb" training paradigm in which adults listened to and recalled sequenced strings that were presented multiple times over a training period, such as "*sa-fa-ra*". These repeated sequences were embedded in random lists of syllables. Lexical competition for existing competitors (e.g., "*safari*") was not observed immediately after this more implicit form of training, but emerged after 12 hours, regardless of whether sleep had occurred. This suggests that with a more implicit style of training the integration of newly acquired items might not be sleep-dependent, although again in this study the integration effect was not present immediately.

A third pattern of emergence of lexical competition was found by Fernandes, Kolinsky and Ventura (2009), who used an artificial language learning paradigm (Saffran, Aslin & Newport, 1996) to examine the influence of implicit exposure to novel competitor words on adults' auditory lexical decision responses to real word neighbours. Adults listened to a continuous stream of artificial speech consisting of concatenated syllables, in which cues for word boundaries could be extracted from statistical information contained within the speech stream. Inhibitory effects (i.e., slower lexical decision responses to real word competitors compared to control words) were observed immediately after exposure to the speech stream, and these effects remained one week later, suggesting that new words acquired through sensitivity to statistical segmentation cues are quickly integrated with existing lexical knowledge in adults. This finding raises the question of whether consolidation is necessary for lexical integration when more implicit forms of training are used (cf. Nemeth et al., 2010).

The primary question addressed here is whether lexical integration occurs during learning (as evidenced by lexical competition immediately after exposure) or after a period of off-line consolidation in both children and adults when new words are learned in a more naturalistic situation (i.e., through listening to stories) that does not rely solely on direct instruction. Although previous studies have supported a role for consolidation when children learn new words from listening to stories (e.g., Senechal & Cornell, 1993; Wilkinson & Houston-Price, 2013; Williams & Horst, 2014) they have relied upon explicit measures of new word knowledge (e.g., recognition and recall).

One hypothesis could be that encountering new words in meaningful stories (in which the words are encountered in multiple sentential contexts) may facilitate word learning, and speed up the process of lexical integration. Previous studies (e.g., Wilkinson & Houston-Price, 2013) have used spreading activation models (e.g., Collins & Loftus, 1975) to account for how children capitalise on word learning opportunities during story exposure. Such models propose that words are represented in the lexicon as networks of related concepts. Each time a familiar word is encountered activation spreads through this network, culminating in the activation of the word’s meaning as well as activation of interrelated word meanings. Encountering a new word within a story in varying sentential contexts may therefore enable more immediate connections to be formed with related concepts (Carey, 1978; Carlo et al., 2004; Mol et al., 2009; although see Horst, Parsons & Bryan, 2011; Wilkinson & Houston-Price, 2013). This could work to strengthen the mapping between the new spoken word form and its meaning and in this way facilitate lexical integration. As discussed above, Henderson et al (2013b) observed lexical integration effects only after a 24-hour delay and not immediately, even when new words were embedded into a defining sentence during training. However, the words were trained using explicit phonics-based tasks and were paired with only a single sentence, which differs considerably to encountering new words incidentally in varying sentential contexts within a story.

A second key question is whether children's ability to consolidate (and integrate) new word forms from story exposure is linked to their existing corpus of vocabulary knowledge. Previous research converges on the view that children with superior vocabulary knowledge are more likely to learn new words when listening to stories than children with poor vocabulary knowledge (e.g., Dockrell, Braisby & Best, 2007; Ewers & Brownson, 1999; Joshi, 2005; Penno et al., 2002; Reese & Cox, 1999; Robbins & Ehri, 1994; Senechal, 1997; Wilkinson & Houston-Price, 2013). However, this so-called "Matthew Effect" (Stanovich, 1986) has not been consistently replicated (Elley, 1989; Walsh & Blewitt, 2006). Furthermore, previous studies have focused on the relationship between established vocabulary knowledge and children's recognition of new words from stories, rather than whether established vocabulary impacts on the consolidation and/or integration of new vocabulary. A number of studies suggest that children with smaller vocabularies are less sensitive to part-word probability (i.e., the extent to which a word contains sound sequences that overlap with other words; Metsala & Walley, 1998; Storkell & Hoover, 2011) and that low levels of oral language are associated with differences in lexical competition during spoken word recognition (Nation, 2014). Hence, it may be hypothesised that children's existing vocabulary knowledge will be associated with the emergence of lexical integration (as indexed by lexical competition). The nature of this association is most likely to be reciprocal: An existing richer network of vocabulary may permit new words to be more easily integrated with the lexicon, but in addition, superior lexical integration may allow for richer vocabulary growth.

The present study

Children and adults listened to a fictitious story that contained 12 novel nonword competitors (e.g., "*daffodat*", a new competitor for "*daffodil*") that occurred five times in varying sentential contexts. Participants' ability to recall and recognise the new words and integrate them with existing knowledge was tested immediately after hearing the story and 24 hours later. Arguably, the presence of the immediate test may contribute to the emergence of lexical competition and any improvements in explicit memory at the 24 hour test; however, previous studies using direct vocabulary instruction have suggested that lexical competition effects are equivalent after 24 hours, regardless of whether repeat testing has occurred (e.g., Henderson et al., 2013b).

The pause detection task (Mattys & Clark, 2002) was used to measure lexical integration (following Dumay & Gaskell, 2007; Henderson et al., 2012, 2013a, 2013b). Short 200 ms pauses are inserted into the existing neighbours at the point of deviation from new competitors (e.g., "*daffo_dil*") as well as in a set of matched control words. Participants decide whether a pause is present or absent as quickly as possible. Pause detection latencies are faster when a single lexical candidate has been isolated by the time the pause is encountered (e.g., "*cathe_dral*") than when there are several alternatives for completion (e.g., "*cabb_age/in/inet*"). When an existing word that previously had no competitors (e.g., *"daffodil"*) acquires a new competitor (e.g., *"daffodat"*), pause detection latencies to the existing word should slow down, but only once the new competitor has been integrated with existing lexical networks. This task is arguably more sensitive to lexical competition than alternatives such as lexical decision (Henderson et al., 2013a); it provides an 'on-line' measure of lexical activity as speech is unfolding in real time and does not require a linguistic judgment which may decrease task demands and increase strategic processing.

This study addresses two main hypotheses: (1) In line with the dual-systems account (Davis & Gaskell, 2009) the strengthening and integration of new words learned from storybook encounters should require a period of off-line consolidation. Alternatively, it is also possible that less reliance upon explicit exposure and/or encountering new words in varying sentential contexts could work to speed up the process of lexical integration (e.g., Fernandes et al., 2009). (2) Based on previous findings of a Matthew Effect in word learning from stories (e.g., Stanovich, 1986; Wilkinson & Houston-Price, 2013), existing vocabulary knowledge should be associated with the extent to which explicit memory for newly learned words is strengthened over time, and with overnight changes in lexical integration.

Method

Participants

A total of 73 participants took part: 40 children aged 7.06-10.60 years (18 males; mean age 8.71 years, SD 1.10 years) and 33 adults aged 18-30 years (13 males; mean age 21.39 years, SD 2.15 years), allowing us to examine developmental changes in the time course of lexical integration. Children were recruited from primary schools situated in areas representing a range of socioeconomic backgrounds. The age range of 7-10 years was selected to facilitate comparison with previous findings (Brown et al., 2012; Henderson et al., 2012; Henderson et al., 2013a). Parental consent was obtained for all children. Adults were recruited from the University of York and provided written consent. Adult participants and parents of child participants confirmed an absence of diagnosed learning or neurological disabilities, that they had normal or corrected to normal vision and hearing, and were native English monolingual speakers. Both groups showed a normal distribution of ability on standardised tests of verbal ability from the *Wechsler Abbreviated Scales of Intelligence* (children's mean T score on Vocabulary subtest 56.25, SD = 9.71, 33-75; adults mean T score 63.54, SD = 6.04, 66-80) and nonverbal ability (children's mean T score on Matrix Reasoning 53.35, SD = 8.20, 32-69; adults mean T score 57.88, SD = 6.11, 48-75) (all Kolmogorov-Smirnov Z scores were >.75, *p*s > .57). The Vocabulary subtest captured existing expressive vocabulary knowledge (or vocabulary 'depth') and required participants to produce definitions of words that increased in difficulty.

Stimuli

Novel words. Participants were randomly allocated to one of two lists of 12 stimulus triplets comprising a "base" (known) word (e.g., *daffodil*), a novel competitor (e.g., *daffodat*) and a novel foil for the 2AFC task (e.g., *daffodan*) (see S1). Two lists (List 1, List 2) were necessary to allow the base words from the unlearned list to act as control words in the pause detection task. The base words were picturable nouns that were deemed to be familiar to the age range (see Henderson et al., 2012). They came from a range of semantic categories (e.g., animals, food, plants, clothing, ornaments, and landmarks) that were equally distributed across lists. The initial fragments of the base words (e.g., "*daffo_*") had no other possible completions. The base words in each list were matched on number of syllables (List 1 mean 2.42, SD 0.52; List 2 mean 2.33, SD 0.49) and phonemes (List 1 mean 6, SD 0.73; List 2 mean 6.33, SD 0.89), spoken length (List 1 mean 812ms, SD 71ms; List 2 mean 835ms, SD 87ms), initial word fragment length (List 1 mean 275ms, SD 37ms; List 2 mean 271ms, SD 36ms), and frequency (Children's Printed Word Database: List 1 mean 21.75, SD 37.22; List 2 mean 35.75, SD 59.54; Celex Frequency: List 1 mean 4, SD 3.30; List 2 mean 4.5, SD 4.47): All p values >.35.

Story. The novel words from each list were embedded into a story written by the experimenters, entitled "A Day at the Intergalactic Zoo" (see S2). Two versions of the story were recorded; one containing the novel words from List 1 and the other containing the novel words from List 2. Every effort was made to record the stories using the same pitch, intonation, and prosody, and they were matched for spoken length (~6 minutes long, 1016 words). The recorded story was presented to participants via headphones. Each novel word occurred five times at distributed positions throughout the story and no novel word occurred in more than two consecutive sentences. None of the words used in the story had an age of acquisition that exceeded 7 years (Kuperman et al., 2012) to ensure that the content was familiar to the age range. The novel words were embedded into sentences such that inferences had to be made in order to understand their meaning (e.g., "*Chop chop," called Mum, "*G*rab your dolpheg in case you get cold*"). The novel words had similar meanings to familiar words but differed via one or two novel features. For example, a "*daffodat*" is (1) a flying space vehicle, (2) has seatbelts, and (3) runs on stardust. Each novel word had three key semantic features (see S3). Five teachers who taught children of the target age range unanimously agreed that the story was gender neutral and age appropriate in terms of interest and language content.

Procedure

Adults were tested in a lab at the University of York; children were tested in a quiet room in their schools. Participants were exposed to the novel words in the story. Immediately following this, measures of lexical integration, cued recall and novel word recognition were administered in that fixed order (Henderson et al., 2012, 2013a, 2013b). Participants were retested on these measures after 24 hours. At the end of the 24-hour session they were asked to define each novel word and their familiarity with the existing base words was assessed. All tasks were delivered via DMDX (Forster & Forster, 2003).

Exposure to the novel words: Participants were told that they were going to listen to a story about an "alien family who live in space and are having a day trip to an alien zoo". They were told that "there may be some words that you have not heard before but please continue to listen and avoid asking questions so that you do not miss any of the story". They were first presented with each novel word (from List 1 or List 2) once via headphones and asked to repeat it aloud. This procedure was adopted as a consequence of a pilot experiment, where children's attention to the spoken story was disrupted when they heard the new words on their first occurrence. Communication with classroom teachers (who also provided feedback on the story prior to the experiment) confirmed that alerting children to new key words prior to using the words in context is common practice in the classroom. Although this pre-exposure arguably acts as a form of direct instruction, the bulk of the training relied upon incidental exposure to the novel words in the story in contrast to previous studies (e.g., Dumay & Gaskell, 2007; Henderson et al., 2012; 2013a, 2013b).

Importantly, it is unlikely that a single explicit exposure to a novel word could by itself support subsequent lexical integration (Gaskell & Dumay, 2003); thus, any lexical integration that emerges can largely be attributed to the learning that takes place whilst listening to the story. Following this single exposure, participants listened to the fictitious story (containing words from List 1 or List 2) and were not given opportunities to replay the story.

Lexical integration task. A pause detection task then measured changes in lexical activity after exposure to the novel word competitors (Mattys & Clark, 2002). Participants decided whether a pause was present or absent for each spoken stimulus (by pressing one of two buttons as quickly and as accurately as possible). Stimuli comprised 24 base words: 12 for which a novel word competitor had been taught (competitor condition) and 12 for which no competitor had been taught (control condition). Twenty-four fillers were also included. Half of the words in the competitor and control conditions and half of the fillers contained a 200ms pause. Four versions of the task were counterbalanced across participants so that each item was equally represented in the four cells of the design (competitor, pause present; competitor, pause absent; control, pause present; control, pause absent; cf. Dumay & Gaskell, 2007). Pauses were positioned before the onset of the final syllable ("*daffo_dil*"). RT was measured from pause onset. Item order was randomised for each participant.

Explicit memory tasks. Three measures of explicit word knowledge were administered. In a *cued recall task*, participants heard the first syllable (e.g., "*daff*") of the 12 novel words and were instructed to complete the cue using one of the words they heard in the story. A *2AFC task* was administered to measure *novel word recognition.* Participants heard the novel word (e.g., "*daffodat*") and the accompanying novel foil (e.g., "*daffodan*") and pressed button 1 if they thought the first word was the one they heard in the story or button 2 if they thought the second word was the one they heard in the story. Order of the novel words and foils was counterbalanced across participants. Accuracy was recorded. These tasks were administered at the immediate and 24-hr tests. At the end of the 24-hour session participants were asked to define the meaning of each novel word (*definitions task)*, to ascertain the extent to which participants had acquired information about the meanings of the novel words. The following instruction was given: "I am going to say each of the new words aloud and I would like you to tell me what each word means. Imagine you are describing the word to someone who has never heard of it before". If participants provided a single word response that was correct (e.g., "spaceship" for "*daffodat*"), they were asked "Can you tell me anything else about a/an xxx?". Each definition was scored out of 3 (max score = 36): Participants received one point for each defining feature (see S3).

Control tasks*.* After hearing the story participants answered six questions to check they had comprehended the story (*comprehension task,* S4). Each question was scored a maximum of two points (total possible score 12). Adults scored better (mean=7.91, SD=2.40, 4-12) than children (mean=5.10, SD=2.30, 0-8), $F(1, 71)=25.95$, $p<.001$, but both groups scored a mean of ~1 per question, suggesting that on average, participants correctly attempted each question. There was no significant difference between Lists, $F(1,71)=0.75$, $p>.05$.

To ensure that participants were familiar with the base words (e.g., "*daffodil*"), a *picture-word matching task* was administered at the end of the final session of the experiment (after the definitions task). For each trial, one target picture (e.g., of a daffodil) and three distracters (2 other trained pictures and 1 untrained distracter) were displayed in separate quadrants on the screen. A base word was played through headphones and the participant pointed to the matching picture. Untrained distracters were matched on age of acquisition to the base words (MRC Psycholinguistic Database, Wilson, 1988). Trial order was randomised but the same distracter images always occurred with the same target for each participant and the position of these four images on screen remained constant. Target pictures were equally distributed across quadrants. Adults were familiar with all base words; children scored a mean 98.90% correct (SD 3.34%). Any unfamiliar items were removed on a participant-by-participant basis from the pause detection analysis.

Results

Lexical integration. The RT and error data (combined across incorrect and missed responses) for the pause detection task are shown in Table 1. Outliers were removed if RTs were more than 2 SDs from the condition mean for each participant separately (a mean of 4.07% trials for adults and 3.96% trials for children, across both sessions). RT data were analysed for correct responses only. RT and error data were combined for pause present and pause absent trials (Henderson et al., 2012; Henderson et al., 2013a, 2013b).

The RT data were entered into a 2 (Condition: Competitor, control) x 2 (Session, 0 hr, 24 hr) x 2 (Group: Children, adults) mixed-design ANOVA (see S5 for analyses with List as an additional variable). There were significant main effects of Session, $F(1, 71)=10.52$, $p<.01$, $\eta_p^2= .13$, Condition, $F(1, 71)=12.10$, $p= .01$, $\eta_p^2=.15$, and Group, $F(1, 71)=28.33$, $p<.001$, $\eta_p^2=.29$: Responses were faster for the 24 hr test than the 0 hr test, for the control than competitor conditions, and for adults than children. Most importantly, there was a significant Session x Condition interaction, $F(1, 71)=10.18$, $p<.01$, $\eta_p^2=.13$: Pause detection latencies were similar for competitor and control conditions at 0-hrs (control mean 943ms, SD = 312ms, competitor mean 949ms, SD = 324ms, mean difference 5.80ms, 95% CIs -32-43ms, $t(72)=0.31$, $p>.05$, $d=.07$), indicating that the acquisition of the novel competitor had not influenced lexical competition at this time point. However, at 24-hrs, participants were slower to respond to the competitor words than to control words (control mean 824ms, SD=215ms, competitor mean 920ms, SD=307ms, mean difference 96ms, 95% CIs 55-136ms, $t(72)=4.73$, $p<.001$, $d=1.11$) suggesting that the competitor effect was present. The three-way interaction between Condition, Session and Age was not significant, indicating that the competitor effect at 24-hrs (but not 0-hrs) was obtained for children (0 hr, mean difference 16ms, SD = 201ms, 95%CIs -48-80ms, $t(39)=0.51$,$p>.05$, $d=0.16$; 24 hr, mean difference 134ms, SD=216ms, 95% CIs 65-203ms, $t(39)=3.92$,$p<.001$, $d=1.26$) and adults (0 hr, mean difference -7ms, SD = 93ms, 95%CIs -39-26ms, $t(32)=-0.41$,$p>.05$, $d=0.14$; 24 hr, mean difference 49ms, SD=81ms, 95% CIs 21-78ms, $t(32)=3.51$,$p<.01$, $d=1.24$). There was a marginal Condition x Group interaction, $F(1, 71) = 3.75$, $p = .057$, $\eta_p^2= .05$, indicative of a larger overall difference between control and competitor conditions for children than for adults; however, the same ANOVA conducted on individuals' standardised RT scores (i.e., z scores) suggests that this interaction occurred as a result of globally slower pause detection RTs for the children, leaving larger room for a difference between competitor and control conditions. In this z-score analysis, the main effects of Condition, $F(1, 71)=13.66$, $p<.001$, $\eta_p^2=.16$, and Session, $F(1, 71)=12.10$, $p<.01$, $\eta_p^2=.15$, and the crucial Condition x Session interaction remained significant, $F(1, 71)=13.88$, $p<.001$, $\eta_p^2=.16$; neither children or adults showed a significant competition effect at 0-hrs (children, $t(39)=0.74$, $p>.05$; adults, $t(32)=-0.45$, $p>.05$) but both groups showed a significant competition effect at 24-hrs (children, $t(39)=4.25$, $p<.001$; adults, $t(32)=3.34$, $p<.01$). However, there was no main effect of Group, $F(1, 71)=0.89$, $p>.05$, $\eta_p^2=.01$, and although children continued to show a numerically larger competition effect when collapsing across sessions (z-score difference between competitor and control conditions = -.19, SD = 0.34, $t(39)=-3.57$, $p<.001$) than adults (z score difference

= -.10, SD=0.32, $t(32) = -1.72$, $p=.095$), the Condition x Group interaction was not significant, $F(1, 71)=1.52$, $p>.05$, $\eta_p^2=.02$. In sum, as shown in Figure 1, neither group showed a significant competition effect immediately after exposure to the novel words but both groups showed a significant competition effect 24 hours.

The error data (combining pause present and absent trials) are also presented in Table 1. Error rates were very low (i.e., ~1 out of 12 items for children and ~0.5 out of 12 items for adults). Errors were equally distributed across pause present ('yes') and pause absent ('no') responses ($p>.05$) and did not significantly differ at the 0-hr test ($p>.05$) or the 24-hr test ($p>.05$). The error data were entered in a 2 (Condition: Competitor, control) x 2 (Session, 0 hr, 24 hr) x 2 (Group: Children, adults) mixed-design ANOVA. Children made more errors than adults (Group, $F(1, 71)=21.30$, $p<.001$, $\eta_p^2=.23$). There were no other main effects or interactions: Condition, $F(1, 71)=0.28$, $p>.05$, $\eta_p^2=.004$, Session, $F(1, 71)=0.06$, $p>.05$, $\eta_p^2=.001$; Condition x Session, $F(1, 71)=0.33$, $p>.05$, $\eta_p^2=.01$, Condition x Group, $F(1, 71)=0.59$, $p>.05$, $\eta_p^2=.01$, Condition x Session x Group, $F(1, 71)=.03$, $p>.05$, $\eta_p^2=.00$.

Explicit memory. Performance on the tasks of explicit memory is shown in Table 1.

Cued Recall. Performance was poor for both groups; to capture partially correct responses, a response was allocated 2 points if the whole word was correctly recalled or 1 point if it was correct apart from one phoneme (e.g., "*daffodet*", "*daffodit*")[2]. The data were entered in a 2 (Session, 0 hr, 24 hr) x 2 (Group: Children, adults) mixed-design ANOVA (see S5). There was a significant main effect of Session, $F(1, 71)=126.82$, $p<.001$, $\eta_p^2=.64$: Participants recalled significantly more syllables of the novel words at the 24 hr test than at 0 hrs. There was a significant main effect of Group, $F(1, 71) = 65.57$, $p < .001$, $\eta_p^2= .48$, and a significant Session x Group interaction, $F(1, 71) = 8.45$, $p < .01$, $\eta_p^2= .11$: Adults showed larger improvements from the 0hr test to the 24 hr test (mean improvement=5.09 syllables, SD=3.25 syllables, 95% CIs 3.94-6.24, $t(32)=8.99$, $p<.001$, $d=3.18$) than children (mean improvement=3 syllables, SD=2.88 syllables, 95% CIs 2.08-3.92, $t(39)=6.58$, $p<.001$, $d=2.11$).

Novel Word Recognition. All adults achieved an accuracy score that exceeded chance (i.e., 50%). Nine children showed recognition scores at 50% or lower at 0-hrs (mean=48.15%, SD=3.67%, range 41.67-50%); however, all of these children showed improvements at 24-hrs (mean 84.26%, SD = 11.37%, range 83.33-100%), confirming that no child performed at 50% or below for both sessions. Accuracy data were entered in a 2 (Session, 0 hr, 24 hr) x 2 (Group: Children, adults) mixed-design ANOVA (see S5). There was a main effect of Session, $F(1, 71)=8.62$, $p<.01$, $\eta_p^2=.11$: Participants recognised more words at the 24 hr test than at 0 hrs. There was also a significant main effect of Group, $F(1, 71) = 47.25$, $p<.001$, $\eta_p^2 = .40$, and a significant Session x Group interaction, $F(1, 71) = 3.93$, $p<.05$, $\eta_p^2= .05$: Adults showed smaller improvements from the 0hr test to the 24 hr test (mean improvement=2.02%, SD=10.21%, 95% CIs 1.60-5.64%, $t(32)=1.14$, $p>.05$, $d=0.40$) than children (mean=10.42%, SD=22.47%, 95% CIs 3.23-17.60%, $t(39)=2.93$, $p<.01$, $d=0.94$), likely due to a ceiling effect in the adult group. Indeed, 24 out of 33 of the adults in the sample showed accuracy at 92% or higher at the 0hr test (10 adults scored 100% and a further 14 scored 92%).

Definitions. A 2 (Group: Children, adults) x 2 (List, 1, 2) between-subjects ANOVA was performed on the definitions scores (Table 1). Performance was low for both groups, but adults performed significantly better than children ($F(1, 72)=52.17$, $p<.001$, $\eta_p^2=.43$). For the majority of responses, participants did not attempt a definition or provided a vague one-word guess (e.g., “clothing”, “alien”). Even when the one-word responses were correct, participants very rarely provided more information when prompted. Significant positive correlations between story comprehension scores and definitions scores were found for children ($r(40)=.31$, $p<.05$) and adults ($r(33)=.46$, $p<.01$), suggesting that individuals who had retained more meaningful information about the novel words also showed superior immediate comprehension of the story.

Correlations between consolidation effects and existing vocabulary knowledge.

Consolidation effects (i.e., overnight changes in performance) were calculated for lexical competition effects (in ms) and explicit phonological memory by calculating the difference between the competition effect or cued recall scores at 24 hrs and 0 hrs. Two-tailed Pearson's *r* correlations were performed between these scores and raw scores on the standardised test of expressive vocabulary, for children (*n* 40) and adults (*n* 33) separately.

Lexical integration. Overnight changes in lexical competition correlated with existing vocabulary for children ($r=.40$, $p=.01$, and when age was controlled, $r=.33$, $p=.043$) but not adults ($r=-.04$, $p=.81$): Consistent with the hypothesis, children with larger existing vocabularies showed larger gains in lexical competition after 24 hours (Figure 2). Children's lexical competition effect at 0 hrs did not correlate with vocabulary ($r=-.13$, $p=.42$) but children with larger lexical competition effects at 24 hrs exhibited larger existing vocabularies ($r=.43$, $p=.006$). This suggests that it is the level of competition between new and existing phonological knowledge after off-line consolidation that is linked to children's existing vocabulary knowledge, rather than performance on the pause detection task generally.

Explicit phonological memory. Overnight improvements in cued recall positively correlated with expressive vocabulary for children ($r=.36$, $p<.023$, and when age was controlled, $r=.34$, $p=.03$), but not for adults (cued recall, $r=-.05$, $p=.80$) (Figure 2). Children's cued recall performance at 0 hrs did not correlate with existing vocabulary (*r*=.25, *p* = .123) but children with better cued recall at 24 hrs showed larger vocabularies (*r* = .38, *p* = .017). Correlations were not performed between existing vocabulary knowledge and novel word recognition on the 2AFC task due to the ceiling effects in this data (see S6).

These correlations suggest that the consolidation of new phonological forms is associated with a richer network of established vocabulary knowledge during childhood.

Correlations between consolidation effects and novel vocabulary knowledge. Exploratory correlations were performed between overnight changes in lexical competition and explicit phonological memory and scores on the novel word definitions task to examine associations between phonological and semantic aspects of novel word learning (Perfetti & Hart, 2002). There were no correlations between overnight changes in lexical competition and children's or adults' definitions of the novel words at the 24 hour session (children, *r*=.02, *p*=.92; adults, *r*=-.23, *p*=.20). However, for children there was a significant positive correlation between overnight increases in cued recall and their definitions of the novel words (*r*=.41, *p*=.008; when controlling for age, *r*=.41, *p*=.01; adults, *r*=.11, *p*=.56). Children's performance on the novel word definitions task significantly correlated with cued recall performance at 24 hrs (*r*=.43, *p*=.005) but not at 0 hrs (*r*=.29, *p*=.07). Thus, children who were better at recalling the new phonological forms at 24hrs (and showed bigger overnight improvements) were also better at recalling their meanings 24 hrs after hearing the story (S6).

Discussion

This study presents novel evidence that children and adults' memories for words learned through listening to stories are strengthened and integrated with existing knowledge over 24-hours. Consistent with previous research, the majority of children (aged 7-10 years) and adults were able to recognise a substantial proportion of the new words immediately after hearing the story (Dickinson, 1984; Elley, 1989; Penno et al., 2002; Robbins & Ehri, 1994; Senechal & Cornell, 1993; Wilkinson & Houston-Price, 2013), and their explicit phonological memory was enhanced after off-line consolidation (Wilkinson & Houston-Price, 2013; Williams & Horst, 2014). These data advance previous findings by showing that the novel words became integrated with existing lexical knowledge only after off-line consolidation: Lexical competition was not evident shortly after story exposure, but a robust effect was found for both children and adults after 24 hours. This is important because it has been argued that measures of recognition or recall of new words do not measure the full extent of lexical engagement because they cannot address whether or not the new information has been stored within lexical networks (Gaskell & Dumay, 2003; Leach & Samuel, 2007). The data add further weight to the dual-memory systems account of vocabulary acquisition (Davis & Gaskell, 2009) and build on previous studies with children (Brown et al., 2012; Henderson et al., 2012; Henderson et al., 2013a, 2013b) and adults (Dumay & Gaskell, 2007; Gaskell & Dumay, 2008; Tamminen et al., 2008), suggesting that the delayed emergence of lexical integration for new words is *not* a consequence of direct instruction of novel word forms but can also be observed after more implicit learning.

Although the pattern of results confirmed our hypothesis that there would be a competition effect after 24 hrs but not immediately after training, it is important to note that there was substantial variability in the observed lexical competition effects, particularly for children at the 0-hr session. On average children were 16ms slower to respond to competitor words (e.g., “*daffo_dil*”) than control words at the 0-hr session, but the standard deviation was large (201ms) and the difference between conditions ranged from -575ms to 626ms. Further research is needed to understand the factors that underpin this variance.

Nevertheless, the findings contrast with an artificial language learning experiment by Fernandes et al. (2009) who found clear evidence of lexical competition immediately after learning from novel words extracted implicitly from a continuous syllable stream. An interpretion of this result is that some forms of implicit learning can bypass the extended consolidation process observed here (Coutanche & Thompson-Schill, 2014; Sharon et al., 2011). However, there are also several other differences between the two experiments that may have been influential. Fernandes et al incorporated high levels of exposure to the new words during training (i.e., up to 189 exposures in Experiment 1). It is possible that extreme levels of exposure can lead to the formation of a (possibly hippocampal) novel word representation that is sufficiently strong to affect recognition of existing neighbours prior to consolidation/integration. Furthermore, the extended and interleaved nature of the training session used by Fernandes et al may have facilitated lexical integration (cf. Lindsay & Gaskell, 2013). Finally, Fernandes et al used lexical decision as a test of lexical integration. Lexical decision response times do not provide an on-line measure of lexical competition as the speech string is being heard (Goldinger, 1998; Wagenmakers et al., 2004; Gaskell & Dumay, 2003; Henderson et al., 2013a; Szmalec et al., 2012) and hence, lexical decision response times are more influenced by strategic decision processes (e.g., awareness of the overlap between new and existing words) which could account for the immediate competition effect. Szmalec et al (2012) used an implicit word learning procedure (i.e., Hebb learning) with pause detection as a more 'on-line' measure of lexical integration during speech recognition, and similar to this study, revealed that lexical competition between new and existing words emerged only 12 hours after training and not immediately.

In comparison to previous studies that have used more direct forms of instruction and incorporated greater numbers of exposures (e.g., Brown et al., 2012; Henderson et al., 2012; 2013a; 2013b; Dumay & Gaskell, 2007), the levels of explicit phonological and semantic memory for the novel words were low. Not surprisingly, participants' ability to recognise the novel words was typically better than their ability to recall and produce the novel words (Brown et al., 2012; Henderson et al., 2012, 2013a; Dumay & Gaskell, 2007; McMurray et al., 2012); however, children only recalled 11% of novel words and adults recalled only 29% of novel words immediately after training. Children and adults of a similar age range in Henderson et al (2013a) showed mean cued recall rates of 18% and 61% immediately after training, respectively. The relatively low recall rate in this study is not surprising. These figures align with Swanborn and de Glopper (1999), who present a meta-analysis concluding that only 15% of unknown words are learned from story exposure (where 'learning' is measured by performance on explicit tasks of memory) (see also Biemiller & Boote, 2006). Dockrell et al (2007) also obtained low rates of learning when unfamiliar science words were presented incidentally to children in video clips. Thus, although children clearly learn a great deal of their vocabulary implicitly through activities such as story exposure, the volume of words acquired from a single occasion will be less than might be acquired from a more direct training session. Low levels of recall may suggest that too many novel words were incorporated into the story or that there were too few exposures. Further research is needed to understand the optimal conditions by which children can consolidate new vocabulary from story exposure (Horst, 2013).

Given these relatively poor levels of novel word recall, it is striking that adults and particularly children showed statistically robust lexical competition effects at the 24-hour test. This could be suggestive of a dissociation, such that learning conditions that are more implicit have a negative impact on the level of explicit knowledge gained but little impact on the ‘implicit’ competition effect. Such a dissociation has been suggested in previous adult studies. Tamminen et al (2010) found a correlation between slow wave sleep and overnight gains in recognition speed to novel words but a correlation between sleep spindle density and lexical competition suggestive of different mechanisms underpinned by different components of the sleep architecture. Evidence also suggests that these two types of consolidation (for explicit memory and for integration) can dissociate in children with autism spectrum disorders (Henderson, Powell, Gaskell & Norbury, 2014). Aligning with this dissociation, explicit phonological recall of the new words at the 0-hr session was not associated with the lexical competition effect at the 0-hr session for children (r=.08) or adults (r=.13) or at the 24-hr session for children (r=.20) or adults (r=.04). Thus, explicit memory of new words and lexical integration may be governed by different neurological mechanisms. One could speculate that explicit recollection of new word forms remains more dependent upon the hippocampus whilst lexical integration is more dependent upon gradually strengthened neocortical representations (Eichenbaum et al., 2007).

Another important contribution of the present study was that although the majority of children in the current sample showed evidence of consolidation, children with better existing expressive vocabulary knowledge showed larger consolidation effects, both in terms of strengthening of explicit memory and lexical integration. This supports our second hypothesis, and adds to evidence supporting the “Matthew Effect” in vocabulary acquisition (Ewers & Brownson, 1999; Penno et al., 2002; Senechal et al., 1995; Stanovich, 1986;

Wilkinson & Houston-Price, 2013). The present findings advance this body of evidence, suggesting that existing vocabulary knowledge supports the *integration* of new and established word knowledge, as well as the retention of new word knowledge. This is supported by a recent finding that more proficient bilingual speakers showed a larger N400 event related potential (an electrophysiological marker of semantic integration) to newly learned words presented in context than compared to less proficient bilingual speakers who had less existing knowledge to support learning (Elgort, Perfetti, Rickles, Stafura, 2014). It is also possible, however, that children who are better at integrating new words subsequently develop larger vocabularies. These explanations are not mutually exclusive, and it is most likely that the relationship between lexical integration and vocabulary knowledge is reciprocal. Intriguingly, the same correlation between overnight changes in lexical integration and existing vocabulary knowledge has not been found in our previous studies that have used more direct word-form training (Henderson et al., 2012; Henderson et al., 2013a) even when words are trained with their meanings (Henderson et al., 2013b). This suggests that the value of existing vocabulary knowledge for lexical integration (or vice-versa) is greater when unfamiliar words are learned through stories.

Previous studies examining the association between existing vocabulary knowledge and word learning through stories have tended to use measures of receptive vocabulary knowledge (or vocabulary ‘breadth’; e.g., Dockrell, Braisby & Best, 2007; Ewers & Brownson, 1999; Robbins & Ehri, 1994; Senechal et al., 1995; Wilkinson & Houston-Price, 2013). We used a measure of expressive vocabulary knowledge (or vocabulary ‘depth’) that required children to provide definitions of words. Hence, our data suggest that it is not just the quantity of established vocabulary knowledge that is associated with spoken word learning when listening to stories, but also the quality or richness of semantic representations.

New phonological knowledge may be more easily integrated if it can be associated with a rich network of existing semantic knowledge. Indeed, newly learned information can be incorporated more easily if it is compatible with existing schematic knowledge (Bartlett, 1932; Lewis & Durant, 2011; Tse et al., 2007; van Kesteren et al., 2010). Thus, the observed relationship between lexical integration and vocabulary knowledge might be explained as a form of schema integration effect. In accordance with spreading activation models (Collins & Loftus, 1975) novel words that are associated with a richer network of related concepts may be more likely to be incorporated within the lexicon. Children with better existing knowledge may have also been more able to capitalise on the multiple sentence contexts in which the novel words occurred, which may have further facilitated the strengthening of connections with related concepts (Mol et al., 2009).

However, there are a number of alternative explanations as to why children with better vocabulary showed bigger consolidation effects. For instance, the novel words were embedded in the story such that children had to make inferences to extract word meaning. Studies have supported an association between children's ability to make inferences and their vocabulary knowledge (Cain, Oakhill & Bryant, 2004); hence, the children in this study who had a deeper understanding of the words in the text as a whole may have been better able to make inferences about the meanings of the new words (Robbins & Ehri, 1994; Wilkinson & Houston-Price, 2013) and this may have worked to facilitate consolidation. Children with richer vocabularies may also be more practised at listening to stories and/or motivated to extract new lexical information from stories. This is reflected by the relationship between vocabulary growth during childhood and the quality and quantity of story exposure in the home environment (Hamilton, 2014).

Another important finding of the present study was that children's overnight improvements in their ability to recall the novel words were positively associated with their ability to produce definitions of the novel words at the end of the experiment. This supports the hypothesized association between the consolidation of explicit semantic and phonological knowledge pertaining to a new word. The same association with semantic knowledge was not found for overnight changes in lexical competition, again suggesting that explicit knowledge of new words and lexical integration depend upon different underlying processes. This finding also resonates with Henderson et al (2013b): Children who were taught the meanings of new science words showed better performance on a cued recall task one week after training than children who were not taught the meanings but the presence of word meaning during training had no significant impact on lexical integration (as measured by lexical competition using a pause detection task). Together, these findings emphasise the importance of training phonological and semantic word knowledge in tandem, to facilitate the efficient retrieval of robust long-term lexical representations. This is consistent with previous research from the 'fast mapping' literature, showing that children only retain mappings between novel words and objects (5 minutes after exposure in a referent selection task) when explicit naming is incorporated into training (Horst & Samuelson, 2008).

In conclusion, we present novel evidence that a period of off-line consolidation remains important when children and adults learn novel word forms through spoken stories. Moreover, children with poorer expressive vocabulary knowledge showed less evidence of consolidation after 24 hours, supporting the view that the "rich get richer, and the poor get poorer" with respect to vocabulary acquisition (Ewers & Brownson, 1999; Penno et al., 2002; Senechal et al., 1995; Stanovich, 1986; Wilkinson & Houston-Price, 2013). The findings highlight the importance of story exposure and the value of classroom story time as a device for consolidating new words. Classroom story time may be particularly valuable for children from disadvantaged backgrounds who may experience lower levels of story exposure at home (Hamilton, 2014). However, the low levels of semantic learning observed in this study also underlines the importance of supplementing story readings with more explicit information about target word meaning (Blake, Macdonald, Bayrami, Agosta, & Milian, 2006; Senechal, Thomas & Monker, 1995; Wilkinson & Houston-Price, 2013) particularly for children with vocabulary difficulties (Brett, Rothlein, & Hurley, 1996; Coyne et al., 2004).

Footnotes

[1] This differs from previous studies in which the number of words correctly produced is typically reported (e.g., Henderson et al., 2012, 2013a). When cued recall responses were scored as correct (1 point) or incorrect (0 points), children's means at 0 hrs and 24 hrs were 1.31 (SD=1.07) and 3.02 (SD=2.98) and adults' means were 3.47 (SD=2.87) and 6.06 (SD=2.49), respectively. The main effects of Group and Session were significant (*p*s<.05).

Acknowledgements

This research was funded by the Leverhulme Trust (grant F/00 224/AO). We would like to thank all of the children who took part in this study and the parents and schools for making this possible. We would also like to thank Laura Porter and Rebecca Heron for their valuable assistance with the materials and data collection.

References

Akhtar, N. (2004). Contexts of early word learning. In O. G. Hall & S. R. Waxman (Eds.), *Weaving a lexicon* (pp. 485–507). Cambridge, MA: MIT Press.

Alloway, T.P., Williams, S., Jones, B., & Cochrane, F. (2014). Exploring the impact of television watching on vocabulary skills in toddlers. *Early Childhood Education Journal, 42,* 343-349. DOI 10.1007/s10643-013-0618-1.

Baumann, J.F. and Kameenui, E.J. (1991). Research on vocabulary instruction: Ode To Voltaire. In J. Flood, J.J. Lapp, and J.R. Squire (Eds.), Handbook of research on teaching the English language arts (pp. 604-632). New York: MacMillan.

Bartlett, F.C. (1932). *Remembering: A Study in Experimental and Social Psychology*. Cambridge, England: Cambridge University Press.

Beck, I .L., & McKeown, M. (1991). Conditions of vocabulary acquisition. In R. Barr, M. L. Karmil, P. Mosenthal, & D. D. Pearson (Eds.), *Handbook of reading research* (Vol. 1, pp. 789–814). New York: Longman.

Biemiller, A. (2003). Vocabulary: Needed if more children are to read well. *Reading Psychology, 24*, 323-335. DOI 10.1080/02702710390227297

Biemiller, A. & Boote, C. (2006). An effective method for building meaning vocabulary in primary grades. *Journal of Educational Psychology*, *98*, 44-62. DOI 10.1037/0022-0663.98.1.44

Blake, J., Macdonald, S., Bayrami, L., Agosta, V., & Milian, A. (2006). Book reading styles in dual-parent and single-mother families. *British Journal of Educational Psychology*, *76*(3), 501-515. DOI 10.1348/000709905X49719

Brett, A., Rothlein, L., & Hurley, M. (1996). Vocabulary acquisition from listening to stories and explanations of target words. *The Elementary School Journal, 96*(4), 415± 422. DOI 0013-5984/96/9604-0003$01.00

Brown, H., Weighall, A., Henderson, L.M., & Gaskell, M.G. (2012). Enhanced recognition and recall of new words in 7- and 12-year old children following a period of offline consolidation. *Journal of Experimental Child Psychology, 112*, 56-72. DOI 10.1016/j.jecp.2011.11.010

Cain, K., Oakhill, J., & Bryant, P. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability and component skills. *Journal of Educational Psychology, 96*, 31-42. DOI 1037/0022-0663.96.1.31

Carey, S. (1978). The child as a word learner. In M. Halle, J. Bresnan, & G. A.Miller (Eds.), *Linguistic theory and psychological reality* (pp. 264–293). Cambridge, MA: MIT Press.

Carlo, M. S., August, D., McLaughlin, B., Snow, C. E., Dressler, C., Lippman, D.N., et al. (2004). Closing the gap: Addressing the vocabulary needs of English-language learners in bilingual and mainstream classrooms. *Reading Research Quarterly, 39*, 188–215. DOI 10.1598/RRQ.39.2.3

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological review*, *82*, 407. DOI 10.1037/0033-295X.82.6.407

Coutanche, M. N. & Thompson-Schill, S. L. (2014). Fast mapping rapidly integrates information into existing memory networks. *Journal of Experimental Psychology: General*, 143, 2296-2303. DOI 10.1037/xge0000020

Coyne, M. D., Simmons, D. C., Kame'enui, E. J., & Stoolmiller, M. (2004). Teaching vocabulary during shared storybook readings: An examination of differential effects. *Exceptionality*, *12*, 145-162. DOI 10.1207/s15327035ex1203_3

Cunningham, A. E., & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology, 33*, 934–945. DOI 10.1037/0012-1649.33.6.934

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

Davis, M. H., Di Betta, A. M., Macdonald, M. J. E., & Gaskell, M. G. (2009). Learning and consolidation of novel spoken words. *Journal of Cognitive Neuroscience*, *21*, 803-820. DOI 10.1162/jocn.2009.21059

Dickinson, D. K. (1984). First impressions: Children's knowledge of words gained from a single exposure. *Applied Psycholinguistics, 5*, 359–373. DOI 10.1017/S0142716400005233

Dockrell, J.E., Braisby, N., & Best, R.M. (2007). Children's acquisition of science terms: Simple exposure is insufficient. *Learning and Instruction, 17*, 577 – 594. DOI 10.1016/j.learninstruc.2007.09.005

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

Eichenbaum, H., Yonelinas, A. R., & Ranganath, C. (2007). The medial temporal lobe and recognition memory. *Annual review of neuroscience*, *30*, 123. DOI 10.1146/annurev.neuro.30.051606.094328

Elgort, I., Perfetti, C. A., Rickles, B., & Stafura, J. Z. (2014). Contextual learning of L2 word meanings: second language proficiency modulates behavioural and event-related brain potential (ERP) indicators of learning. *Language, Cognition and Neuroscience*, (ahead-of-print), 1-23. DOI 10.1080/23273798.2014.942673

Elley,W. B. (1989). Vocabulary acquisition from listening to stories. *Reading Research Quarterly, 24*, 174–187. DOI 10.2307/747863

Ewers, C. A.,& Brownson, S. M. (1999). Kindergartners' vocabulary acquisition as a function of active vs. passive storybook reading, prior vocabulary, and working memory. *Reading Psychology, 20*, 11–20. DOI 10.1080/027027199278484

Fernandes, T., Kolinsky, R., & Ventura, P. (2009). The metamorphosis of the statistical segmentation output: Lexicalization during artificial language learning. *Cognition*, *112*, 349-366. DOI 10.1016/j.cognition.2009.05.002

Forster, J. C., & Forster, K. I. (2003). DMDX: A Windows display program with millisecond accuracy. *Behaviour Research Methods, Instruments, & Computers, 35*, 116-124. DOI 10.3758/BF03195503

Gaskell, M.G., & Dumay, N. (2003). Lexical competition and the acquisition of novel words. *Cognition, 89*, 105 – 132. DOI 10.1016/S0010-0277(03)00070-2

Gaskell, M. G. & Marslen-Wilson, W. D. (2002). Representation and competition in the perception of spoken words. *Cognitive Psychology,* 45, 220-266. DOI 10.1016/S0010-0285(02)00003-8

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review, 105,* 251–279. DOI 10.1037/0033-295X.105.2.251

Grainger, J., & Jacobs, A.M. (1996). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review, 103*, 518-565. DOI 10.1037/0033-295X.103.3.518

Graves, M.F. (1986). *Vocabulary learning and instruction*. In E.Z. Rothkopf (Ed.), Review of Research in Education, 13, 49-89.

Gupta, P., & MacWhinney, B. (1997). Vocabulary acquisition and verbal short-term memory: Computational and neural bases. *Brain and Language*, 59, 267-333. DOI doi:10.1006/brln.1997.1819

Hamilton, L. (2014). The role of the home literacy environment in the early literacy development of children at family-risk of dyslexia. *PhD Thesis, University of York.*

Henderson, L., Powell, A., Gaskell, M. G., & Norbury, C.F. (2014). Learning and consolidation of new spoken words in autism spectrum disorder. *Developmental science*, 17, 858-871. DOI 10.1111/desc.12169

Henderson, L., Weighall, A., Brown, H., & Gaskell, G. (2013a). Online lexical competition during spoken word recognition and word learning in children and adults. *Child development*, *84*, 1668-1685. DOI 10.1111/cdev.12067

Henderson, L.M., Weighall, A., Brown, H. & Gaskell, M. G. (2012). Consolidation of vocabulary is associated with sleep in children. *Developmental Science*, *15*, 674-687. DOI 10.1111/j.1467-7687.2012.01172.x

Henderson, L., Weighall, A., & Gaskell, G. (2013b). Learning new vocabulary during childhood: Effects of semantic training on lexical consolidation and integration. *Journal of experimental child psychology*, *116*(3), 572-592. DOI doi:10.1016/j.jecp.2013.07.004

Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development, 74*, 1368–1378. DOI 10.1111/1467-8624.00612

Horst, J.S. (2013). Context and repetition in word learning. *Frontiers in Psychology, 4*: 149. DOI 10.3389/fpsyg.2013.00149

Horst, J. S., Parsons, K. L., & Bryan, N. M. (2011). Get the story straight: contextual repetition promotes word learning from storybooks. *Frontiers in psychology*, *2*. DOI 10.3389/fpsyg.2011.00017

Horst, J.S., & Samuelson, L.K. (2008). Fast mapping but poor retention by 24-month-old infants. *Infancy, 13*, 128-157. DOI 10.1080/15250000701795598

Joshi, R. M. (2005).Vocabulary: A critical component of comprehension. *Reading & Writing Quarterly, 21*, 209–219. DOI 10.1080/10573560590949278

Justice, L. M., Meier, J., & Walpole, S. (2005). Learning new words from storybooks: An efficacy study with at-risk kindergartners. *Language, Speech & Hearing Services in Schools, 36*, 17–32. DOI 10.1044/0161-1461(2005/003)

Keenan, J. M., Betjemann, R. S., Wadsworth, S. J., DeFries, J. C., & Olson, R. K. (2006). Genetic and environmental influences on reading and listening comprehension. *Journal of Research in Reading, 29*, 75–91. DOI 10.1111/j.1467-9817.2006.00293.x

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behaviour Research Methods*, *44*, 978-990. DOI 10.3758/s13428-013-0348-8.

Leach, L., & Samuel, A. G. (2007). Lexical configuration and lexical engagement: When adults learn new words. *Cognitive psychology*, *55*, 306-353. DOI 10.1016/j.cogpsych.2007.01.001

Lewis, P., & Durrant, S.J. (2011). Overlapping memory replay during sleep builds cognitive schemata. *Trends in Cognitive Science, 15*, 343-351. DOI 10.1016/j.tics.2011.06.004

Lindsay, S., & Gaskell, M. G. (2013). Lexical integration of novel words without sleep. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*, 608. DOI 10.1037/a0029243

Marslen-Wilson, W.D. (Ed.) (1989). *Lexical Representation and Process*, Cambridge, MA: MIT Press.

Mattys, S. L., & Clark, J. H. (2002). Lexical activity in speech processing: Evidence from pause detection. *Journal of Memory and Language, 47*, 343-359. DOI 10.1016/S0749-596X(02)00037-2

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology, 10,* 1-86. DOI 10.1016/0010-0285(86)90015-0

McMurray, B., Horst, J., and Samuelson, L. (2012). Word learning emerges from the interaction of online referent selection and slow associative learning. *Psychological Review, 119(4),* 831-877. DOI 10.1037/a0029872

McMurray, B., Samelson, V.M., Hee Lee, S., & Tomblin, J.B. (2010). Individual differences in online spoken word recognition: Implications for SLI. *Cognitive Psychology, 60*, 1-39. DOI Individual differences in online spoken word recognition: Implications for SLI

Metsala, J. L., & Walley, A. C. (1998). Spoken vocabulary growth and the segmental restructuring of lexical representations: precursors to phonemic awareness and early reading ability. In J. L. Metsala & L. C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 89-120). Mahwah, NJ: Erlbaum.

Mol, S. E., Bus, A. G., & de Jong, M. T. (2009). Interactive book reading in early education: A tool to stimulate print knowledge as well as oral language. *Review of Educational Research*, *79*, 979-1007. DOI 10.3102/0034654309332561

Nagy, W. E., & Herman, P. A. (1987). *Breadth and depth of vocabulary knowledge: Implications for acquisition and instruction*. In M. G. McKeown & M. E. Curtis (Eds.), *The nature of vocabulary acquisition*. Hillsdale, NJ: Erlbaum.

Nation, K. (2014). Lexical learning and lexical processing in children with developmental language impairments. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*, 20120387. DOI 10.1098/rstb.2012.0387

Nation, K., & Angell, P. (2006). Learning to read and learning to comprehend. *London Review of Education, 4*, 77–87. DOI 10.1080/13603110600574538

Nation, K., & Snowling, M. J. (2004). Beyond phonological skills: Broader language skills contribute to the development of reading. *J of Research in Reading, 27*, 342–356. DOI 10.1111/j.1467-9817.2004.00238.x

Nemeth, D., Janacsek, K., Londe, Z., Ullman, M. T., Howard, D. V., & Howard Jr, J. H. (2010). Sleep has no critical role in implicit motor sequence learning in young and old adults. *Experimental brain research*, *201*, 351-358. DOI 10.1007/s00221-009-2024-x

Norris, D. (1994). Shortlist: a connectionist model of continuous speech recognition. *Cognition, 52,* 189-234. DOI 10.1016/0010-0277(94)90043-4

Perfetti, C.A., & Hart, L. (2002). The lexical quality hypothesis. In L. Vehoeven. C. Elbro, & P. Reitsma (Eds.), *Precursors of functional literacy* (pp. 189-213). Amsterdam/Philadelphia: John Benjamins.

Penno, J. F., Wilkinson, I. A. G., & Moore, D. W. (2002). Vocabulary acquisition from teacher explanation and repeated listening to stories: Do they overcome the Matthew effect? *Journal of Educational Psychology, 94*, 23–33. DOI 10.1037/0022-0663.94.1.23

Reese, E., & Cox, A. (1999). Quality of adult book reading affects children’s emergent literacy. *Developmental Psychology, 35*, 20–28. DOI 10.1037/0012-1649.35.1.20

Robbins, C., & Ehri, L. C. (1994). Reading storybooks to kindergartners helps them learn new vocabulary words. *Journal of Educational Psychology, 86*, 54–64. DOI 10.1037/0022-0663.86.1.54

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274,* 1926-1928. DOI 10.1126/science.274.5294.1926

Senechal, M. (1997). The differential effect of storybook reading on preschoolers’ acquisition of expressive and receptive vocabulary. *Child Language, 24*, 123–138. DOI?

Senechal, M., & Cornell, E. H. (1993). Vocabulary acquisition through shared reading experiences. *Reading Research Quarterly, 28*, 361–373. DOI http://www.jstor.org/stable/747933

Sénéchal, M., Thomas, E., & Monker, J. A. (1995). Individual differences in 4-year-old children's acquisition of vocabulary during storybook reading. *Journal of Educational Psychology*, *87*, 218. DOI 10.1037/0022-0663.87.2.218

Sharon, T., Moscovitch, M., & Gilboa, A. (2011). Rapid neocortical acquisition of long-term arbitrary associations independent of the hippocampus. *Proceedings of the National Academy of Sciences*, *108*(3), 1146-1151. DOI 10.1073/pnas.1005238108

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*, 360–401. Doi http://www.jstor.org/stable/747612

Stanovich, K. E., Cunningham, A. E., & Feeman, D. J. (1984). Intelligence, cognitive skills, and early reading progress. *Reading Research Quarterly, 19*, 278–303. Doi http://www.jstor.org/stable/747822

Storkel, H. L., & Hoover, J. R. (2011). The influence of part-word phonotactic probability/neighborhood density on word learning by preschool children varying in expressive vocabulary. *Journal of Child Language, 38*, 628-643. Doi 10.1017/S0305000910000176

Swanborn, M.S. L., & de Glopper, K. (1999). Incidental word learning while reading: A meta-analysis. *Review of Educational Research, 69*, 261–285. Doi Incidental word learning while reading: A meta-analysis

Szmalec, A., Page, M., & Duyck, W. (2012). The development of long-term lexical representations through Hebb repetition learning. *Journal of Memory and Language*, *67*, 342-354. DOI 10.1016/j.jml.2012.07.001

Takashima, A., Bakker, I., Van Hell, J. G., Janzen, G., & McQueen, J. M. (2014). Richness of information about novel words influences how episodic and semantic memory networks interact during lexicalization. *NeuroImage, 84*, 265-278. DOI 10.1016/j.neuroimage.2013.08.023

Tamminen, J. & Gaskell, M.G. (2008). Newly learned spoken words show long-term lexical competition effects. *The Quarterly Journal of Experimental Psychology, 61*, 361-371. DOI 10.1080/17470210701634545

Tamminen, J., Payne, J. D., Stickgold, R., Wamsley, E. J., & Gaskell, M. G. (2010). Sleep Spindle Activity is Associated with the Integration of New Memories and Existing Knowledge. *Journal of Neuroscience, 30*, 14356-14360. DOI 10.1523/JNEUROSCI.3028-10.2010

Tse, D. et al. (2007). Schemas and memory consolidation. *Science, 316*, 76-82. DOI 10.1126/science.1135935

van Kesteren, M., Fernandez, G., Norris, D.G., & Hermans, E.J. (2010). Persistent schema-dependent hippocampal-neocortical connectivity during memory encoding and postencoding rest in humans. *Proceedings of the National Academy of Science, USA, 107*, 7550-7555. DOI 10.1073/pnas.0914892107

Wagenmakers, E.J.M., Steyvers, M., Raaijmakers, J.G.W., Shiffrin, R.M., Van Rijn, H., & Zeelenberg, R. (2004). A model for evidence accumulation in the lexical decision task. *Cognitive Psychology, 48,* 332-367. DOI 10.1016/j.cogpsych.2003.08.001

Waisk, B. A., & Bond, M. A. (2001).Beyond the pages of a book: Interactive book reading and language development in preschool classrooms. *Journal of Educational Psychology, 93*, 243–250. DOI 10.1037/0022-0663.93.2.243

Walsh, B. A., & Blewitt, P. (2006). The effect of questioning style during storybook reading on novel vocabulary acquisition of preschoolers. *Early Childhood Education Journal, 33*, 273–278. DOI 10.1007/s10643-005-0052-0

Wilkinson, K. S., & Houston-Price, C. (2013). Once upon a time, there was a pulchritudinous princess...: The role of word definitions and multiple story contexts in children's learning of difficult vocabulary. *Applied Psycholinguistics*, *34*, 591-613. DOI 10.1017/S0142716411000889

Wilson, M. (1988). MRC Psycholinguistic Database – machine-usable dictionary, version 2.00. *Behaviour Research Methods Instruments & Computers, 20*, 6 – 10. DOI 10.3758/BF03202594

Williams, S. E., & Horst, J. S. (2014). Goodnight book: sleep consolidation improves word learning via storybooks. *Frontiers in psychology*, *5*. DOI 10.3389/fpsyg.2014.00184

Table 1. Descriptive statistics for measures of explicit novel word knowledge and mean (and SD) pause detection RT (in ms) and errors (out of 12) for competitor and control conditions at the –0 hr and 24 hr tests, for children and adults

			Children		Adults	
			Mean (SD)	Range	Mean (SD)	Range
Cued Recall (max 24)	0 hr		2.93 (2.25)	0-9	8.55 (4.46)	1-20
	24 hr		5.93 (4.24)	0-14	13.64 (4.09)	7-22
Word Recognition (% correct)	0 hr		71.04 (16.98)	41.67-100	90.91 (8.17)	75-100
	24 hr		81.46 (15.84)	33.33-100	92.93 (7.25)	75-100
Definitions	24 hr		2.20 (1.68)	0-6	6.36 (3.06)	1-12
RT (ms)	0 hr	Competitor	1096 (333)		770 (203)	
		Control	1080 (321)		777 (204)	
	24 hr	Competitor	1052 (317)		760 (204)	
		Control	917 (190)		711 (189)	
Mean Errors (out of 12)	0 hr	Competitor	1 (1.24)		0.40 (0.61)	
		Control	1.05 (1.06)		0.30 (0.59)	
	24 hr	Competitor	1 (1.04)		0.52 (0.76)	
		Control	1 (0.88)		0.33 (0.54)	

Figure 1. Lexical competition effects (competitor RT – control RT) at the 0 hr and 24 hr tests for children and adults (error bars display 95% confidence intervals)

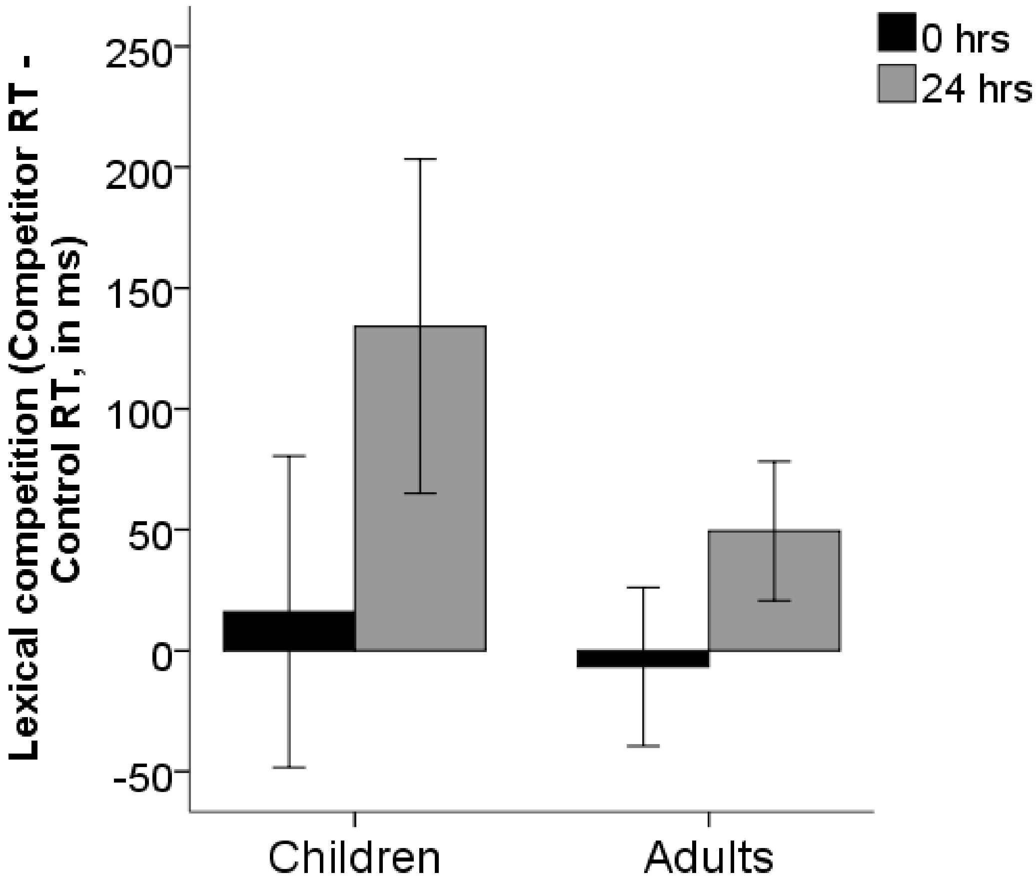

Figure 2. Scatterplots showing the correlations between children's existing vocabulary and overnight changes in lexical competition and cued recall.