CHILD DEVELOPMENT PERSPECTIVES

Learning to Remember Words: Memory Constraints as Double-Edged Sword Mechanisms of Language Development

Haley A. Vlach

University of Wisconsin, Madison

ABSTRACT—***Word learning researchers have historically been interested in elucidating the mechanisms that allow children to encode words. Recent research has moved beyond the moment of encoding to examine the processes underlying children's retention and retrieval of words across time. This work has revealed significant memory constraints on children's word learning. That is, children struggle to retain and retrieve newly learned words. This review outlines research suggesting that describing these processes as memory constraints may mischaracterize how memory shapes language development. Instead, memory constraints are more accurately characterized as double-edged sword mechanisms; limited memory abilities likely hinder and promote children's word learning simultaneously. The review concludes with suggestions for developing a theory of how children learn to remember words.***

KEYWORDS—*forgetting; retrieval effort; word learning; memory development; language development; cognitive development*

Haley A. Vlach, Department of Educational Psychology, University of Wisconsin, Madison.

I thank my colleagues for their insights, feedback, and continual support over the years. Special thanks to Megan Kaul and Melina Knabe for reviewing this article. The process of writing this article was supported in part by National Institutes of Health grant R21 HD092867-01.

Correspondence concerning this article should be addressed to Haley A. Vlach, Department of Educational Psychology, 859 Educational Sciences, 1025 W. Johnson Street, Madison, WI 53706; e-mail: hvlach@wisc.edu.

DOI: 10.1111/cdep.12337

Linguists, philosophers, and psychologists have historically characterized language learning as seemingly impossible (e.g., Quine, 1960). Word learning is particularly difficult because children need to resolve high degrees of referential ambiguity to determine word-referent pairings. For instance, if children are at the park and hear the word *dog*, this word could refer to the creature they see wagging its tail and barking, but it could also refer to a seemingly infinite number of other referents, such as a color, action, and abstract concept. Thus, a central pursuit in cognitive science, developmental psychology, and other disciplines has been to understand the mechanisms that make word learning possible.

Over the last few decades, researchers have focused on studying how children learn to associate words with referents in a single moment. In these studies, children are exposed to a novel word and a novel object, then are immediately tested on whether they mapped the word to the correct object. For instance, in a typical experiment (see Figure 1), children are presented with a novel linguistic label (*wug*) and a novel object. At test, children are shown several objects and asked to identify the *wug*. If children recognize or generalize the word *wug* at an immediate test, researchers conclude that children successfully encoded the word presented to them. Indeed, researchers have found that children can readily map words to referents after just a few learning trials, a behavior termed *fast mapping* (see Carey, 2010, for a review).

Three classes of theories have been proposed to explain children's rapid mapping of words to referents: domain-general theories (Smith, 2002), constraints-principles theories (Markman, 1991), and social-pragmatic theories (Tomasello, 1992). Taken together, these theories suggest that children use a variety of tools to determine word mappings. For instance, children use associative/statistical learning to track regularities in perceptual information (Krogh, Vlach, & Johnson, 2013), apply rules (e.g., mutual exclusivity; Kalashnikova, Mattock, & Monaghan, 2016), and rely on the social cues of others (Yurovsky & Frank, 2017)

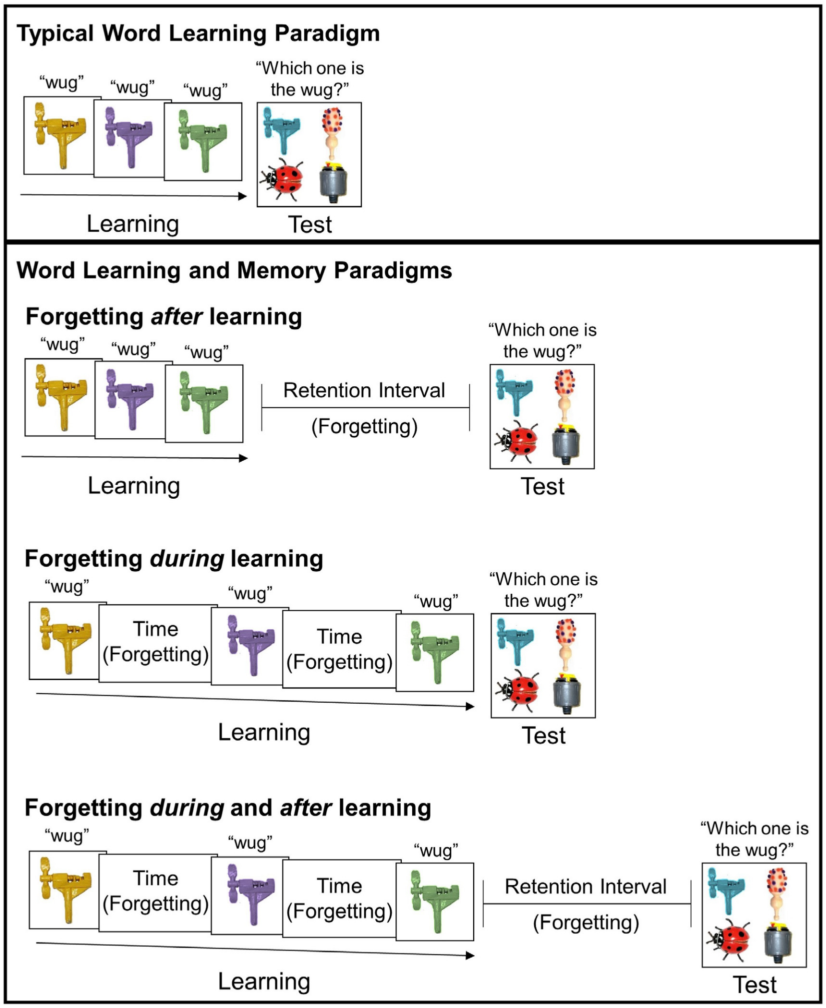

Figure 1. **Paradigms used by researchers to study children's word learning. Top panel provides an example of a typical paradigm to study children's encoding of new words. Bottom panel provides examples of recent paradigms used to study children's memory for words. Researchers insert time/retention intervals during learning, after learning, and/or during and after learning in order to determine how forgetting affects children's language acquisition. [Color figure can be viewed at wileyonlinelibrary.com]**

to learn words. This body of work suggests that referential ambiguity is a problem children solve readily; children have powerful tools for learning the properties and meanings of words.

One limitation of the classic fast mapping paradigm is that word learning is assessed at an immediate test. Consequently, these studies tell us when and how children encode new words, but they tell us little about word learning over time (often termed *slow* or *extended mapping*; Carey, 2010). To address this limitation, researchers who study children's language development have conducted tests with a delay between learning and testing (e.g., Horst & Samuelson, 2008; Karaman & Hay, 2018; Vlach, Ankowski, & Sandhofer, 2012; Wojcik, 2013). These studies have sought to determine whether young learners can encode, retain, and retrieve words across time. In this work, infants and children have struggled to remember words across timescales, including seconds (Vlach & Johnson, 2013), minutes (Horst & Samuelson, 2008), days (Vlach & Sandhofer, 2012), and weeks or months (Gordon et al., 2016; Vlach & Sandhofer, 2012; Wojcik, 2017). That is, in these experiments, infants and children failed to reliably identify the *wug* from a series of objects after a delay. Thus, researchers have concluded that there are significant memory constraints on word learning and language acquisition (e.g., Adams & Gathercole, 2000; Barr, 2013; Endress, Nespor, & Mehler, 2009; Horst & Samuelson, 2008; Vlach & Johnson, 2013).

Many memory constraints likely shape language development; hence, in this review, I focus on one constraint as a case example: children's forgetting of information during and after word mapping. In the first section of the article, I describe how forgetting has been characterized historically as a constraint on language acquisition. While rapidly forgetting words would seem to be an obstacle to word learning, I counter this assumption in the next section, proposing that rapid forgetting of new words is a key factor of children's language development. Thus, forgetting serves as an example of how memory constraints are best characterized as double-edged sword mechanisms of learning. Finally, I conclude with recommendations for how researchers can construct a theory of how children learn to remember words.

FORGETTING AS A BARRIER TO WORD LEARNING

One of the most significant memory constraints on human cognition is the fact that we continuously forget learned information. Forgetting has been studied since the beginning of psychological science and is operationally defined as the diminishing ability to retrieve information across time (Ebbinghaus, 1964; Wixted, 2004). Learners forget information according to a curvilinear function, referred to as a *forgetting curve*. Because forgetting has been studied widely and we know how forgetting unfolds over time, it provides a fruitful way to understand whether and how memory processes affect language development. That is, research on memory suggests that children should forget words in a pattern similar to other types of nonlinguistic information (i.e., according to a forgetting curve) and, if this occurs, that it suggests that language development relies heavily on memory systems. To understand the role of forgetting in children's word learning, researchers insert time intervals between learning events or between learning and testing. The time intervals allow children to forget information, and researchers can observe how this forgetting changes the course of word learning (for examples of these paradigms, see Figure 1).

During the first 2 years of life, small intervals of time between word mapping events constrain infants' language acquisition. As infants grow older, their forgetting slows (Rovee-Collier & Cuevas, 2008) and their word learning is less likely to be constrained by small time intervals. For instance, in one study (Vlach & Johnson, 2013), 16- and 20-month-olds were presented with a cross-situational word learning task in which half the words and objects were presented six times in immediate succession (see Figure 1, classic fast mapping paradigm) and half were presented six times with other learning trials interleaved between presentations (creating 26s time intervals; see Figure 1, forgetting during learning). At an immediate word mapping test, the 16-month-olds demonstrated that they learned words for the objects presented in immediate succession, but not for the objects with inserted time intervals. This suggests that the small-time intervals between presentations caused the 16-month-olds to forget rapidly, to the point where they could no longer successfully retrieve information at subsequent learning events or at the test. However, the 20-month-olds learned words for both the massed and interleaved objects, suggesting that the 26s time intervals did not constrain their word learning.

Although infants may be able to quickly overcome the demands of retaining and retrieving information across short time intervals during learning, retaining and retrieving knowledge after learning appears has a more protracted course of development. Even brief delays between learning and testing can deter children's word learning (Horst & Samuelson, 2008; Vlach et al., 2012; see Figure 1, forgetting after learning). Moreover, years may pass between the time when children can successfully encode words and the time when they can successfully retain and retrieve words across time (Vlach & DeBrock, 2019). For example, children can learn new words via cross-situational statistical learning, which requires them to track co-occurrences between words and objects to encode words, as early as 12 months (Escudero, Mulak, & Vlach, 2016; Smith & Yu, 2008). However, children cannot reliably retain and retrieve words across a 5-minute testing delay during cross-situational statistical learning until age 4 (Vlach & DeBrock, 2019). Indeed, the general time course of word learning is that children encode word mappings as early as 6 months (Bergelson & Swingley, 2012), but do not easily retrieve and produce the same words until a year or more later (Fenson et al., 1994). This work demonstrates that overcoming the challenge of retaining and retrieving words after learning could take up to several years.

Why do we see different developmental trajectories in word mapping versus retaining and retrieving words across time? One reason for the dissociation between encoding and subsequent storage and retrieval of words is that these processes may involve different mechanisms or networks in the brain. In studies of adults with brain injuries, patients reliably mapped words to objects, suggesting that they can encode new words rapidly (Sharon, Moscovitch, & Gilboa, 2011; Warren, Tranel, & Duff, 2016). However, at a delayed test, patients failed to recognize or retrieve the words learned earlier in the experiment. In contrast, healthy adults in a control group did not struggle to retrieve words in a delayed test. These results have led researchers to conclude that certain networks in the brain, such as hippocampal networks, are unnecessary for immediate word mapping but critical to the long-term ability to retain and retrieve words. Thus, underdeveloped hippocampal networks may be the source of children's rapid forgetting of words after learning, and these hippocampal networks may take several years of experience to develop.

Given the findings of this work, researchers have taken a special interest in identifying ways to support children's long-term memory for words and enhance the development of their hippocampal networks. Based on this research, we know that making small changes to the learning environment—such as providing additional environmental cues (Vlach & Sandhofer, 2012), reinstantiating the context in which the word was initially learned (Horst, 2013; Horst, Parsons, & Bryan, 2011; Vlach & Sandhofer, 2011), and distributing learning across time (Vlach et al., 2012)—can improve children's memory for words. However, even when providing these supports, children still forget newly learned words according to a typical forgetting curve. For instance, in one study (Vlach & Sandhofer, 2012), 3-year-olds were presented with a naturalistic fast mapping task in which they were given a novel label while measuring a novel object with a ruler, and then tested on the word mapping immediately, 1 week later, or 1 month later. Across four conditions, children were either given no memory cues, one memory cue, two memory cues, or three memory cues. Providing additional memory cues improved children's ability to retain and retrieve words after delays of 1 week and 1 month, but children still forgot words. Thus, like other types of learned information (Bjork & Bjork, 1992), improving the encoding strength of words can slow forgetting, but it cannot prevent forgetting from happening entirely.

This work might suggest that forgetting makes retrieving words more difficult and that forgetting is therefore best characterized as a constraint on language development. However, in the next section, I propose that this memory constraint promotes the ability to learn new words. In particular, I argue that the years infants and children spend struggling to retrieve words is precisely what gives rise to the more fluent retrieval of language observed in older childhood and adulthood.

FORGETTING AS A DRIVER OF LANGUAGE DEVELOPMENT

Forgetting makes retrieving information across time more difficult. However, this difficulty in retrieval has a silver lining: Learners must use more cognitive effort to retrieve information, which in turn slows the future rate of forgetting. In memory research on desirable difficulties (Bjork, 1994; also see retrieval effort hypothesis, Pyc & Rawson, 2009), learning conditions that engender challenging retrieval promote long-term memory for information.

In several studies, desirable difficulties promoted children's language development (Vlach et al., 2012; Vlach & Sandhofer, 2011). For instance, in one study (Vlach et al., 2012), 2-year-olds were presented with a novel noun generalization task that had three learning conditions: simultaneous (all objects were presented simultaneously, preventing forgetting during learning), massed (see Figure 1, forgetting after learning), and spaced (see Figure 1, forgetting during and after learning). Children's ability to retrieve and generalize the novel word (e.g., *wug*) was tested during learning and after a 15-minute delay (see Figure 2 for information on performance across the three conditions). Children in the simultaneous and massed conditions did not struggle to retrieve information during learning (see Figure 2, top figure). In contrast, children in the spaced condition failed to retrieve the words after the first learning trial. Over the course of the learning phase, their ability to retrieve improved with practice, but was still less successful than in the other conditions. At the 15-minute delayed test, children in the spaced condition were the only group that could retain the words learned earlier in the experiment (see Figure 2, bottom figure). Thus, children who forgot during learning had the strongest memory for words.

I propose that the years infants and young children spend failing, struggling, and then successfully retrieving words is what give rise to the more fluent retrieval observed in older childhood and adulthood (for a depiction of how this process unfolds, see Figure 3). At first, infants encode words, but then forget them so

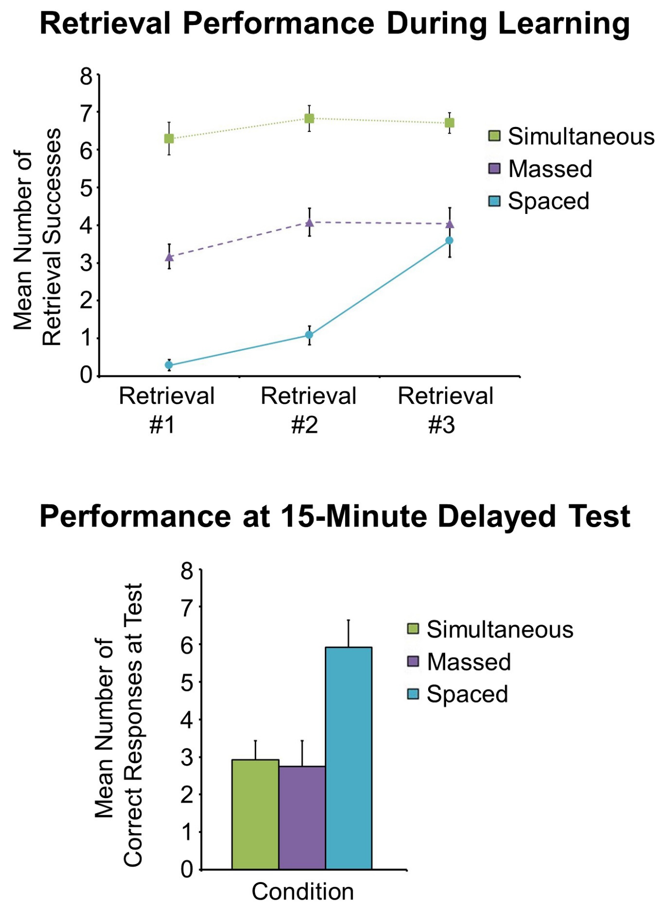

Figure 2. Results adapted from Vlach et al. (2012; Experiment 2). The top figure represents the mean number of retrieval successes by retrieval event (first retrieval event at second wug presentation, second retrieval event at third wug presentation, and third retrieval event at fourth wug presentation) and condition (simultaneous, massed, and spaced). The bottom figure represents mean number of correct responses at the 15-minute delayed test by presentation timing condition (simultaneous, massed, or spaced). Children in the spaced condition struggled to retrieve words during learning, but had the highest test performance across time. Error bars in both figures represent standard errors. [Color figure can be viewed at wileyonlinelibrary.com]

quickly that they cannot be retrieved after even a brief delay (see the yellow and green forgetting curves in Figure 3). That is, infants and children do not have access to any part of this specific representation (e.g., the shape of the object, the auditory label *wug*) and these representations cannot be retrieved in the future. Consequently, each time children learn the word *wug*, it is not bound in memory to previous learning events.

As hippocampal networks improve with maturation and practice retrieving (see the difference in the rate between the yellow and green forgetting curves in Figure 3), older infants and children begin to successfully retrieve components of learned words at subsequent learning events (e.g., retrieving information about the purple *wug* when presented with the blue *wug*, such as the shape of the object). At first, retrieving these words successfully requires much cognitive effort (see the purple vertical dotted line in Figure 3). However, by successfully retrieving words learned

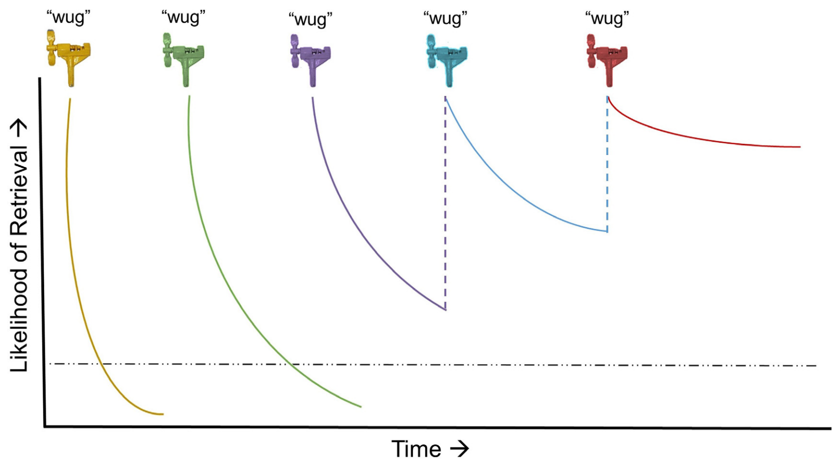

Figure 3. **Hypothetical forgetting curves for words. Early in development, children rapidly forget words and are unable to retrieve previous learning (e.g., yellow and green forgetting curves). The horizontal dotted line represents the theoretical point at which information is no longer able to be retrieved from memory. As hippocampal networks mature with practice, forgetting slows. Children are then able to retain and retrieve words from subsequent learning events (e.g., retrieving purple wug when presented with the blue wug). The vertical dotted lines represent the cognitive effort required to retrieve and bind the previous memory with the current word learning event. This cognitive effort results in increased retrieval strength for the word, slowing the future forgetting rate. With each repetition of the word this process continues, eventually leading to slow forgetting and fluent retrieval of words. [Color figure can be viewed at wileyonlinelibrary.com]**

earlier, children reactivate the memory of the word and bind it in memory to the current word learning event. This cognitive effort in memory reactivation in turn improves the retrieval strength of the word and slows the future rate of forgetting for that word.

This process repeats itself with each new exposure to the word; with increased exposure comes more practice retrieving words, leading to slower forgetting (see the last forgetting curve in Figure 3) and fluent retrieval (see the shortest dotted vertical line in Figure 3). Children readily retrieve multiple dimensions of the representation (e.g., shape, color, and textures of the object, other related words) with ease. Indeed, repetition is more than just an opportunity to enhance encoding; the forgetting between repetitions gives children experience remembering words. Thus, forgetting is a mechanism that drives children's word learning and language development.

In summary, it is often assumed that rapid forgetting uniformly constrains infants' and children's word learning. Indeed, forgetting causes young learners to fail to retrieve words. However, research shows that forgetting promotes children's long-term memory for words (Vlach, 2014; Vlach et al., 2012). A more accurate characterization of forgetting is that it is a double-edged sword mechanism of word learning: While it causes language development to be slow and difficult initially, it also causes children to dedicate cognitive effort to storing and remembering information, eventually leading them to retrieve words with ease.

NEXT STEPS: A THEORY OF HOW CHILDREN LEARN TO REMEMBER WORDS

We know relatively little about how infants and children successfully store and retrieve words from long-term memory. As an important next step in the field of language development, we need to shift our focus from debating philosophical arguments of how children encode words (e.g., associative accounts vs. hypothesis-testing accounts) to constructing a theory of how children learn to remember words. This new theoretical account should describe memory constraints as double-edged sword mechanisms; a single memory process, such as forgetting, is likely to simultaneously hinder and promote word learning. Indeed, researchers need to identify both sides of the sword for an accurate characterization of how language acquisition develops.

In building a new theory of how children learn to remember words, I suggest that researchers take the following five steps: First, they should continue studying forgetting and word learning. Forgetting has been researched for a long time (Ebbinghaus, 1964; Wixted, 2004), and researchers have only begun to connect these two areas of work. Second, special consideration should be given to studying memory consolidation; although I have not reviewed that process here, recent research suggests that sleep-dependent consolidation is a critical process in early language development (Dionne et al., 2011; Simon et al., 2016; Werchan & Gómez, 2014). Third, researchers should examine whether and how the nature of the word-referent representation changes with time. Do children retrieve the same representation they encoded during learning? How does retrieving the representation change the representation? Fourth, researchers should identify other memory processes that are considered constraints on language development and determine if these constraints are actually double-edged sword mechanisms. For example, many researchers have claimed that infants' and children's limited working or short-term memory capacities constrain language learning (Adams & Gathercole, 2000; Gathercole & Baddeley, 2014). However, studies with adult learners suggest possible benefits to having limited working or short-term memory (DeCaro, Thomas, & Beilock, 2008; Gaissmaier, Schooler, & Rieskamp, 2006). Finally, researchers should examine whether findings from laboratory-based tasks scale up to real contexts.

Focusing on memory processes in word learning also affords an opportunity to bridge psychological science with applied settings. Although investigating philosophical arguments about language has generated many theories of language learning, this work has led to very few recommendations for how to teach children language in real contexts. Speech-language pathologists and educators have called for more research on memory and language development because this type of work would be valuable for designing language interventions (Storkel, 2015); many patients and students struggle to retain and retrieve words across time (Sharon et al., 2011; Storkel, 2015; Warren et al., 2016). Indeed, a theory of how children learn to remember words may be the greatest gift language researchers can give practitioners.

REFERENCES

Adams, A. M., & Gathercole, S. E. (2000). Limitations in working memory: Implications for language development. *International Journal of Language & Communication Disorders*, *35*, 95–116. https://doi.org/10.1080/136828200247278

Barr, R. (2013). Memory constraints on infant learning from picture books, television, and touchscreens. *Child Development Perspectives*, *7*, 205–210. https://doi.org/10.1111/cdep.12041

Bergelson, E., & Swingley, D. (2012). At 6–9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences*, *109*, 3253–3258. https://doi.org/10.1073/pnas.1113380109

Bjork, R. A..(1994). Memory and metamemory considerations in the training of human beings. In J. Metcalf & A. P. Shimura (Eds.), *Metacognition: Knowing about knowing* (pp. 185–205). Cambridge, MA: MIT Press.

Bjork, R. A., & Bjork, E. L. (1992). A new theory of disuse and an old theory of stimulus fluctuation. In A. Healy, S. Kosslyn, & R. Shiffrin (Eds.), *From Learning processes to cognitive processes: Essays in honor of William K. Estes* (Vol 2, pp. 35–67). Hillsdale, NJ: Erlbaum.

Carey, S. (2010). Beyond fast mapping. *Language Learning and Development*, *6*, 184–205. https://doi.org/10.1080/15475441.2010.484379

DeCaro, M. S., Thomas, R. D., & Beilock, S. L. (2008). Individual differences in category learning: Sometimes less working memory capacity is better than more. *Cognition*, *107*, 284–294. https://doi.org/10.1016/j.cognition.2007.07.001

Dionne, G., Touchette, E., Forget-Dubois, N., Petit, D., Tremblay, R. E., Montplaisir, J. Y., & Boivin, M. (2011). Associations between sleep-wake consolidation and language development in early childhood: A longitudinal twin study. *Sleep*, *34*, 987–995. https://doi.org/10.5665/SLEEP.114

Ebbinghaus, H. (1964). Memory: A contribution to experimental psychology. In H. A. Ruger, C. E. Bussenius, & E. R. Hilgard Trans (Eds.), *Publications. (Original work published 1885)*. New York, NY: Dover.

Endress, A. D., Nespor, M., & Mehler, J. (2009). Perceptual and memory constraints on language acquisition. *Trends in Cognitive Sciences*, *13*, 348–353. https://doi.org/10.1016/j.tics.2009.05.005.

Escudero, P., Mulak, K. E., & Vlach, H. A. (2016). Infants encode phonetic detail during cross-situational word learning. *Frontiers in Psychology*, *7*, 1419. https://doi.org/10.3389/fpsyg.2016.01419

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., & Pethick, S. J.. (1994). Variability in Early Communicative Development. *Monographs of the Society for Research in Child Development*, *59* (serial 242), 1–173. https://doi.org/10.2307/1166093

Gaissmaier, W., Schooler, L. J., & Rieskamp, J. (2006). Simple predictions fueled by capacity limitations: When are they successful? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 966–982. https://doi.org/10.1037/0278-7393.32.5.966

Gathercole, S. E., & Baddeley, A. D. (2014). *Working memory and language*. London, England: Psychology Press.

Gordon, K. R., McGregor, K. K., Waldier, B., Curran, M. K., Gomez, R. L., & Samuelson, L. K. (2016). Preschool children's memory for word forms remains stable over several days, but gradually decreases after 6 months. *Frontiers in Psychology*, *7*, 1–15. https://doi.org/10.3389/fpsyg.2016.01439

Horst, J. S. (2013). Context and repetition in word learning. *Frontiers in Psychology*, *4*, 1–11. https://doi.org/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*, 1–11. https://doi.org/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. https://doi.org/10.1080/15250000701795598

Kalashnikova, M., Mattock, K., & Monaghan, P. (2016). Flexible use of mutual exclusivity in word learning. *Language Learning and Development*, *12*, 79–91. https://doi.org/10.1080/15475441.2015.1023443

Karaman, F., & Hay, J. F. (2018). The longevity of statistical learning: When infant memory decays, isolated words come to the rescue. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*, 221–232. https://doi.org/10.1037/xlm0000448

Krogh, L., Vlach, H. A., & Johnson, S. P. (2013). Statistical learning in development: Flexible yet constrained. *Frontiers in Psychology*, *3*, 598. https://doi.org/10.3389/fpsyg.2012.00598

Markman, E. M. (1991). *Categorization and naming in children: Problems of induction*. Cambridge, MA: MIT Press.

Pyc, M. A., & Rawson, K. A. (2009). Testing the retrieval effort hypothesis: Does greater difficulty correctly recalling information lead to higher levels of memory? *Journal of Memory and Language*, *60*, 437–447. https://doi.org/10.1016/j.jml.2009.01.004

Quine, W. V. O. (1960). *Word and object*. Cambridge, MA: MIT Press.

Rovee-Collier, C., & Cuevas, K. (2008). The development of infant memory. In M. Courage & N. Cowan (Eds.), *The development of memory in infancy and childhood* (pp. 23–54). London, England: Psychology Press.

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*, 1146–1151. https://doi.org/10.1073/pnas.1005238108

Simon, K. N., Werchan, D., Goldstein, M. R., Sweeney, L., Bootzin, R. R., Nadel, L., & Gómez, R. L. (2016). Sleep confers a benefit for retention of statistical language learning in 6.5 month old infants. *Brain and Language*, *167*, 3–12. https://doi.org/10.1016/j.bandl.2016.05.002

Smith, L. B. (2002). Learning how to learn words: An associate crane. In R. M. Golinkoff, K. Hirsh-Pasek, L. Bloom, L. Smith, A. Woodward, & N. Akhtar . . . G. Hollich (Eds.), *Becoming a word learner: A debate on lexical acquisition*. New York, NY: Oxford University Press.

Smith, L. B., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, *106*, 1558–1568. https://doi.org/10.1016/j.cognition.2007.06.010

Storkel, H. L. (2015). Learning from input and memory evolution: Points of vulnerability on a pathway to mastery in word learning. *International Journal of Speech-Language Pathology*, *17*, 1–12. https://doi.org/10.3109/17549507.2014.987818

Tomasello, M. (1992). The social bases of language acquisition. *Social Development*, *1*, 67–87.

Vlach, H. A. (2014). The spacing effect in children's generalization of knowledge: Allowing children time to forget promotes their ability to learn. *Child Development Perspectives*, *8*, 163–168. https://doi.org/10.1111/cdep.12079

Vlach, H. A., Ankowski, A. A., & Sandhofer, C. M. (2012). At the same time or apart in time? The role of presentation timing and retrieval dynamics in generalization. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *38*, 246–254. https://doi.org/10.1037/a0025260

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*, 700–711. https://doi.org/10.1037/xlm0000611

Vlach, H. A., & Johnson, S. P. (2013). Memory constraints on infants' cross-situational statistical learning. *Cognition*, *127*, 375–382. https://doi.org/10.1016/j.cognition.2013.02.015

Vlach, H. A., & Sandhofer, C. M. (2011). Developmental differences in children's context-dependent word learning. *Journal of Experimental Child Psychology*, *108*, 394–401. https://doi.org/10.1016/j.jecp.2010.09.011

Vlach, H. A., & Sandhofer, C. M. (2012). Fast mapping across time: Memory mechanisms support children's ability to retain words. *Frontiers in Psychology*, *3*, 46. https://doi.org/10.3389/fpsyg.2012.00046

Warren, D. E., Tranel, D., & Duff, M. C. (2016). Impaired acquisition of new words after left temporal lobectomy despite normal fast-mapping behavior. *Neuropsychologia*, *80*, 165–175. https://doi.org/10.1016/j.neuropsychologia.2015.11.016

Werchan, D. M., & Gómez, R. L. (2014). Wakefulness (not sleep) promotes generalization of word learning in 2.5-year-old children. *Child Development*, *85*, 429–436. https://doi.org/10.1111/cdev.12149

Wixted, J. T. (2004). The psychology and neuroscience of forgetting. *Annual Review of Psychology*, *55*, 235–269. https://doi.org/10.1146/annurev.psych.55.090902.141555

Wojcik, E. H. (2013). Remembering new words: Integrating early memory development into word learning. *Frontiers in Psychology*, *4*, 1–12. https://doi.org/10.3389/fpsyg.2013.00151

Wojcik, E. H. (2017). 2.5-year-olds' retention and generalization of novel words across short and long delays. *Language Learning and Development*, *13*, 300–316. https://doi.org/10.1080/15475441.2016.1275644

Yurovsky, D., & Frank, M. C. (2017). Beyond naïve cue combination: Salience and social cues in early word learning. *Developmental Science*, *20*, e12349. https://doi.org/10.1111/desc.12349