Rapid Semantic Integration of Novel Words Following Exposure to Distributional Regularities

Abstract

Our knowledge of words consists of a lexico-semantic network in which different words and their meanings are connected by relations, such as similarity in meaning. This research investigated the integration of new words into lexico-semantic networks. Specifically, we investigated whether new words can rapidly become linked with familiar words given exposure to distributional regularities that are ubiquitous in real-world language input, in which familiar and new words either: (1) directly co-occur in sentences, or (2) never co-occur, but instead share each other's patterns of co-occurrence with another word. We observed that, immediately after sentence reading, familiar words came to be primed not only by new words with which they co-occurred in sentences, but also by new words with which they shared co-occurrence. This finding represents a novel demonstration that new words can be rapidly integrated into lexico-semantic networks from exposure to distributional regularities.

Keywords: word learning; semantic priming; distributional semantics; semantic integration

Introduction

Starting early in development and continuing through adulthood, we amass sizable vocabularies commonly estimated to contain tens of thousands of words (Schmitt & McCarthy, 1997). Beyond the size of the resulting vocabulary, word learning is remarkable both because much of it unfolds merely by encountering words in linguistic contexts without explicit instruction, and because it leads to the formation of an organized lexico-semantic network in which different words and their meanings are linked by relations. For example, our lexico-semantic networks contain links both between words that can be combined to form meaningful utterances (e.g., *eat* and *apple*), and words similar in meaning (e.g., *apple* and *grape*) (Jones, Willits, Dennis, & Jones, 2015). These links are a fundamental facet of our lexico-semantic knowledge that influence behavior even without awareness, reasoning, or recall of relevant information from episodic memory (as is evident from phenomena such as priming). How do the new words we encounter become integrated into our lexico-semantic networks?

The purpose of the present research is to investigate the rapid integration of new words into existing lexico-semantic networks purely on the basis of regularities with they are distributed with other words in linguistic input. As demonstrated by the seminal work of Landauer and Dumais (1997) and many subsequent modeling efforts (Fremann & Lapata, 2015; Huebner & Willits, 2018; Jones & Mewhort, 2007; Rohde, Gonnerman, & Plaut, 2004), sensitivity to distributional regularities may represent a powerful mechanism for building lexico-semantic networks. First, links between words that can be combined to form meaningful utterances such as *eat* and *apple* can be formed from the regularity with which they co-occur in language. Critically, although words similar in meaning such as *apple* and *grape* may not reliably co-occur, links between them can also be formed from the regularity with which they *share* each other's patterns of co-occurrence (e.g., *apple* and *grape* may not reliably co-occur, but do share each other's co-occurrence with *eat*, *juicy*, etc.). These distributional regularities are sufficiently abundant in language that mechanistic models that form representations of words purely on the basis of these regularities capture the majority of links present in human lexico-semantic networks (Jones et al., 2015).

In spite of the extensive evidence from modeling research supporting the potential contributions of sensitivity to distributional regularities, we know little about whether exposure to these regularities actually drives the integration of new words into lexico-semantic networks in human learners. Accordingly, the present research was designed to assess whether adults semantically integrate novel words with familiar words after reading sentences rich in distributional regularities. Specifically, we investigated whether familiar words came to be semantically primed not only by novel words with which they co-occurred, but also by novel words with which they never co-occurred, and instead shared patterns of co-occurrence with another word.

In what follows, we first review existing evidence about human learner's sensitivity to distributional regularities. In this review, we highlight the paucity of prior research that is informative about the role of distributional regularities abundant in language in building human lexico-semantic networks. We then present an experiment designed to illuminate this role.

Human Sensitivity to Distributional Regularities in Language

Extensive evidence from statistical learning research suggests that humans are sensitive to some forms of distributional regularities in some modalities. Specifically, numerous studies have revealed that we are sensitive to the regularity with which items such as speech sounds or shapes co-occur, either simultaneously, sequentially, or separated by some number of other items (Conway & Christiansen, 2005; Fiser & Aslin, 2002; Gomez, 2002; Saffran, Johnson, Aslin, & Newport, 1999).

However, this evidence cannot directly illuminate whether distributional regularities of words in language can drive lexico-semantic integration for two reasons. First, very little statistical learning research conducted to date has investigated whether we form links between items that never occur together, and instead share each other's patterns of co-occurrence with other items (to our knowledge, only one study visual domain, Schapiro, Rogers, Cordova, Turk-Browne, & Botvinick, 2013, is suggestive of this form of learning). However, this process is a critical facet of the potential importance of sensitivity to distributional regularities for building lexico-semantic networks, because many words similar in meaning do not reliably co-occur, and can instead only be linked based on their shared patterns of co-occurrence (Jones et al., 2015). Second, statistical learning research has focused on learning links between items in domains that intentionally do not carry meaning, such as speech sounds, acoustic sounds, and shapes, and tactile stimuli. Because statistical learning phenomena vary even across these studied domains (Conway & Christiansen, 2005), it is unclear whether they generalize to the formation of semantic links between novel and familiar words in language.

To our knowledge, the only evidence relevant to the role of distributional regularities in semantic integration comes from a handful of studies conducted by McNeill (McNeill, 1963, 1966). In these studies, novel words were organized into triads, in which one novel word A co-occurred in sentences with either of two other novel words, B and C. Accordingly, the distributional regularities consisted of both the direct co-occurrence of A- B and A-C, and the shared co-occurrence of B-C (which never actually co-occurred, but both co-occurred with A). By administering a free association task at multiple points during sentence reading in which participants were asked to produce the first novel word that came to mind when prompted with another, McNeill observed that participants first formed links between novel words that directly co-occurred (i.e., A-B and A-C), and then between those that shared co-occurrence (i.e., B-C). This finding provides evidence that people can learn the distributional regularities of words in sentences online, as they are experienced. These regularities therefore represent a viable candidate for drivers of semantic integration. However, these studies were not designed to investigate the semantic integration of novel words into existing lexico-semantic networks, because novel words only ever shared distributional regularities with each other, and not with familiar words. Moreover, the use of a free association task to assess learning leaves open the possibility that these links participants apparently formed were based on retrieving the episodic experiences of reading the sentences from memory, rather than on the formation of automatically-activated semantic links. The role of distributional regularities in lexico-semantic integration therefore formed the focus of the present experiments.

Present Experiments

The present experiments were designed to investigate whether distributional regularities can drive the rapid integration of new words into existing lexico-semantic networks. Specifically, participants read sentences in which were embedded triads of words that consisted of a novel pseudoword (e.g., foobly) that regularly preceded a familiar word (e.g., apple) in some sentences, and another novel pseudoword (e.g., mipp) in other sentences. Accordingly, the sentences contained distributional regularities with which a familiar word (e.g., apple) both directly co-occurred with one novel psuedoword (foobly), and shared this pattern of co-occurrence with another (mipp) (Fig. 1). The sentences otherwise contained no information from which the meanings of the novel pseudowords could be derived. For example, participants might read "My sister loves to see a foobly apple" and "I saw a foobly mipp on vacation".

Immediately following a short session of sentence reading, we then assessed lexico-semantic integration by testing whether the familiar word came to be primed by both the novel pseudoword with which it co-occurred, and the novel pseudoword with which it shared this pattern of co-occurrence. To show both patterns of priming, participants must: (1) Learn the novel word forms, (2) Form links between novel and familiar words that directly co-occur, and (3) Derive links between novel and familiar words that never co-occur, but instead share each other's patterns of co-occurrence.

Method

Participants

Participants were 45 undergraduate students from a Midwestern university who received course credit. An additional five participants were excluded due to failure to complete the experiment.

Stimuli and Design

Training. The training stimuli were two triads of words (1: foobly-apple-mipp; 2: dodish-horse-geck) that each consisted of a pseudoadjective (e.g., foobly) that consistently preceded one familiar noun (e.g., apple) and one pseudonoun (e.g., mipp) in different sentences. Each word pair from these triads (foobly-apple, foobly-mipp, dodish-horse, dodish-geck) was embedded in 10 unique sentence frames, for a total of 40 training sentences. These sentences therefore conveyed both direct co-occurrences between words in the same pair from the same triad, and shared co-occurrences between familiar and pseudonouns from the same triad. The sentences did not convey any other cues to pseudoword meaning (Figure 1).

Test. For testing purposes, we added two new pseudowords (nuppical; boff) and 2 pictures: One of an apple and one of a horse.

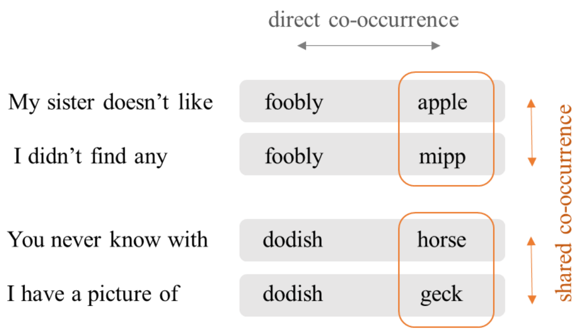

Figure 1: Illustration of training sentence structure.

Using these stimuli, we generated five types of Prime-Target word pairs. Primes were always novel pseudowords, and Targets were always one of the two familiar nouns used during training (apple or horse). First, we generated two types of Related pairs that were consistent with the training triads: Related Direct, in which a pseudoadjective preceded the familiar noun that it had preceded during training (e.g., foobly-apple), and Related Shared, in which a pseudonoun preceded the familiar noun with which it had shared co-occurrence during training (e.g., mipp-apple). Second, we generated corresponding Unrelated Direct and Unrelated Shared pairs in which the Primes from Related pairs were switched, such that they violated the regularities present during training (e.g., foobly-horse). Finally, we generated Neutral pairs, in which the new pseudowords that were only present during Test (nuppical; boff) preceded each familiar noun.

Procedure

The experiment had 2 phases: Training and Test.

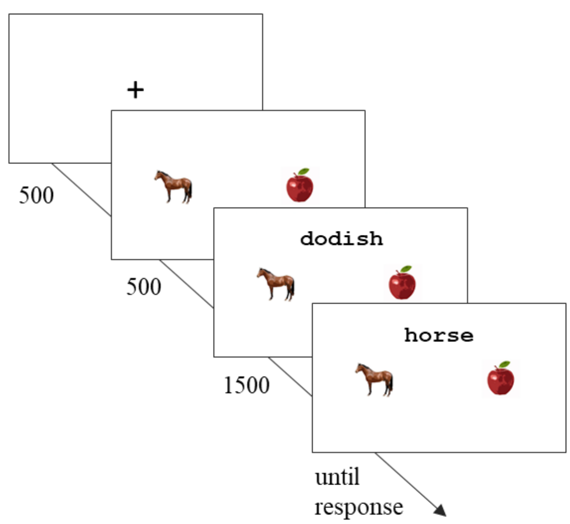

Figure 2: Timing of events during the primed visual search task used in the Test phase.

Training. The Training consisted of three blocks. In each block, participants first read all of the 40 training sentences in a random order at their own pace. To check whether participants were attending to the sentences, three control questions appeared following random sentences in which participants were prompted to type the novel words from the last sentence they had read. The reading component of each block was followed by a free association task in which participants were asked to respond with the first novel (pseudo) word they could think of when prompted with each of the pseudowords from the training sentences. Each of the pseudowords (foobly, dodish, mipp, geck) was presented 3 times in a randomized order.

Test. For the test phase, participants performed a primed visual search task (see Figure 2 for timing of events in trials). At the start of each trial, participants saw a fixation cross followed by two images, one on either side of the screen: A horse, and an apple. Two words (a Prime and Target) were then consecutively presented as text on the top of the screen. Participants' task was to read both words, but choose the image labeled by the second (i.e., Target) word using the mouse. During a practice phase consisting of 8 trials, the two words consisted of Neutral word pairs (i.e., a new pseudoword followed by apple or horse). During the actual task consisting of 144 trials, the two words consisted of Related Direct, Related Shared, Unrelated Direct, Unrelated Shared, and Neutral pairs.

Participants were given an unlimited time to make their responses, but were prompted to respond quickly and were shown a message saying that they were too slow if their response time on a trial was > 800ms.

Results and Discussion

Preliminary analyses: Free association

To test whether participants were attending to the sentences, we analyzed participants' responses on the free association task. Participants responded as instructed by responding with one of the training pseudowords on an average 90.6% of all free association trials. Participants tended to respond with training pseudowords that had directly co-occurred with the prompt pseudoword: 88% of all responses to pseudoadjective prompts were with the noun that followed the pseudoadjective during training, and 77% of responses to pseudonoun prompts were the pseudoadjective that preceded it during training. Only 2.5% of all responses to pseudonouns were based on shared co-occurrence. This confirmed that participants read the sentences and learned the word forms.

Main analyses: Priming[1]

The purpose of the main analyses was to investigate whether the novel pseudowords were semantically

[1] The full pattern of effects on reaction time have also been replicated with two samples (*N*s= 25 and 28) of participants recruited from Amazon Mechanical Turk: Once as an exact replication, and once as a conceptual replication in which

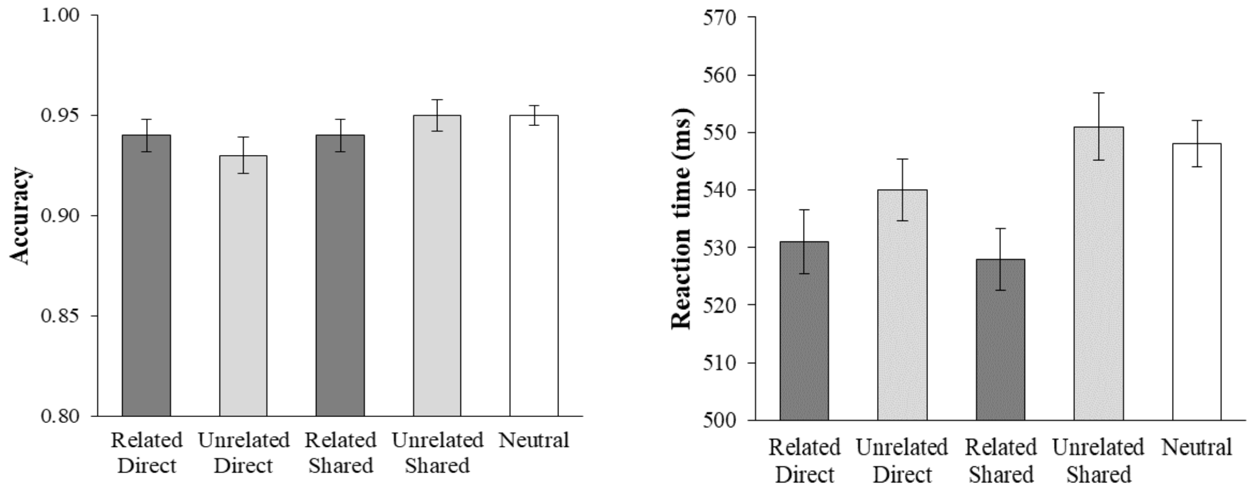

Figure 3: Mean accuracy (left) and reaction times (right) across five conditions. Dark gray bars represent Related (Direct and Shared) conditions, and light gray bars represent Unrelated (Direct and Shared) conditions. The Neutral condition (new pseudoword) is presented in white. Error bars indicate the standard errors of the means.

integrated with familiar words with which they shared distributional regularities (i.e., direct or shared co-occurrence) in training sentences. We accomplished this investigation by measuring whether the novel pseudowords affected the speed and accuracy of processing familiar words in the priming task used during the Test phase. Specifically, we compared the speed and accuracy with which participants identified whether the Target word was apple or horse when it was preceded by a novel pseudoword in the Related Direct, Related Shared, Unrelated Direct, Unrelated Shared, and Neutral conditions. Related pseudowords were expected to facilitate Target word identification, whereas Unrelated pseudowords were expected to inhibit identification. Moreover, these facilitation and inhibition effects may be greater for Direct versus Shared co-occurrences.

Prior to conducting this analysis, we first eliminated data from 8 participants with extremely short reaction times (more than 2/3rds of RTs < 100ms), leaving a sample size of 38 participants. Accuracies and Reaction Times are presented in Figure 3.

Accuracy. We analyzed effects on accuracy using a linear mixed effects regression model in which Relatedness (Related vs Unrelated) and Type (Direct vs Shared) were fixed effects, and Participant was a random effect. This model revealed no effect of either Relatedness or Type on accuracy (Relatedness: B = - 0.004, SE = 0.008, t = - 0.55, p = .59, d = 0.004, Type: B = - 0.012, SE = 0.008, t = - 1.48, p = .15, d = 0.012).

Reaction Time. For analyses of reaction time, we removed data from incorrect trials, and trials with extremely short (<100 ms) and extremely long response latencies (>1500 ms), resulting in removal of 8.1 % of trials.

We then generated a linear mixed-effects model with Relatedness (related; unrelated) and Type (direct; shared) as fixed effect factors and Participants as a random effect. This model revealed no significant effect of Type (neither as a main effect nor in interaction with Relatedness). Thus, Type was excluded from the final model. A log-likelihood ratio test indicated that the best fitting random effects structure included only a random intercept for participants. Thus, the final model included Relatedness as a fixed effect factor and a random intercept for participants. This model revealed a significant effect of Relatedness on reaction times, B = 14.82, SE = 5.10, t = 2.91, p < .01, d = 0.096 (see Brysbaert & Stevens, 2018 for effect size estimate approach). Participants were 14.9 ms faster in related than in unrelated conditions (Figure 3, right panel). The model explained 16% of total variance (R-squared based on Nakagawa & Schielzeth, 2013).

The follow-up analyses compared Related and Unrelated conditions to the Neutral condition. A linear mixed-effects model with Condition (Neutral; Related Direct, Related Shared, Unrelated Direct, Unrelated Shared) as a fixed effect factor and a random intercept for Participants revealed that only Related conditions were significantly different than the Neutral (Related Direct: B = - 16.11, SE = 6.30, t = - 2.56, p = .01, d = .104; Related Shared: B = -16.56, SE = 6.32, t = - 2.62, p < .01, d = .104). There was no significant difference in RT between the Neutral condition and Unrelated conditions. In other words, participants were faster to respond when the Target was preceded by a pseudoword that either directly co-occurred with the Target (Related Direct) or shared the pattern of co-

the pseudoadjectives were changed from foobly/dodish to foobing/doding.

occurrence (Related Shared) than when it was preceded by a new pseudoword that only appeared in the Test and not the Training phase. Primes that were incongruent with the regularities presented during the training (Unrelated Direct, Unrelated Shared) did not affect speed.

General Discussion

The present experiment provides a novel demonstration that new words can be rapidly integrated into existing lexico-semantic networks based on the distributional regularities of words in sentences. Specifically, immediately following a short session of sentence reading, familiar words came to be primed by both novel words with which they co-occurred in sentences, and novel words with which they never co-occurred, but instead shared a pattern of co-occurrence with another novel word. Given that these distributional co-occurrence regularities are ubiquitous in language (Jones et al., 2015), the present results provide evidence that sensitivity to these regularities may represent a critical way in which new words are rapidly integrated into lexico-semantic knowledge.

Implications for Lexico-Semantic Integration

The present findings build upon prior research in two key ways. First, prior evidence about the potentially powerful contributions of distributional regularities to building lexico-semantic networks comes primarily from modeling research. The present findings therefore substantially underline this potential by demonstrating that new words can be added to actual human lexico-semantic networks through mere exposure to distributional regularities.

Second, this evidence also adds to our understanding of how rapidly new words can be integrated into our existing lexico-semantic networks. Specifically, extensive prior research has investigated the lexico-semantic integration of novel words through different kinds of input, such as studying definitions of novel words, or repeatedly observing words co-occurring with images of specific familiar objects (Breitenstein, Zwitserlood, de Vries et al., 2007; Clay, Bowers, Davis, & Hanley, 2007; Dagenbach, Horst, & Carr, 1990; Dobel, Junghöfer, Breitenstein et al., 2010; Tamminen & Gaskell, 2013). Much of this research has suggested that newly learned words are only gradually integrated into existing lexico-semantic networks, following at least one day and up to several weeks of consolidation. In contrast, a handful of recent findings (Borovsky, Elman, & Kutas, 2012; Mestres-Missé, Rodriguez-Fornells, & Münte, 2006; Zhang, Ding, Li, & Yang, 2019) have suggested that lexico-semantic integration of novel words can occur more rapidly when learning is driven by reading sentences in which novel words appear in a position typically occupied by a specific, familiar word (e.g., “It was a windy day, so Peter went to the park to fly his *dax*”). The present findings add to this evidence that novel words can be integrated into existing lexico-semantic networks very rapidly, immediately following an initial learning experience.

Future Directions

The evidence provided by the present experiment highlights a new avenue for future research to investigate *how* distributional regularities foster semantic integration. For example, do direct co-occurrences foster integration more rapidly than shared co-occurrences, or do these processes unfold in parallel? This question could be addressed by measuring integration (e.g., using the priming approach taken in the present experiment) at multiple points throughout training. Moreover, addressing this and related questions could help to generate and arbitrate between different potential mechanistic accounts of distributional regularity-driven semantic integration.

Summary

Throughout our lives, we amass a sizable and interconnected body of knowledge of words and their meanings. The present research highlights how the formation of these lexico-semantic networks may be critically facilitated by the rapid integration of new words via sensitivity to the regularities with which words occur with other words in linguistic input.

References

Borovsky, A., Elman, J. L., & Kutas, M. (2012). Once is enough: N400 indexes semantic integration of novel word meanings from a single exposure in context. *Language Learning and Development, 8*, 278-302.

Breitenstein, C., Zwitserlood, P., de Vries, M. H., Feldhues, C., Knecht, S., & Dobel, C. (2007). Five days versus a lifetime: Intense associative vocabulary training generates lexically integrated words. *Restorative Neurology and Neuroscience, 25*, 493-500.

Brysbaert, M., & Stevens, M. (2018). Power analysis and effect size in mixed effects models: A tutorial. *Journal of Cognition, 1*, 9.

Clay, F., Bowers, J. S., Davis, C. J., & Hanley, D. A. (2007). Teaching adults new words: the role of practice and consolidation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 970-976.

Conway, C. M., & Christiansen, M. H. (2005). Modality-constrained statistical learning of tactile, visual, and auditory sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 24.

Dagenbach, D., Horst, S., & Carr, T. H. (1990). Adding new information to semantic memory: How much learning is enough to produce automatic priming? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 581-591.

Dobel, C., Junghöfer, M., Breitenstein, C., Klauke, B., Knecht, S., Pantev, C., & Zwitserlood, P. (2010). New names for known things: on the association of novel word forms with existing semantic information. *Journal of Cognitive Neuroscience, 22*, 1251-1261.

Fiser, J., & Aslin, R. N. (2002). Statistical learning of new visual feature combinations by infants. *Proceedings of the National Academy of Sciences, 99*, 15822-15826.

Frermann, L., & Lapata, M. (2015). Incremental Bayesian Category Learning From Natural Language. *Cognitive Science, 40,* 1333–1381.

Gomez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science, 13*, 431-436.

Huebner, P. A., & Willits, J. A. (2018). Structured semantic knowledge can emerge automatically from predicting word sequences in child-directed speech. *Frontiers in Psychology, 9*.

Jones, M. N., & Mewhort, D. J. (2007). Representing word meaning and order information in a composite holographic lexicon. *Psychological review, 114*, 1-37.

Jones, M. N., Willits, J., Dennis, S., & Jones, M. (2015). Models of semantic memory. In J. Busemeyer & J. Townsend (Eds.), *Oxford Handbook of Mathematical and Computational Psychology* (pp. 232-254). New York, NY: Oxford University Press.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review, 104*, 211.

McNeill, D. (1963). The origin of associations within the same grammatical class. *Journal of Verbal Learning and Verbal Behavior, 2*, 250-262.

McNeill, D. (1966). A study of word association. *Journal of Memory and Language, 5*, 548.

Mestres-Missé, A., Rodriguez-Fornells, A., & Münte, T. F. (2006). Watching the brain during meaning acquisition. *Cerebral Cortex, 17*, 1858-1866.

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, *4*, 133-142.

Rohde, D. L., Gonnerman, L. M., & Plaut, D. C. (2004). An improved method for deriving word meaning from lexical co-occurrence. *Cognitive Psychology, 7*, 573-605.

Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, *70*, 27-52.

Schapiro, A. C., Rogers, T. T., Cordova, N. I., Turk-Browne, N. B., & Botvinick, M. M. (2013). Neural representations of events arise from temporal community structure. *Nature Neuroscience*, *16*, 486-492.

Schmitt, N., & McCarthy, M. (1997). *Vocabulary: Description, acquisition and pedagogy*: Cambridge University Press.

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

Zhang, M., Ding, J., Li, X., & Yang, Y. (2019). The impact of variety of episodic contexts on the integration of novel words into semantic network. *Language, Cognition and Neuroscience, 34*, 214-238.