Do Young Children Learn Words from the Company they Keep?

Layla Unger (layla.unger@york.ac.uk)
Department of Psychology, University of York &
Department of Psychology, Ohio State University

Emma James (emma.james@york.ac.uk)
Department of Psychology, University of York

Abstract

The challenge of early word learning is often framed as one of individually mapping words to their referents. Yet children do not experience words just as individual labels, but as parts of broader language contexts, such as conversations and stories. In principle, word contexts might support word learning because words similar in meaning tend to occur in similar contexts. Thus, a child who knows some fruit words and has heard them in the context of "juicy" might learn that a "juicy mango" is also a fruit even without ever seeing a mango. Although children can use such contextual support to learn words in the lab, we do not know whether they harness contextual support in everyday language for real-world word learning. We quantified words' contextual support in children's everyday language input and found that it reliably predicted normative word learning, even accounting for other established predictors such as word frequency.

Keywords: word learning; distributional semantics; language models

Introduction

Humans are unique among animals in acquiring thousands of communicative signals – i.e., words – that express meaning. The feat of word learning is often framed as one of individually mapping words to their meanings. Indeed, it is exceedingly common for research on word learning to introduce the topic using Quine's (1960) *gavagai* example, in which the challenge of learning a word is to map it onto its intended referent. Although mapping is certainly useful for anchoring some words to their real-world counterparts (for a striking recent demonstration, see Vong, Wang, Orhan, & Lake, 2024), children do not just experience words as individual labels for things. Instead, children overwhelmingly experience words with other words, in conversations, stories, books and other media. How does the way children experience words in language shape word learning?

In what follows, we first review our understanding of how word learning is shaped by characteristics of the way words occur in language, such as a word's frequency. In this review, we highlight that these characteristics might predict word learning (e.g., more frequent words tend to be learned earlier in development) but do not provide a route for learning what words mean (e.g., merely encountering a word frequently does not offer information about what it means). We then highlight and examine a potentially vital route through which children might learn words from the way they occur with other words.

Language Characteristics that Contribute to Word Learning

To date, research has uncovered key characteristics of children's everyday experiences with words that may contribute in some way to word learning. For example, children learn words earlier in development (and thus more easily) when words are more frequent and tend to occur in short utterances (Swingley & Humphrey, 2018; Braginsky, Yurovsky, Marchman, & Frank, 2019). Moreover, there is an active debate surrounding the contribution to word learning of *contextual diversity*: whether words are learned more readily when they tend occur with words relating to a a diversity of topics (e.g., occur with words related to mealtimes, playtimes, bedtime, and so on) or a consistent topic (e.g., consistently accompanied by words related to mealtimes) (Hills, Maouene, Riordan, & Smith, 2010; Roy, Frank, DeCamp, Miller, & Roy, 2015; Unger, Chang, Savic, Bergen, & Sloutsky, 2024).

Yet, none of these characteristics directly relate to learning what words mean. For example, hearing a word more frequently does not itself convey information about what it means. Frequent occurrences might help familiarize a child with the *form* of a word. A word's frequency might also *indirectly* capture routes for learning what it means, such opportunities to map it to a referent. Likewise, a word's tendencies to occur in short utterances or in diverse versus consistent contexts do not directly reflect routes for learning what it means.

Learning Words from the Company they Keep?

A prime candidate for learning word meanings from the way they occur in language comes from an idea often described with the phrase, "You shall know a word by the company it keeps" (Firth, 1957). More formally, this *distributional hypothesis* posits that words similar in meaning tend to occur in similar surrounding contexts of other words. For example, different words for fruits such as "apple" and "orange" tend to occur in the context of "juicy" and "sweet", and different words for emotions such as "happy" and "sad" tend to occur in the context of "feel" and "face". Multiple analyses of language point to these distributional regularities in language (Miller & Charles, 1991; Rubenstein & Goodenough, 1965). Further evidence for the existence of these regularities comes from numerous computational models that form representations of words based on the other words that they tend to co-occur with (Lund & Burgess, 1996; Landauer & Dumais, 1997; Jones & Mewhort, 2007; Mikolov, Chen, Cor-

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society*

rado, & Dean, 2013; Pennington, Socher, & Manning, 2014). Such models form similar representations of words similar in meaning, indicating that words similar in meaning do indeed tend occur in similar contexts. Moreover, this outcome has been observed even when models are trained on relatively small samples of the language that infants and young children day to day, suggesting that these regularities are common in children's language input (Huebner & Willits, 2018).

These regularities could in principle allow children to leverage their existing knowledge of words (perhaps acquired in part from mapping) to learn new words. For example, consider a child who already knows the words "apple" and "orange" and has heard them accompanied by the word "juicy". Just from hearing "I hope the mangoes are juicy this year", a child might get the sense that mangoes are something like apples and oranges, even if they have never encountered mangoes before. Put broadly, children might learn new words by connecting them to known words similar in meaning that occur in similar contexts. For brevity, we refer to this source of word learning as *context leverage*. This idea goes beyond the notion of distributional semantics: it proposes not just that words similar in meaning occur in similar contexts, but moreover that new words can be learned by *inheriting* known meanings from known words that are similar in meaning based on their occurrence in similar contexts.[1]

Model simulations have shown that in principle, it is possible to accurately infer the meanings and referents of new words based on their contextual similarity with similar known words (Johns & Jones, 2012; Riordan & Jones, 2011; Aho, Roads, & Love, 2022). Moreover, from researcher-designed language input in the lab, school-age children (Maguire et al., 2018) and possibly toddlers (Yuan, Fisher, Kandhadai, & Fernald, 2011; Sullivan & Barner, 2016; Ferguson, Graf, & Waxman, 2018; Syrett, LaTourrette, Ferguson, & Waxman, 2019) can use context leverage to learn words. Yet, we do not know whether context leverage contributes to real-world word learning in childhood.

Present Study

The goal of the present study is to investigate whether context leverage is an important source of early word knowledge. Critically, although there is an overall tendency for words similar in meaning to occur in similar contexts (e.g., Landauer & Dumais, 1997; Huebner & Willits, 2018), this tendency may be stronger for some words than others. If context leverage is indeed an important source of word knowledge, then the availability of context leverage support for learning a word should predict how easy it is for children to learn that word. Specifically, the likelihood that a child will learn a new word should depend in part on the new word's tendency to occur in similar contexts to words similar in meaning that the child already knows. For example, a child should be more likely to learn a new word for a fruit if it tends to occur in similar contexts to earlier-learned fruit words.

The present study tested this prediction. Central to this test was to quantify the amount of context leverage support that is available for learning different words in children's everyday language input. We developed two metrics of context leverage support: one explicitly designed to capture the notion of context leverage, and one derived from the models described above that use distributional regularities in language to form representations of words. We then tested whether the greater availability of context leverage support predicts earlier word learning. To disentangle context leverage from other key predictors of word learning, we tested whether this relationship held above and beyond characteristics such as word frequency.

Methods

Word Learning Outcome: Age of Acquisition

The present study examined predictors of word learning in childhood. Currently, the best and most commonly used normative measure of early word learning comes from the widely administered MacArthur-Bates Communicative Index (MCDI), which is a checklist in which parents indicate the words that their child knows (Fenson et al., 2007; Frank, Braginsky, Yurovsky, & Marchman, 2016) (see the General Discussion for alternative measures that may become available in future). We focused on data from the Words and Sentences form of the checklist for 16 - 30 month-old children, in which parents indicate whether their child produces each word. To estimate word learning, for each word, we fit a logistic regression to estimate changes in the likelihood that children know the word with age. We then calculated the age of acquisition (AoA) as the age at which at 50% of children are likely to know a word.

Corpus of Children's Everyday Language

Characteristics of the way words occur in language were measured from transcripts of children's everyday language experiences in the CHILDES database (MacWhinney, 2000). Prior to measuring these characteristics, we first followed common corpus pre-processing procedures, including removing punctuation and standardizing morphological variants (e.g., singular and plural forms of nouns and tenses of verbs) to a single word form using the textstem package (Rinker, 2018) in the R environment. To focus on language input available to the child, we removed the child's own utterances.

Context Leverage Metrics

We developed two metrics of context leverage support available in children's everyday language. Both metrics followed the same logic. First, each metric used a different approach to measure the degree to which any two words occurred in similar contexts. For each word, we then contrasted: (A) its

[1]Note that the concept of context leverage is not explicitly related to the similar-sounding concept of contextual diversity. As noted above, contextual diversity simply refers to the occurrence of *individual* words in diverse versus consistent contexts. In contrast, context leverage refers to the meaning that can be gleaned from the occurrence of words in similar contexts to other, similar words.

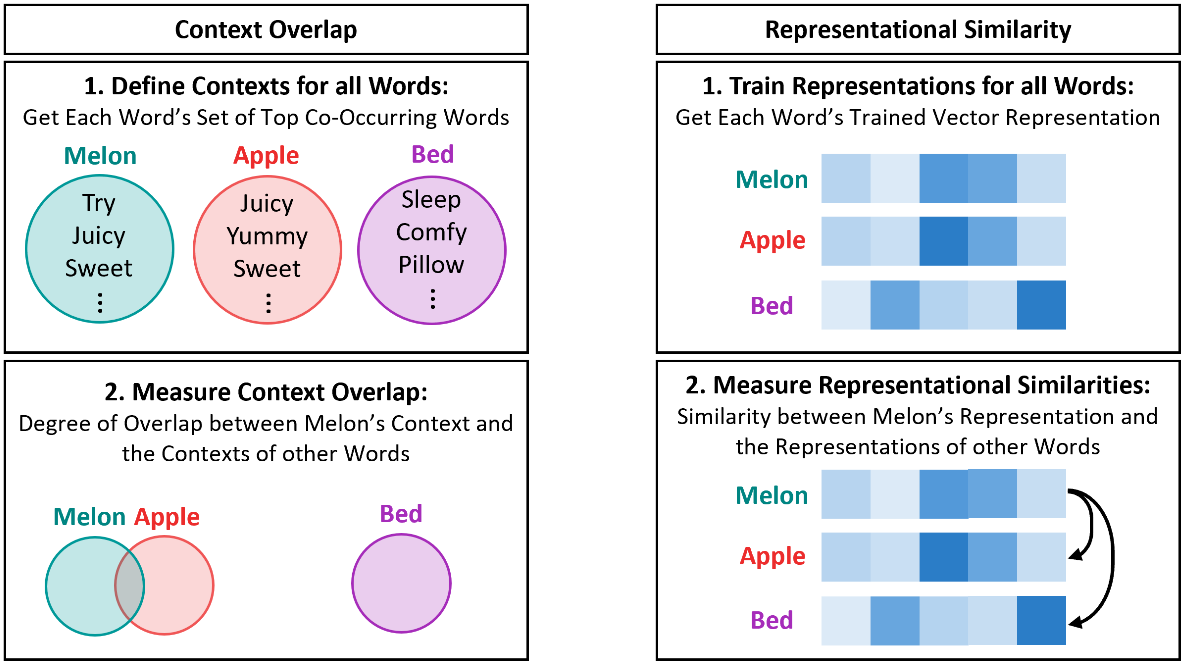

Figure 1: Calculation of Context Overlap and Representational Similarity illustrated for the word *melon*.

tendency to occur in similar contexts to earlier-learned words similar in meaning, versus (B) its tendency to occur in similar contexts to earlier-learned words different in meaning. This contrast quantified the amount of context leverage support for learning a word. Here, we first describe how each metric measured context similarity between words. For both metrics, we then describe how we compared a word's contextual similarity to earlier learned words that are similar versus different in meaning.

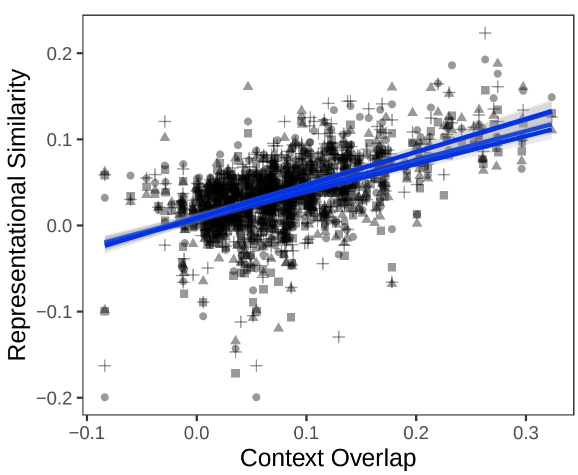

Figure 2: Relationship between context leverage measured from Context Overlap and Representational Similarity. Representational Similarity was derived from four randomly initialized models, which each correspond to a different symbol in the plot. The metric values derived from the randomly initialized models yielded convergent results.

Context Similarity Measure: Context Overlap

Context Overlap was designed to explicitly capture the degree to which any two words occur in similar contexts. We calculated Context Overlap using a series of steps illustrated in Figure 1.

The first step was to take each word in the children's language corpus, and get the context of other words that it tends to reliably occur with. To get the context of words that each word tends to co-occur with, we calculated the reliability with which different words in the language corpus co-occurred with each other separated by up to 10 intervening words. Critically, we did not simply calculate the frequency of word co-occurrence, because words that are overall more frequent are also likely to have frequent co-occurrences with other words just by chance. To instead calculate co-occurrence regularities, we calculated positive pointwise mutual information (PPMI) between words, which captures how likely a given pair of words are to occur together above and beyond the frequency that would be expected by chance (Evert, 2008). We then defined the context of each word as the other words that it most reliably co-occurred with. For example, the context of "melon" might include "try", "juicy" and "sweet", the context of "bed" might include "sleep", "pillow" and "comfy", and so on. Each word's context was thus a list of the top 100 words it most regularly co-occurred with (note: lists of different sizes yielded similar results, though prediction of Age of Acquisition was weaker when lists contained fewer than 25 words).

As illustrated in Figure 1, different words vary in the degree to which their contexts overlap. For example, the contexts "melon" and "apple" includes some overlaps such as "juicy" and "sweet". The second step of calculating the Con-

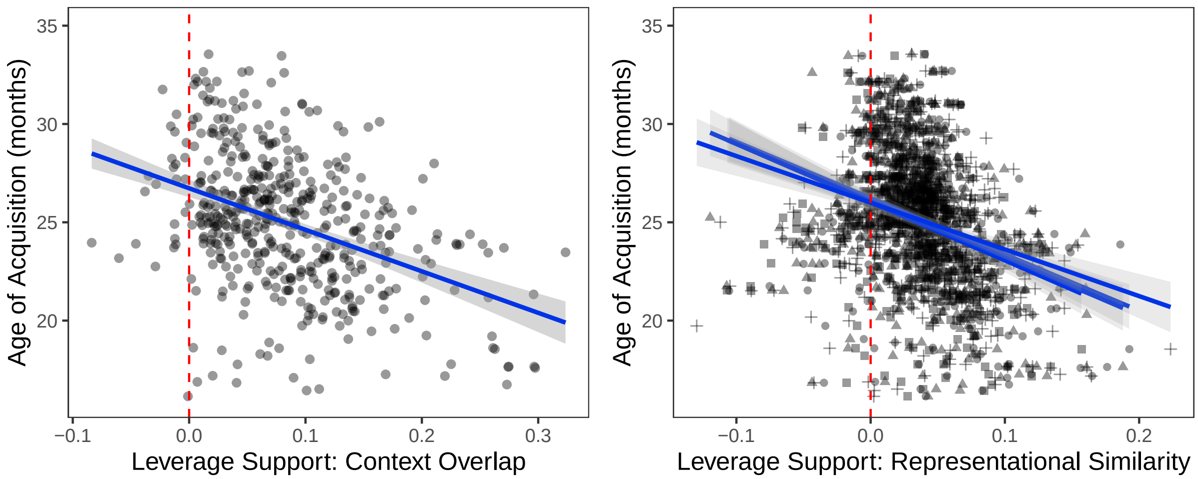

Figure 3: Relationship between a words' Age of Acquisition and leverage support measured from Context Overlap (left) and Representational Similarity (right). The plot depicting the relationship with Representational Similarity metric combines the metrics taken from four randomly initialized models (depicted as different symbols), which revealed similar results.

text Overlap metric was to quantify this overlap. Specifically, for each pair of words, we calculated the Jaccard Index of their contexts, which is a common measure of overlap between sets. This measure is simply the number of words that occurred in both words' contexts (the overlap, e.g., "juicy" and "sweet") divided by the total number of words in the contexts. It is worth noting that more nuanced measures of overlap in which different context words were given different weights (e.g., based on PPMI) did not yield stronger predictions of AoA.

Context Similarity Measure: Representational Similarity

Representational Similarity is a measure of contextual similarity drawn from the computational modeling research described above, in which models are trained to form representations of words based on the way they occur with other words. Of the several such models that have been developed, we focused on the Recurrent Neural Network models (Elman, 1990) used by Huebner and Willits (2018) who demonstrated that RNN models trained on child language input form similar representations of words that are similar in meaning.

A full description of these models lies beyond the scope of this work, so we provide a high-level description here. To give an intuitive sense, the RNN is like a listener who hears the language input that children hear, and uses the words it has heard recently to try to predict the words it will hear next. Training involves learning to make more accurate predictions. In practice, words are represented as vectors of numbers, so that the vectors representing the words that have just been "heard" are transformed into an internal representation that is used to predict the word that will occur next. Incorrect predictions update the way that the input word vectors are transformed into internal representations to increase the likelihood of correct future predictions. Critically, this means that the internal representation of a word starts off as random before training, and then becomes shaped by the way the word tends to occur with other words with training. Thus, the trained internal representation of a word reflects its context.

Following Huebner and Willits (2018), we trained RNNs on the child language corpus. For generality, we trained 4 randomly initialized models. From each trained model, we obtained its internal representations of words in the corpus. To capture the degree to which words occurred in similar contexts, we used the commonly used cosine similarity measure to measure similarity between words' internal representations.

Context Leverage

Context Overlap and Representational Similarity each quantify the degree to which words occur in similar contexts. We used these to estimate context leverage support for word learning: i.e., the degree to which a word occurs in similar contexts to earlier-learned words similar in meaning. Here, we treated similarity in meaning as binary. Specifically, we identified words as similar in meaning if they belonged to the same "semantic category", such as animals, body parts or clothing, and different in meaning if they did not. To identify a word's semantic category, we started with the semantic categories identified for words from the developers of the MCDI. However, many of these categories are very broad, such as the category of "descriptive words" that includes adjectives for color, size, weight, and other attributes. We therefore refined these categories, and validated the new categories against adult ratings of similarities in meaning between words in the same category versus words in different categories.

Control Variables

To quantify context leverage support for word learning, we used the Context Overlap and Representational Similarity measures to perform a simple comparison. Specifically, for

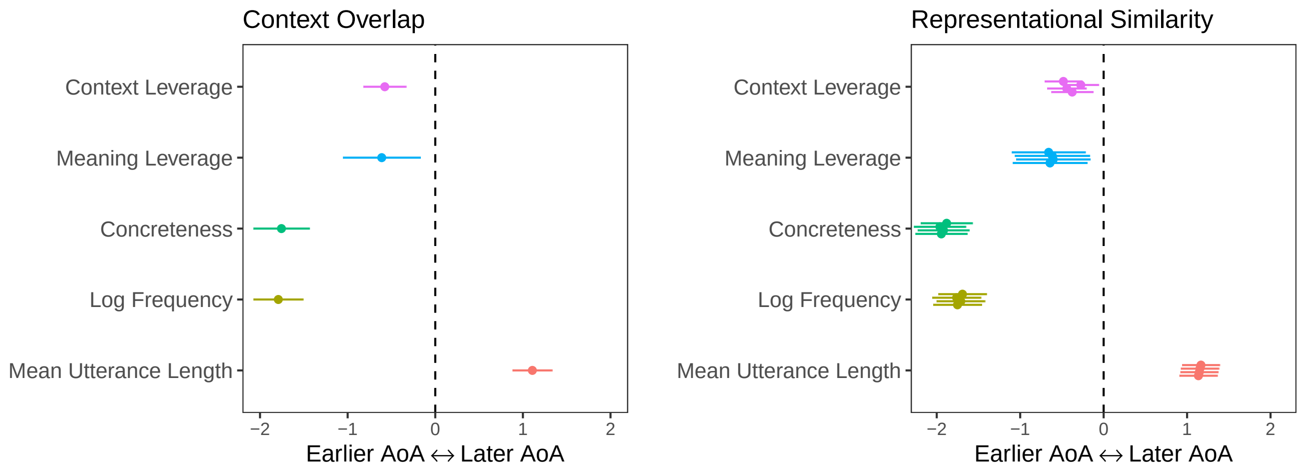

Figure 4: Standardized coefficients for the prediction of Age of Acquisition from context leverage and the control variables. All variables significantly predicted Age of Acquisition.

each word, we calculated the difference between: (A) its mean contextual similarity with earlier-learned words *similar* in meaning, versus (B) its mean contextual similarity with earlier-learned words *different* in meaning. Larger positive values indicate a stronger tendency to occur in similar contexts to earlier-learned words similar in meaning.

To evaluate whether the two metrics converge on the same construct, we examined the relationship between them (see Figure 2). Context Overlap correlated with Representational Similarity derived from each of the 4 trained models at a consistent level of $r \approx 0.55$ (r range = 0.53-0.56, all $ps < .0001$), suggesting convergence between these metrics.

To disentangle any contribution of context leverage to word learning from other contributors, we measured other characteristics that have been associated with Age of Acquisition (AoA), including a word's log-transformed frequency and the mean length of utterances in which a word occurs. In addition, we also included a measure of word concreteness (Brysbaert, Warriner, & Kuperman, 2014). Finally, it is possible that the context leverage metrics in part reflect the degree to which a word is itself similar in meaning to earlier-learned words, rather than its occurrence in similar contexts to such earlier-learned words. Indeed, there is evidence that children are more likely to learn a new word that is similar in meaning to words they already know (Borovsky, Ellis, Evans, & Elman, 2016). Therefore, for each word, we also calculated the proportion of earlier-learned words that are similar in meaning out of all earlier-learned words. For brevity, we refer to this variable as meaning leverage.

Results

The main goal of analyses was to test whether the two metrics of context leverage support predicted word learning above and beyond other predictors identified in prior research (e.g., Braginsky et al., 2019). The relationship between the context leverage metrics and AoA is shown in Figure 3. As depicted in Figure 3, more context leverage support was associated with earlier AoA for both metrics.

To test whether this relationship held above and beyond other predictors of word learning, analyses were conducted as linear models that each predicted word AoA based on one of the context leverage metrics and all control variables. All predictors were scaled prior to analyses. The standardized coefficients are depicted in Figure 4. All variables were significantly associated with AoA. These results replicate prior findings that words are learned earlier when they are frequent, more concrete, occur in shorter utterances, and are similar in meaning to words that are already known. Critically, stronger context leverage support further predicts earlier word learning in addition to these other predictors.

General Discussion

The goal of the present study was to illuminate how word learning is shaped by the way children experience words in language. We particularly focused on a context leverage route through which children might pick up word meanings by connecting new words to similar words they already know that occur in similar contexts. If this route does indeed contribute to word learning, then words should tend to be learned earlier when there is more support for learning via this route in children's everyday language. Specifically, a word should tend to be learned earlier in development when it has a stronger tendency to occur in similar contexts to similar words that are already known. We developed two metrics to test this possibility. We observed that the availability of context leverage support did indeed predict earlier word learning. Moreover, this prediction held while controlling for the contributions of other strong predictors of word learning, such as the frequency with which children hear words in everyday language.

These findings add new weight to the decades-old distributional hypothesis that the company of other words that a word keeps is informative about its meaning. Numerous analyses (Rubenstein & Goodenough, 1965; Miller & Charles, 1991), models (Lund & Burgess, 1996; Landauer & Dumais, 1997; Mikolov et al., 2013; Pennington et al., 2014) and simulations (Johns & Jones, 2012; Riordan & Jones, 2011; Aho et al., 2022) have pointed to the availability of this information in language, including the language children hear day-to-day (Huebner & Willits, 2018). Moreover, young children show some ability to use this information to learn words in the lab (Yuan et al., 2011; Sullivan & Barner, 2016; Ferguson et al., 2018; Syrett et al., 2019). Our results provide evidence that this information is further harnessed in real-word,early word learning.

Open Questions

The present study provides initial evidence that children learn words from the company they keep. A primary motivation for this study is that this source of information provides a plausible route for learning word meanings by connecting new words to words similar in meaning that are already known. For example, a child might learn that the meaning of "mango" is something like the meaning of "apple", "orange" and "strawberry" just from encountering it with other words like "juicy". However, the measure of word learning used in the present study does not directly capture children's knowledge of word meanings. Instead, we measured word learning from word ages of acquisition taken from an instrument in which parents indicate the words their child produces. This measure is useful because it is the most widely administered (and thus normative) index of early word learning available. However, emerging efforts to compile data from numerous prior eye tracking studies of word recognition may yield more informative measures in future (Zettersten et al., 2022). In these studies, gaze is tracked while participants hear a word (e.g., "apple") and see a pictures of both the word's referent and a distractor (e.g., an apple and a horse). Participants' tendency to accurately look at the labelled picture may provide a more direct window onto their knowledge of what a word means. To date, the compiled dataset and analysis techniques may be too preliminary to reliably detect effects of corpus-derived predictors of individual word learning. However, this ongoing effort may provide a rich measure of normative word learning in future.

The present study also does not provide insight into the learning processes involved in picking up words from their surrounding contexts. For example, the RNNs used to derive one of the metrics of context leverage support use a learning process in which the model learns to use recently "heard" words to predict upcoming words. However, the degree to which children learn via similar mechanisms or even engage in prediction during language processing is a topic of active debate (e.g., Mani & Huettig, 2012). Here, we only use such models as a way to measure the contextual leverage support available in language, and do not assert that children use similar learning mechanisms to harness this information. Other facets of the learning process remain similarly open for future research, such as the time course of learning. For example, do children get a strong sense of what a word might mean just from hearing it once in an informative context, or is this sense built up from repeated exposures to contextual leverage support in everyday language?

Finally, in principle, context leverage could allow a learner to pick up the meaning of a word in the absence of a real-world counterpart. At the same time, context leverage could also allow a learner to identify a real-world counterpart from several potential candidates (Sullivan & Barner, 2016). For example, hearing "juicy mangoes" could help a learner infer that mangoes are fruits even when there are no fruits present. Alternatively, when there are both mangoes and other potential referents visible at the same time, "juicy mangoes" could help a learner infer that "mangoes" refers to the fruits. The present study does not disentangle these potential contributions of context leverage to children's word learning.

Acknowledgments

This work was supported by National Institutes of Health Grant R01HD111458 awarded to LU.

References

Aho, K., Roads, B. D., & Love, B. C. (2022). System alignment supports cross-domain learning and zero-shot generalisation. *Cognition*, *227*, 105200.

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

Braginsky, M., Yurovsky, D., Marchman, V. A., & Frank, M. C. (2019). Consistency and variability in children's word learning across languages. *Open Mind*, *3*, 52–67.

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known english word lemmas. *Behavior research methods*, *46*, 904–911.

Elman, J. L. (1990). Finding structure in time. *Cognitive science*, *14*(2), 179–211.

Evert, S. (2008). Corpora and collocations. *Corpus linguistics. An international handbook*, *2*, 1212-1248.

Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., Reznick, J. S., & Bates, E. (2007). *Macarthur-bates communicative development inventories*. Paul H. Brookes Publishing.

Ferguson, B., Graf, E., & Waxman, S. R. (2018). When veps cry: Two-year-olds efficiently learn novel words from linguistic contexts alone. *Language Learning and Development*, *14*(1), 1–12.

Firth, J. R. (1957). A synopsis of linguistic theory, 1930-1955 [Book Section]. In J. R. Firth (Ed.), *Studies in linguistic analysis.* Basil Blackwell.

Frank, M. C., Braginsky, M., Yurovsky, D., & Marchman, V. A. (2016). Wordbank: An open repository for developmental vocabulary data. *Journal of Child Language*, *44*, 677-694.

Hills, T. T., Maouene, J., Riordan, B., & Smith, L. B. (2010). The associative structure of language: Contextual diversity in early word learning. *Journal of Memory and Language*, *63*(3), 259-273.

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*.

Johns, B. T., & Jones, M. N. (2012). Perceptual inference through global lexical similarity. *Topics in Cognitive Science*, *4*(1), 103-120.

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

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*(2), 211-240.

Lund, K., & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instruments, & Computers*, *28*(2), 203-208.

MacWhinney, B. (2000). *The CHILDES project: The database* (Vol. 2). Psychology Press.

Maguire, M. J., Schneider, J. M., Middleton, A. E., Ralph, Y., Lopez, M., Ackerman, R. A., & Abel, A. D. (2018). Vocabulary knowledge mediates the link between socioeconomic status and word learning in grade school. *Journal of Experimental Child Psychology*, *166*, 679–695.

Mani, N., & Huettig, F. (2012). Prediction during language processing is a piece of cake—but only for skilled producers. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(4), 843.

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.

Miller, G. A., & Charles, W. G. (1991). Contextual correlates of semantic similarity. *Language and Cognitive Processes*, *6*(1), 1-28.

Pennington, J., Socher, R., & Manning, C. D. (2014). Glove: Global vectors for word representation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (emnlp)* (p. 1532-1543).

Quine, W. V. O. (1960). *Word and object (studies in communication)*. Technology Press of MIT.

Rinker, T. W. (2018). textstem: Tools for stemming and lemmatizing text [Computer software manual]. Buffalo, New York. Retrieved from `http://github.com/trinker/textstem` (version 0.1.4)

Riordan, B., & Jones, M. N. (2011). Redundancy in perceptual and linguistic experience: Comparing feature-based and distributional models of semantic representation. *Topics in Cognitive Science*, *3*(2), 303-345.

Roy, B. C., Frank, M. C., DeCamp, P., Miller, M., & Roy, D. (2015). Predicting the birth of a spoken word. *Proceedings of the National Academy of Sciences*, *112*(41), 12663–12668.

Rubenstein, H., & Goodenough, J. B. (1965). Contextual correlates of synonymy. *Communications of the ACM*, *8*(10), 627-633.

Sullivan, J., & Barner, D. (2016). Discourse bootstrapping: Preschoolers use linguistic discourse to learn new words. *Developmental Science*, *19*(1), 63–75.

Swingley, D., & Humphrey, C. (2018). Quantitative linguistic predictors of infants' learning of specific english words. *Child Development*, *89*(4), 1247–1267.

Syrett, K., LaTourrette, A., Ferguson, B., & Waxman, S. R. (2019). Crying helps, but being sad doesn't: Infants constrain nominal reference online using known verbs, but not known adjectives. *Cognition*, *193*, 104033.

Unger, L., Chang, T., Savic, O., Bergen, B. K., & Sloutsky, V. M. (2024). When is a word in good company for learning? *Developmental Science*, e13510.

Vong, W. K., Wang, W., Orhan, A. E., & Lake, B. M. (2024). Grounded language acquisition through the eyes and ears of a single child. *Science*, *383*(6682), 504–511.

Yuan, S., Fisher, C., Kandhadai, P., & Fernald, A. (2011). You can stipe the pig and nerk the fork: Learning to use verbs to predict nouns. In *Proceedings of the 35th annual boston university conference on language development* (pp. 665–677).

Zettersten, M., Yurovsky, D., Xu, T. L., Uner, S., Tsui, A. S. M., Schneider, R. M., ... others (2022). Peekbank: An open, large-scale repository for developmental eye-tracking data of children's word recognition. *Behavior Research Methods*, 1–16.