Does Contextual Diversity Hinder Early Word Acquisition?

Tyler A. Chang (tachang@ucsd.edu)
Department of Cognitive Science
Halıcıoğlu Data Science Institute
UC San Diego, La Jolla, CA 92093, USA

Benjamin K. Bergen (bkbergen@ucsd.edu)
Department of Cognitive Science
UC San Diego, La Jolla, CA 92093, USA

Abstract

Previous work has found competing evidence for how contextual diversity influences early word learning. In support of contextual diversity facilitating learning, the corpus-derived diversity metric from Hills, Maouene, Riordan, and Smith (2010) was found to correlate with earlier ages of acquisition in children. We extend this work to five languages, accounting for a nonlinear relationship between the raw contextual diversity metric and word log-frequencies, and we account for additional covariates such as word length, concreteness, and lexical class. In contrast with the original result, we find that contextually diverse words are acquired later by children across languages. Our findings support the hypothesis that contextual diversity introduces an excess of possible meanings for contextually diverse words, adding noise to the word learning process. This hindering effect overshadows any benefits of syntactic or semantic bootstrapping during early word learning, when children are still in the early stages of vocabulary and conceptual development.

Keywords: language acquisition; word learning; contextual diversity; lexical acquisition

Introduction

Learning early words requires children to map spoken word forms to meaning representations. There are a variety of mechanisms that may help them accomplish this task, ranging from statistical learning using word distributions (Mintz, 2006), to social cues such as eye gaze and pointing (Çetinçelik, Rowland, & Snijders, 2021). However, one linguistic feature that is believed to affect early word acquisition, but remains controversial, is contextual diversity. Previous work has argued that when words appear in diverse contexts, their meaning is easier to narrow down using semantic and syntactic bootstrapping (Hills et al., 2010). On the other hand, it has been suggested that contextual diversity might simply add noise to the word learning process (Roy, Frank, DeCamp, Miller, & Roy, 2015). In our work, we seek to distinguish between these competing hypotheses.

One way to examine how a given factor influences word learning is to consider when children acquire individual words. Specifically, Braginsky, Yurovsky, Marchman, and Frank (2016) defined a word's age of acquisition as the age at which half of children produce that word, as judged by caregivers, usually parents. Across seven languages, earlier ages of acquisition were associated with higher word frequencies, shorter words, and higher concreteness ratings. Earlier ages of acquisition were also associated with shorter mean lengths of utterances (MLU, the mean length of utterances containing a word), a coarse metric for the syntactic complexity of contexts in which a word appears (Roy et al., 2015). Although the relative importance of different factors varied across lexical classes and across languages, there were consistent directions of effects for each factor, suggesting that early word learning may involve similar mechanisms across languages.

However, using this technique and others has produced mixed results for contextual diversity. Hills et al. (2010) quantified a word's contextual diversity as the number of distinct words co-occurring with the target word in the CHILDES corpus of child-directed speech (MacWhinney, 2000). Using this metric, Hills et al. (2010) found that contextually diverse words generally had earlier ages of acquisition. They proposed that contextually diverse words are learned earlier because they are more connected to other words either in a child's internal semantic network or in the learning environment; viewed from the perspective of syntactic and semantic bootstrapping, children might be able to better determine the meaning of a new word if it often co-occurs with previously learned words or concepts. By integrating information about known words and the current environment, children can place constraints on the possible meanings for a novel word. Then, across multiple encounters, novel words that appear in conjunction with a more diverse set of contexts and known words would have greater constraints placed upon their possible meanings, leading to faster acquisition.

In further support of contextual diversity facilitating word learning, diverse contexts have consistent beneficial effects on word learning in adults and school-age children. Adults are faster and more accurate at recognizing novel words in stories if they appear in more diverse semantic contexts (Johns, Dye, & Jones, 2016), and third grade children learn words more effectively when they appear in more diverse texts in classroom settings (Rosa, Tapia, & Perea, 2017). Because these studies involve more mature speakers, the theoretical focus is less on the initial acquisition of a word's meaning, and more on the retrieval of the word's meaning after initial exposure. It is argued that more contextually diverse words are retrieved in a wider variety of contexts, facilitating more efficient retrieval for future exposures to the word in novel contexts; retrieval of contextually specific words would be reliant on the specific context in which they were learned (Johns et al., 2016).

However, these retrieval arguments are less applicable to

early word acquisition, when one of the primary challenges is to initially map a word to its meaning, without having the word's meaning stored previously in semantic memory. Thus, the arguments in favor of contextual diversity in word learning propose that contextual diversity facilitates two complementary mechanisms: semantic and syntactic bootstrapping to initially identify a word's meaning (more relevant for early word acquisition; Hills et al., 2010), and context-general retrieval to efficiently retrieve meanings during subsequent exposures (more relevant for word acquisition in mature speakers; Johns et al., 2016).

At the same time, there is contradictory evidence that contextual diversity may in fact impede word learning in children. Roy et al. (2015) defined a word's contextual "distinctiveness" using probability distributions over a word's occurrences in time, space, and semantic topics. On average, more distinctive words were learned earlier, suggesting that exposure to a word in highly specific contexts is beneficial to word learning. Of course, the contextual diversity metric in Roy et al. (2015) is clearly distinct from measures of purely linguistic contextual diversity in previous studies, but one would expect the different measures to correlate, resulting in similar effects on early word acquisition. Instead, Roy et al. (2015)'s results suggest that contextual diversity may impede early word learning.

Under this account, contextual diversity might add unnecessary noise to the word learning process for each word. When a word appears in many different contexts, it co-occurs with a wider variety of linguistic and non-linguistic stimuli. This variety of stimuli may make it more difficult for children to associate the word with a specific meaning, particularly if their existing vocabulary is insufficient to narrow down the field of potential meanings. Thus, we are left with two competing accounts of contextual diversity in early word learning: (1) contextual diversity facilitates word learning through syntactic and semantic bootstrapping, or (2) contextual diversity hinders word learning by introducing unnecessary noise and potential word meanings.

In this work, we seek to further clarify whether contextual diversity facilitates or hinders early word acquisition. We adopt the general approach of Hills et al. (2010), predicting words' ages of acquisition from a corpus-derived metric for contextual diversity. We extend this approach to five different languages, and we account for a nonlinear effect of word frequency on the raw contextual diversity metric. In direct contrast with the original result (Hills et al., 2010), we find that contextually diverse words consistently have *higher* ages of acquisition across languages, consistent with the claim that contextual diversity hinders early word learning. These results are consistent with an interpretation that even if increased contextual diversity confers benefits on word learning, through syntactic and semantic bootstrapping, or improved retrieval of contextually diverse words, these benefits do not overcome the noise-introducing effects of contextual diversity in early word acquisition.

Method

We assessed whether contextual diversity was predictive of words' ages of acquisition in multiple languages after accounting for covariates such as frequency and lexical class.[1] We selected the five languages with the largest corpora of transcribed child-directed speech in the CHILDES dataset (MacWhinney, 2000) that also had child age of acquisition data in the Wordbank database (Frank, Braginsky, Yurovsky, & Marchman, 2017): English, German, Mandarin, Spanish, and French. The size of the CHILDES corpus for each language is listed in Table 1.

Age of Acquisition

We calculated word acquisition curves using the Wordbank database (Frank et al., 2017). Wordbank contains compiled data from parents reporting when their child produced each word in the MacArthur-Bates Communicative Development Inventories (CDI; Fenson et al., 2007); we pulled data for all available words on the CDI Words and Sentences instrument. As in Braginsky et al. (2016), for each word, we fitted a logistic curve predicting the proportion of children producing that word at different ages. We defined a word's age of acquisition as the age at which half of children were predicted to produce the word. We excluded words with ages of acquisition below zero or above five years old. This excluded only one to four words per language, usually words such as "daddy," "mommy," and "ouch." Because the CDI Words and Sentences instrument is designed for children aged 18 to 30 months old, some words' ages of acquisition were outside the CDI administration age range, extrapolated based on the Wordbank data. We obtained ages of acquisition for words in English (649 words), German (417 words), Mandarin (419 words), Spanish (371 words), and French (468 words).

Contextual Diversity Metric

We quantified contextual diversity using the metric from Hills et al. (2010). For each word w in the CDI, we counted the number of distinct word lemmas that appeared within a window of size k around w in the CHILDES dataset for the corresponding language. Windows were inclusive of the target word; for instance, a window of size two considered only adjacent words. To best match the setup from Hills et al. (2010), we initially used a window size of 5. Results for other window sizes are discussed in later sections.

When counting co-occurring lemmas, raw tokens in the CHILDES corpus were lemmatized using the spaCy toolkit for English, Spanish, French, and German (Honnibal, Montani, Van Landeghem, & Boyd, n.d.).[2] Because spaCy does not support lemmatization in Mandarin, and because Mandarin exhibits relatively little morphological inflection, we

[1] Code is available at `https://github.com/tylerachang/contextual-diversity`.

[2] We obtained similar results when computing contextual diversities using raw tokens instead of lemmas, but we used lemmas to best match Hills et al. (2010).

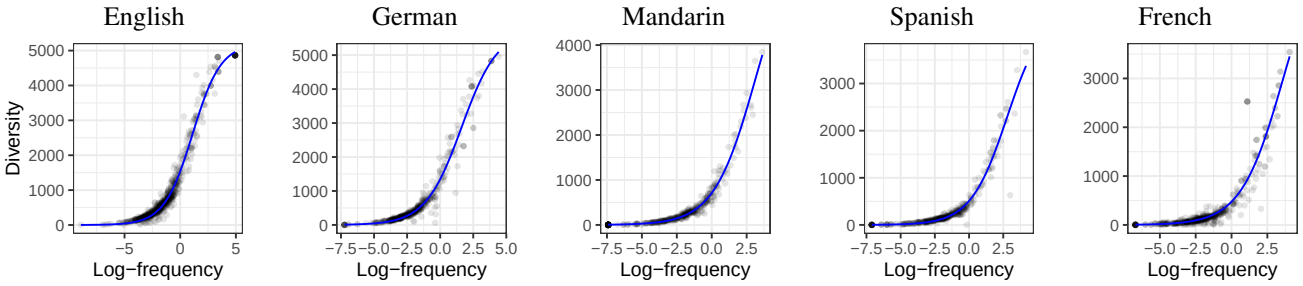

Figure 1: The raw contextual diversity metric was highly correlated with word log-frequency. Blue curves represent fitted sigmoid functions. Residuals of the sigmoid regression were used as our frequency-adjusted contextual diversity metric. The high diversity outlier point in French was the verb "avoir" ("to have").

used un-lemmatized word forms for Mandarin. In each language, we considered all words in the CDI along with the 5,000 most frequent lemmas in the CHILDES corpus for that language; the remaining rare lemmas were excluded from our contextual diversity counts. As in Hills et al. (2010), our raw contextual diversity metric was simply the number of distinct lemmas that occurred within window size k of the target word in the CHILDES corpus.

However, we found that the raw contextual diversity metric was highly correlated with a word's log-frequency in the CHILDES dataset (Pearson's $r = 0.72$ to 0.92 for the five languages; $p < 0.001$ for all languages). Unsurprisingly, more frequent words co-occurred with a greater number of distinct lemmas. Notably, when predicting words' ages of acquisition directly from raw contextual diversities and log-frequencies, the variance inflation factors (VIFs) for contextual diversity ranged from 2.0 to 6.7, exceeding the common cutoff of 5.0 for one language (English). To account for these colinearities, we sought to adjust the raw contextual diversity metric to account for any effects of frequency.[3] Because the relationship between raw contextual diversity and log-frequency was roughly sigmoidal or exponential in all languages (see Figure 1), we fitted a sigmoid curve for each language, predicting raw contextual diversities from words' log-frequencies in the CHILDES corpus.[4] We used the residuals after the sigmoid regressions to quantify words' contextual diversities after accounting for expected co-occurrence counts based on word frequencies. These residuals served as our final metric for words' contextual diversities in each language, removing outliers further than three standard deviations from the mean contextual diversity. In future sections, the contextual diversity metric is frequency-adjusted unless stated otherwise.

[3]Hills et al. (2010) accounted for log-frequency by including both contextual diversity and log-frequency in a linear regression predicting words' ages of acquisition in English. However, the signs of coefficients for highly correlated variables in linear regressions can be unreliable. Furthermore, as seen in Figure 1, the relationship between contextual diversity and log-frequency was nonlinear.

[4]As can be seen in Figure 1, the sigmoid curves were able to fit both sigmoidal and exponential relationships between raw contextual diversity and log-frequency.

		Predictors	
Language	CHILDES corpus size (tokens)	Diversity alone	+log-frequency +n-chars +concreteness +lexical class +MLU
English	7.9M	0.17^{***} (+)	0.37^{***} (+)
German	4.2M	0.09^{***} (+)	0.27^{**} (+)
Mandarin	1.7M	0.04^{***} (+)	0.27^{***} (+)
Spanish	1.2M	0.02^{**} (+)	0.13^{*} (+)
French	0.9M	0.01^{**} (+)	0.13

Table 1: Linear regression results predicting words' ages of acquisition in children. Numbers indicate adjusted R^2 values, and asterisks indicate significant effects of contextual diversity. Significance levels were determined by the coefficient confidence intervals for column three, and by likelihood ratio tests (LRTs) in column four. When significant, signs indicate the sign of the coefficient for contextual diversity.

Results

We first ran a linear regression for each language, predicting words' ages of acquisition directly from their contextual diversities (after frequency adjustment). Results are shown in the third column of Table 1. In all five languages, there was a significant positive coefficient for contextual diversity, suggesting that children learn more contextually diverse words later. This contrasts with the findings of Hills et al. (2010), who found a negative coefficient for contextual diversity in English before accounting for the nonlinear relationship between word frequency and raw contextual diversity. Not only did our frequency-adjusted contextual diversity metric account for more variance in words' ages of acquisition in English than the raw contextual diversity metric (adjusted $R^2 = 0.17$ compared to 0.02), it accounted for more variance in English than any of the covariates described in the next section. This result suggests that contextual diversity may be a significant driving factor in early word learning, such that contextually diverse words are learned later.

Covariates

To test the robustness of the effects of contextual diversity and to account for possible colinearities, we ran regressions

accounting for other variables that have been shown to influence words' ages of acquisition in children across languages. We considered all five significant predictors of age of acquisition from Braginsky et al. (2016):

- Log-frequency: the natural log of the word's per-1000 word frequency in the CHILDES corpus.
- N-chars: the number of characters in the target word, serving as a coarse proxy for word length. Longer words are likely to be harder for children to parse and produce.
- Concreteness: human-generated concreteness norms from Brysbaert, Warriner, and Kuperman (2014), rated on a five-point scale.
- Lexical class: lexical classes as annotated in Wordbank. The distribution over lexical classes in each language was similar to Hills et al. (2010), except we retained words in the Other class. Our results were similar when removing those words. Across languages, the mean proportion of words in each class were: Noun (50%), Verb (16%), Adjective (9%), Function Word (11%), and Other (13%).
- Mean length of utterance (MLU): the mean length of sentences containing the target word in the CHILDES corpus. MLU has been used as a metric for the complexity of syntactic contexts in which a word appears (Roy et al., 2015).

We ran a linear regression for each language, predicting words' ages of acquisition using linear terms for each predictor. The VIF for contextual diversity was less than 1.5 in every language, suggesting minimal effects of multicolinearity and allowing us to safely interpret our regression results for contextual diversity. For each language, we determined the significance of contextual diversity using a likelihood ratio test (LRT), comparing a model using the five predictors plus contextual diversity with a model using only the five predictors. Results are shown in the last column of Table 1.

In four out of the five languages, the LRT identified a significant effect of contextual diversity even after all five covariates were accounted for, indicating that contextual diversity accounts for a significant amount of variance beyond log-frequency, n-chars, concreteness, lexical class, and MLU in those languages. Importantly, the coefficient for contextual diversity was positive in each language where it remained statistically significant; contextually diverse words were generally acquired later. Even the non-significant coefficient for contextual diversity in French was positive ($p = 0.17$). These results conflict with previous studies suggesting that contextual diversity facilitates early word learning.

Effects of Window Size and Language

We repeated our analyses with windows of size two, three, five, ten, fifteen, and thirty when computing contextual diversities.[5] The varying window sizes can be interpreted as a gradient between syntactic and semantic diversities; Chang and Deák (2020) considered adjacent tokens as syntactic contexts and larger windows as semantic contexts. For instance, a diverse set of adjacent tokens might indicate a relatively flexible part of speech (high syntactic contextual diversity), while a diverse set of nearby but non-adjacent tokens might indicate a wide variety of semantically related topics.

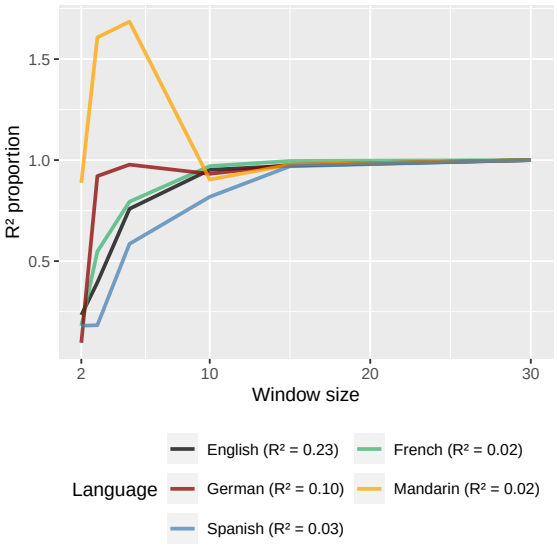

Figure 2: Adjusted R^2 values when predicting words' ages of acquisition from contextual diversity computed with different window sizes in each language. To visualize comparisons across languages, adjusted R^2 values are expressed as proportions of the R^2 value at the largest window size of 30. The raw adjusted R^2 values for window size 30 are indicated next to each language. In all languages except Mandarin, window size 30 exhibited the largest R^2 value.

When predicting words' ages of acquisition from contextual diversity alone, there was a significant positive coefficient for contextual diversity in nearly all languages and window sizes (i.e. more contextually diverse words were learned later). The only non-significant coefficients were in Spanish and French (the two languages with the smallest CHILDES corpora) with window sizes 2-3 and 2 respectively. Even after accounting for covariates as in the previous section, the only negative coefficient for contextual diversity was a non-significant coefficient ($p = 0.82$) in French with window size two after accounting for all five covariates. These results are in direct opposition to previous findings suggesting that contextual diversity facilitates early word learning. Instead, our results are compatible with the hypothesis that contextual diversity, whether it be syntactic or semantic, hinders early word learning.

In the majority of languages, we also found that contextual diversity had larger effects on words' ages of acquisition as

[5]The sigmoid regression predicting raw contextual diversities from log-frequencies did not converge for French with window size two. In this case, we regressed raw contextual diversity quadratically over log-frequency. As before, we used the residuals as our frequency-adjusted contextual diversity metric.

window size increased (i.e. semantic contextual diversity). As shown in Figure 2, in all languages except Mandarin, contextual diversity had the largest effect on age of acquisition when computing diversities with the largest window size of 30. This suggests that in these languages, semantic contextual diversity may have a larger hindering effect on word learning than syntactic contextual diversity. However, in Mandarin, there was a spike in the effect of contextual diversity at window size five. In Mandarin, contextual diversities with smaller window sizes had relatively large effects on words' ages of acquisition, suggesting more substantial effects of syntactic contextual diversity than in the other four languages. Thus, we hypothesize that the spike in Mandarin at window size five may be due to compounded effects of syntactic and semantic diversity captured by our contextual diversity metric for intermediate window sizes in Mandarin.

In summary, it appeared that all five languages exhibited substantial (and statistically significant) hindering effects of semantic contextual diversity on words' ages of acquisition; effects of syntactic contextual diversity were less pronounced except in Mandarin. The more pronounced effects of syntactic contextual diversity in Mandarin could be due to morphological differences between Mandarin and the other languages. For example, the lack of morphological inflection in Mandarin might make it harder for children to learn words that appear in diverse syntactic contexts. Of course, because our contextual diversity metric is derived from text corpora, differing effects of syntactic contextual diversity could also be due to orthographic differences between Mandarin and the other languages; we further note that the remaining four languages were all European languages, and the overall variance explained by contextual diversity differed significantly across languages. Further study is required to assess differences between languages and cultures in the effects of syntactic vs. semantic contextual diversity on early word learning, but our results indicate that contextual diversity in either form consistently correlates with later acquisition of words in children.

Effects of Corpus Size

As shown in Table 1, languages with larger CHILDES corpora tended to exhibit larger effects of contextual diversity on words' ages of acquisition. Intuitively, larger corpora might produce cleaner metrics for words' contextual diversities. To test whether the larger effects of contextual diversity were in part driven by corpus size, we ran our analyses using random subsets of the CHILDES corpus for each language. We computed contextual diversities for subsets of 50K, 100K, 250K, 500K, 1M, 2M, 4M, and 6M tokens, limited by the full corpus size for each language. As in Hills et al. (2010), we used a window size of 5 for all languages.

The adjusted R^2 values when predicting words' ages of acquisition from contextual diversities computed from different corpus sizes in each language are shown in Figure 3. In all languages, contextual diversity accounted for more variance in words' ages of acquisition as corpus size increased. Again, all significant coefficients for contextual diversity were positive, indicating that contextually diverse words were acquired later. These results support the hypothesis that larger corpus sizes lead to cleaner contextual diversity metrics, and that cleaner diversity metrics better capture the hindering effects of contextual diversity on early word learning.

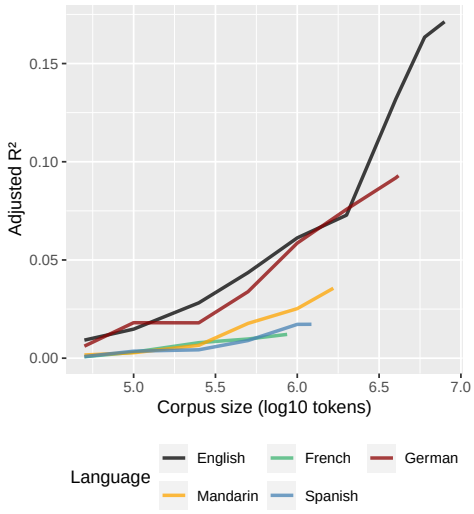

Figure 3: Adjusted R^2 values when predicting words' ages of acquisition from contextual diversity computed with different corpus sizes in each language. Across languages, contextual diversity computed over larger corpora accounted for more variance in words' ages of acquisition.

That said, corpus size did not account entirely for differences between languages in the effects of contextual diversity on words' ages of acquisition. For example, English and German exhibited larger effects of contextual diversity than the other languages even when subsampled to the same corpus sizes. This may be due to the quality or uniformity of the data (e.g. the subsamples were still sampled from the larger original corpora for those languages), but it could also reflect some veridical difference in how contextual diversity affects word acquisition in those languages.

Discussion

We found that more contextually diverse words had significantly later ages of acquisition in children across languages. Our results were consistent across window sizes (syntactic vs. semantic contextual diversities) and corpus sizes when computing contextual diversities. These results support the theory that diverse contexts add noise to the process of mapping novel words to meanings, introducing a wider variety of possible meanings for contextually diverse words. Notably, our results contradict previous findings suggesting that contextual diversity facilitates early word learning (Hills et al., 2010). Previous studies have argued that greater contextual diversity allows children to better infer word meanings based on known words in the surrounding context. However, this bootstrapping process relies on a substantial set of existing

known words. In early childhood, children may know too few words to exhibit any substantial effects of bootstrapping. Indeed, this theory would explain why contextual diversity might hinder word learning in young children but facilitate word learning in older children and adults (see Introduction; Johns et al., 2016; Rosa, Salom, & Perea, 2022). In early childhood, the noise introduced by contextual diversity overshadows any benefits of syntactic or semantic bootstrapping; however, as a child's vocabulary grows, the benefits of bootstrapping become more prominent. Furthermore, as children progress in their conceptual development, knowledge of abstract categories such as nouns (or objects) and verbs (or actions) may compound with the effects of larger vocabularies to further improve children's ability to leverage diverse contexts to infer word meanings.

Implications for Related Work

Our results are in direct contrast with previous corpus-based studies of contextual diversity in early word learning (Hills et al., 2010). Methodologically, our results demonstrate the importance of accounting for covariates in correlational language acquisition research. Previously-identified effects of contextual diversity appear to have been driven primarily by a nonlinear correlation between word frequency and the raw contextual diversity metric. To corroborate our results, future work might use our frequency-adjusted contextual diversity metric to estimate raw production data using mixed-effects logistic regression models as in Braginsky, Yurovsky, Marchman, and Frank (2019), rather than estimating single age of acquisition values. We note that the results from Braginsky et al. (2016) did not change substantially when using these mixed-effects models (Braginsky et al., 2019).

Conceptually, our results emphasize the differences between word learning in young children (e.g. less than three years old) and in more mature speakers. Previous studies have established the benefits of contextual diversity for word learning in school age children and adults (Johns et al., 2016; Rosa et al., 2017, 2022). Even among known words, contextually diverse words are faster to access later in life (Adelman, Brown, & Quesada, 2006), despite findings that earlier-acquired words (which according to our results, may be less contextually diverse) tend to be accessed faster as well (Ellis & Morrison, 1998; Stewart & Ellis, 2008). To reconcile these seemingly contradictory findings, we hypothesize that the advantages for contextually diverse words in adult lexical access overshadow any disadvantages of later acquisition. In the years after childhood, more contextually diverse words are retrieved and consolidated in memory in a wider variety of contexts than other words, facilitating improved retrieval despite potentially later initial acquisition.

Of course, the contextual diversity metric used in this study specifically measured lexical diversity, counting word co-occurrences in a corpus. Intuitively, this metric is likely to correlate with non-linguistic measures of words' contextual diversities. Previous work has sought to quantify contextual diversity across temporal, semantic, and spatial dimensions, finding that contextually "distinctive" words tend to be learned earlier by children under the age of three (Roy et al., 2015). Our results lend further support for this finding.

Previous work has also considered specific types of contexts in which words appear, rather than quantifying overall contextual diversity. For instance, Chang and Deák (2020) found that specific syntactic contexts (e.g. frames for nouns and verbs) and thematic contexts (e.g. food and face-to-face play) were predictive of words' ages of acquisition. In a similar line of work, Mintz (2003) considered contexts as frequent frames, frames consisting of adjacent words around a target word. They found that children categorized words together if they appeared in similar frequent frames. These results suggest that the context in which a word appears has a significant effect on early word learning; our results further suggest that a greater diversity of such contexts may hinder early word learning.

Finally, contextual diversity is closely related to measures of words' lexical ambiguity and polysemy. Words with a wider variety of meanings are more likely to appear in more diverse contexts. Indeed, the degree to which a word form appears in polysemous meanings can be estimated from linguistic usage (Garí Soler & Apidianaki, 2021), and more polysemous words have been found to be acquired later by children (Meylan, Mankewitz, Floyd, Rabagliati, & Srinivasan, 2021). Differentiating effects of contextual diversity and polysemy will depend on how word "meanings" are defined; under a continuous notion of word meaning based on the contexts in which a word appears (Nair, Srinivasan, & Meylan, 2020), operationalizations of contextual diversity and polysemy might be indistinguishable. For example, corpus-derived measures of contextual diversity and polysemy might capture the same co-occurrence statistics in language. Still, from a theoretical perspective, an ideal measure of polysemy would capture more features of people's *internal* word meaning representations than contextual diversity, which is primarily a measure of *external* contexts.

Conclusion

In this work, we extended Hills et al. (2010) to study the effects of contextual diversity on words' ages of acquisition in five languages, after accounting for a nonlinear effect of word frequency on the raw contextual diversity metric. In contrast with the original finding, we found that contextually diverse words were consistently learned later by children across languages. Our work reconciles the findings of Hills et al. (2010) with previous results finding hindering effects of contextual diversity on early word learning (Roy et al., 2015). We hypothesize that the hindering effects of contextual diversity are driven by noise, where diverse contexts introduce an excess of possible word meanings, overshadowing any benefits of syntactic and semantic bootstrapping in diverse contexts. Our work suggests that the consistency of contextual cues is of central importance to early word learning, when children have little existing vocabulary to rely on.

Acknowledgements

We would like to thank the anonymous reviewers for their insightful comments, Thomas Hills for helpful suggestions and advice, and the UCSD Language and Cognition Lab (Sean Trott, James Michaelov, and Cameron Jones) for valuable discussion. Tyler Chang is partially supported by the UCSD HDSI graduate fellowship.

References

Adelman, J. S., Brown, G. D. A., & Quesada, J. F. (2006). Contextual diversity, not word frequency, determines word-naming and lexical decision times. *Psychological Science*, *17*, 814–823.

Braginsky, M., Yurovsky, D., Marchman, V., & Frank, M. (2016). From uh-oh to tomorrow: Predicting age of acquisition for early words across languages. In *Proceedings of the annual meeting of the cognitive science society.*

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

Brysbaert, M., Warriner, A., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*, *46*.

Chang, L., & Deák, G. (2020). Adjacent and non-adjacent word contexts both predict age of acquisition of English words: A distributional corpus analysis of child-directed speech. *Cognitive Science*, *44*(11).

Ellis, A., & Morrison, C. (1998). Real age-of-acquisition effects in lexical retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(2), 515–523.

Fenson, L., Marchman, V., Thal, D., Dale, P., Reznick, S., & Bates, E. (2007). *Macarthur-bates communicative development inventories*. Baltimore, MD: Paul H. Brookes Publishing Company. Retrieved from `https://mb-cdi.stanford.edu`

Frank, M., Braginsky, M., Yurovsky, D., & Marchman, V. (2017). Wordbank: An open repository for developmental vocabulary data. *Journal of Child Language*, *44*(3), 677–694.

Garí Soler, A., & Apidianaki, M. (2021). Let's play monopoly: BERT can reveal words' polysemy level and partitionability into senses. *Transactions of the Association for Computational Linguistics*, *9*, 825–844.

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

Honnibal, M., Montani, I., Van Landeghem, S., & Boyd, A. (n.d.). *spaCy: Industrial-strength natural language processing in python.* Retrieved from `https://spacy.io/`

Johns, B., Dye, M., & Jones, M. (2016). The influence of contextual diversity on word learning. *Psychonomic Bulletin and Review*, *23*, 1214–1220.

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk*. Mahwah, NJ: Lawrence Erlbaum Associates. Retrieved from `https://childes.talkbank.org`

Meylan, S., Mankewitz, J., Floyd, S., Rabagliati, H., & Srinivasan, M. (2021). Quantifying lexical ambiguity in speech to and from English-learning children. *PsyArXiv*.

Mintz, T. (2003). Frequent frames as a cue for grammatical categories in child directed speech. *Cognition*, *90*(1), 91–117.

Mintz, T. (2006). Finding the verbs: Distributional cues to categories available to young learners. In *Action meets word: How children learn verbs* (pp. 31–63). Oxford University Press.

Nair, S., Srinivasan, M., & Meylan, S. (2020). Contextualized word embeddings encode aspects of human-like word sense knowledge. In *Proceedings of the workshop on the cognitive aspects of the lexicon* (pp. 129–141).

Rosa, E., Salom, R., & Perea, M. (2022). Contextual diversity favors the learning of new words in children regardless of their comprehension skills. *Journal of Experimental Child Psychology*, *214*.

Rosa, E., Tapia, J. L., & Perea, M. (2017). Contextual diversity facilitates learning new words in the classroom. *PLoS ONE*, *12*(6).

Roy, B., Frank, M., 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.

Stewart, N., & Ellis, A. (2008). Order of acquisition in learning perceptual categories: A laboratory analogue of the age-of-acquisition effect? *Psychonomic Bulletin and Review*, *15*, 70–74.

Çetinçelik, M., Rowland, C., & Snijders, T. (2021). Do the eyes have it? a systematic review on the role of eye gaze in infant language development. *Frontiers in Psychology*, *11*.