Developmental Science

WILEY

Developmental Science

RESEARCH ARTICLE

The Words Children Hear and See: Lexical Diversity Across-Modalities and Its Impact on Lexical Development

Luan Li[1,2,3] | Ming Song[2,4,5] | Qing Cai[2,4,5]

[1]Speech-Language-Hearing Center, School of Foreign Languages, Shanghai Jiao Tong University, Shanghai, China | [2]Shanghai Key Laboratory of Brain Functional Genomics (Ministry of Education), Affiliated Mental Health Center (ECNU), Institute of Brain and Education Innovation, School of Psychology and Cognitive Science, East China Normal University, Shanghai, China | [3]National Research Centre for Language and Well-Being, Shanghai, China | [4]Shanghai Changning Mental Health Center, Shanghai, China | [5]Shanghai Center for Brain Science and Brain-Inspired Technology, Shanghai, China

Correspondence: Qing Cai (qcai@psy.ecnu.edu.cn)

Received: 8 May 2024 | **Revised:** 8 November 2024 | **Accepted:** 2 December 2024

Funding: This work was funded by the National Science and Technology Innovation 2030 (STI 2030—Major Projects 2021ZD0200500), the National Natural Science Foundation of China (31970987 to Qing Cai), and the National Social Science Fund of China (23CYY040 to Luan Li).

Keywords: lexical development | lexical diversity | middle childhood | multimodal language environments

ABSTRACT

Early vocabulary development benefits from diverse lexical exposures within children's language environment. However, the influence of lexical diversity on children as they enter middle childhood and are exposed to multimodal language inputs remains unclear. This study evaluates global and local aspects of lexical diversity in three 1.6-million-word child-directed corpora, representing average Chinese children's speech, print, and media language environments. Additionally, pseudo-multimodal samples were compiled from the three corpora to compare with the unimodal environments on lexical diversity. We then investigated the associations between lexical diversity and the acquisition of 361 words spanning early-to-middle childhood. The findings show that print and pseudo-multimodal language provided the most diverse lexical environments, whereas speech exhibited the least diversity. However, speech diversity most strongly predicted lexical development, particularly before the onset of middle childhood. Exploratory analysis revealed that lexical diversity of other modalities emerged as stronger predictors thereafter. Early lexical development was best predicted by words' variations in connectivity with other words within an immediate context, whereas in middle childhood, variations in words' occurrences in larger context windows became the primary predictor, implicating children's growing ability to attend to linguistic contexts of increasing sizes. Importantly, higher diversity was consistently associated with earlier word acquisition across measures and developmental phases. These findings underscore the critical role of varied lexical experiences in children's language development.

1 | Introduction

Child vocabulary development is strongly associated with the linguistic input they receive. Particularly, qualitative variations of lexical items in the speech input play a pivotal role in shaping early vocabulary acquisition (Huttenlocher et al. 2010; Montag, Jones, and Smith 2018; Rowe 2012; Rowe and Weisleder 2020; Weizman and Snow 2001). However, as children progress from early to middle childhood (3–9 years of age), they encounter increasingly complex language environments characterized by

Luan Li and Ming Song contributed equally to this work.

Summary

- In early-to-middle childhood, print and multimodal language environments offer the highest lexical diversity, followed by media language input.
- Greater local lexical diversity, as well as the diversity of input modality, is positively associated with lexical development.
- Speech diversity most strongly predicted early lexical development, with its influence declining with the onset of middle childhood.
- The study extends research on lexical diversity and its influence on Chinese, contributing to knowledge in non-Indo-European language development.

multimodal inputs, including daily speech, screen media, and printed words. The specific lexical experiences provided by these input modalities and their impact on lexical development are not fully understood. Further, the extent to which the pattern of variability is present in children's language environment beyond the English-speaking world remains largely unexplored. To address these gaps, we investigate lexical diversity—a hallmark of input quality—in children's speech, media, and print language environments and its influence on lexical development. Our analysis extends beyond the English-speaking context to Chinese, a non-Indo-European language with the largest number of developing language learners globally.

1.1 | Language Input Across Modalities

Throughout development, input modality is an important source of variation in children's language environment. In infancy and toddlerhood, caregiver speech is significantly associated with the rate of children's vocabulary growth (Hart and Risley 1995). As children grow, exposure to language through TV and other screen media becomes substantial. By the age of 3, children can spend up to 3.6 h a day watching TV (Madigan et al. 2019). Further, even before formal education begins, many children engage with written texts through shared book reading, which constitutes another significant source of language input. Individual differences in the amount of print exposure are already present in preschoolers (e.g., Scheele, Leseman, and Mayo 2010), which was found to strongly influence later vocabulary learned through reading (Mol and Bus 2011).

Although all these input modalities are essentially "child-directed," they differ in their linguistic features. For example, compared with child-directed speech, picture books contain more complex sentence constructions (Cameron-Faulkner and Noble 2013; Hsiao, et al. 2023; Montag 2019), a greater variety of word types (Dawson et al. 2021; Montag, Jones, and Smith 2015) and longer and more morphologically complex words (Dawson et al. 2021). Given that print exposure typically starts much later than speech input, the complexity of book language aligns with the developmental trajectory wherein children progress from acquiring simple to complex language structures.

A pertinent question arises regarding the language experiences of children as they start navigating across modalities from early to middle childhood. Media language exposure during this period is substantial, and its impact on child development has received considerable research attention in recent years. The findings, however, seem to suggest a complex relationship: while excessive screen use is often negatively associated with language development (Pagani, Fitzpatrick, and Barnett 2013), it could be beneficial when coupled with interactive content and reasonable usage duration (Jing et al. 2023; Madigan et al. 2019). This implicates that not all media interactions have "equal" influences on children's language development, leaving open questions about whether it is the language per se or the accompanying activity that impacts language learning from screen media exposure. Theoretically, language in children's animated cartoons and movies, while scripted, is conversational, offering a complex language environment akin to books while being as interactive as speech, which might confer benefits for language and cognitive development (Mar, Tackett, and Moore 2010). Despite considerable screen media exposure in children and concerns about its impact on language development, little is known about the language children encounter when using screen media, and comparisons between media language with speech or print language environments have been scant.

Therefore, the primary aim of this study was to compare language across children's speech, print, and media inputs. It is important to note that the actual language environments of most children are essentially multimodal. They all are exposed to speech, print, and media language to varying degrees, and their daily activities are often accompanied by a blend of language input modalities (e.g., watching TV or reading picture books while hearing caregivers talking to them). By analyzing the language modalities as separate input sources, we aimed to query the lexical experience children might have when exposed to each modality of language input but not children's actual language activity.

We focus specifically on a key aspect of input quality: lexical diversity. The term "lexical diversity" here refers to a range of measures of the degree of lexical variation in language input, as seen in prior research. Current word learning theories and models posit that forming connections between words and their contextual surroundings is crucial for lexical development, as the goal of word learning is to establish a densely connected network of lexicalized concepts to support efficient retrieval and use (Hills et al. 2009; Unger and Fisher 2021). Consequently, measures of word diversity in relation to other words and contexts have been widely studied in the context of language development in infancy and reading development in school-age children. Differences in lexical diversity within individuals' language environments or in word learning episodes are strong predictors of vocabulary development or word learning (e.g., Hsiao and Nation 2018; Huttenlocher et al. 2010; Rowe 2012). Yet this crucial aspect of lexical quality has been little examined in the input of children moving from early to middle childhood. We aim to close this gap by investigating a range of lexical diversity measures, which have been separately probed in prior work, within a single study. In the following sections, we first review different measures of lexical diversity and then research on their impact on lexical development.

1.2 | Global and Local Lexical Diversity

To assess the degree of diversity of the lexical experiences provided by a language environment, previous research has utilized *type-token ratio*—the unique number of word types given a certain amount of word token input. This metric is positively linked to the language abilities of individual infants and toddlers (Hoff 2003; Hoff and Naigles 2002; Huttenlocher et al. 2010; Weizman and Snow 2001) and strongly predicts vocabulary growth (Rowe 2012). Computational simulations have also demonstrated that while the quantity of word token input is important for vocabulary learning early in the process, the number of unique word type inputs quickly surpassed it in importance as learning proceeds (Jones and Rowland 2017), highlighting its crucial role, particularly beyond the very early stages of lexical development. Importantly, research has consistently shown that children's picture books contain more unique word types, that is having a higher type-token ratio, than child-directed speech (Dawson et al. 2021; Montag, Jones, and Smith 2015), suggesting that it is a valid measure of the *global lexical diversity* of a language input environment.

Words also vary in their *local lexical diversity*. One intuitive aspect of a word's diversity in relation to its surroundings is its *connectivity*[1] with other words in the local lexical environment, which can be calculated as the probability of co-occurrences with the number of unique word types within a short contextual window. Connectivity is arguably highly related to lexical development, as it fosters the acquisition of relational knowledge and contributes to the formation of a structured lexical organization (Jiménez and Hills 2022a; Mintz 2003; Unger and Fisher 2021). It also significantly predicts words' AoA in infancy (Hills 2013; Hills et al. 2010) and could be used to distinguish between typically developing toddlers and late talkers (Jiménez and Hills 2022a, 2022b).

Additionally, a word's varied use and meaning can be induced from a larger context that incorporates topicalized semantic content, indexed by *contextual diversity (CD)*[2] or *semantic diversity (SD)* measures. The measures differ in that CD is simply the number of unique documents in which a word occurs (Adelman, Brown, and Quesada 2006), while SD also considers the latent semantic similarities between the documents (Hoffman, Lambon Ralph, and Rogers 2013; Hsiao and Nation 2018). Children's experiences with words involve rich contextualized information, leading to variations in words' lexical quality in the mental lexicon—an idea shared by the *Lexical Legacy Hypothesis* (Nation 2017) and the *Semantic Distinctiveness Model* (Johns and Jones 2022; Jones, Johns, and Recchia 2012). They posit that a word's lexical strength is updated each time it appears in a new context. The updated strength is determined by the dissimilarity between the current context and the previous contexts. Consequently, words experienced in a wider variety of meaningful contexts should be more readily acquired and more robustly represented, as they are more likely to be accessed in the future. Therefore, the contextualized details of a word's occurrences should be considered in evaluating its lexical diversity. Although CD and SD metrics have been shown to impact written word acquisition in older children (e.g., Hsiao and Nation 2018), their variation across modalities and effects on lexical development around middle childhood remain unknown.

1.3 | Impact of Lexical Diversity on Word Learning

Whether a repetitive or diverse lexical surrounding is more conducive to word learning has always been of strong interest in language development research. Previous research has consistently shown that global lexical diversity has a significant facilitative effect: a higher type-token ratio in the input is associated with better vocabulary development in individual children (e.g., Rowe 2012). In contrast, the impact of local lexical diversity remains a subject of debate. Although research suggests that higher local connectivity facilitates word learning (Alhama, Rowland, and Kidd 2023; Hills et al. 2010; Jiménez and Hills 2022a, 2022b), there is also evidence that words with more diverse contexts or higher connectivity (Roy et al. 2015; Unger et al. 2024) are acquired later. In a recent study, Unger et al. (2024) proposed two potential causes for this discrepancy: the use of different measures of diversity and the inability to control for input frequency. To address these issues, the researchers computed two diversity measures in toddlers' speech environments—connectivity and SD—and used the residuals of diversity against frequency to disentangle the frequency effect. They found that although high connectivity is related to earlier AoA, the effect reversed after statistically controlling for frequency. Meanwhile, high SD was consistently associated with later acquisition.

Another possible explanation for these discrepant findings is that a diverse environment might hinder word learning in the early stage and benefit it later on. Previous research has primarily focused on word learning in infancy and toddlerhood (before 2.5 years old), a period when the primary goal is to establish mappings between linguistic labels and their referents. Yet beyond early childhood, the aim of lexical development gradually shifts to building a rich, interconnected lexical network that enables efficient lexical retrieval. In a recent review, Raviv, Lupyan, and Green (2022) proposed that high diversity likely poses challenges in the early stages of learning but eventually helps learners maintain a broad "hypothesis" about the usage of newly learned items or skills, thereby accommodating future encounters. Empirical support for this argument comes from a word learning experiment showing that adults learn to read rare words better in less diverse contexts, but when initial familiarization opportunities were provided, word learning benefits from more diverse contexts (Mak, Hsiao, and Nation 2021). From a developmental perspective, a facilitative role of diversity has been consistently found in older children learning to read visual words (Hsiao and Nation 2018; Hsiao et al. 2020), who have advanced well beyond the initial phases of language learning. If the learning phase indeed underlies the effect of diversity on word learning, it is possible that the impact is related to the increasing granularity of children's sensitivity to linguistic surroundings as their attention and memory span grows. Thus, diversity within larger contextual windows, such as CD and SD, may exert a stronger or more significant influence on lexical development in later stages. This possibility, which has not been previously explored due to a focus on infancy, underscores a gap in understanding the timing and nature of the diversity effect.

1.4 | The Present Study

The present study has two aims: (1) to evaluate cross-modality variations in global and local lexical diversity in children's language environment within a non-Indo-European context, and (2) to explore the influence of lexical diversity across modalities on lexical development from early to middle childhood. We leverage two recently released Chinese children's language environment corpora (Li, Yang, et al. 2023; Li et al. 2023) as proxies for average print and screen media language inputs, complemented by the CHILDES corpus (MacWhinney 2000) representing the speech environment. Considering that most children are exposed to multimodal language environments, we additionally compile corpora drawing equally from each of the unimodal corpora to approximate a pseudo-multimodal environment on lexical properties, which was used to compare with the unimodal corpora. In addition to lexical diversity, we computed words' *input frequency*, which is one of the most established indices of children's experience with words and is associated with the difficulty of word learning in toddlerhood (Goodman, Dale, and Li 2008). To address our first aim, we conduct cross-modality comparisons on lexical diversity. For our second aim, we used the diversity metrics, as well as input frequency, of the four modalities to predict the AoA of words acquired between the ages of 2.4 and 11 in a series of regression analyses.

2 | Method

We begin by introducing the corpora representing children's language environments in speech, book, and media modalities. We also elaborate on how the pseudo-multimodal samples were compiled. Within each corpus, we first evaluate global lexical diversity through the lens of type-token ratio and then compute three local lexical diversity measures (connectivity, CD, and SD). Detailed information on their calculations is provided. Analysis of variance (ANOVA) and subsequent Mann-Whitney U tests were conducted for the measures to address variations in lexical diversity among input modalities.

Next, we describe the dataset used to evaluate the impacts of diversity on lexical development. To gauge word learning, we used existing objective AoA data for concrete nouns. To explore the effects of lexical diversity across modalities on different development stages, the data were also partitioned into early-, mid-, and late-acquired words. Linear regression models were employed to examine the unique effects of the diversity metrics. Random forest regression models were also built to estimate the relative importance of the metrics simultaneously in one model. Random forests have been suggested to yield more accurate estimates of predictor importance than beta coefficients of multiple regressions and are robust in handling multicollinearity issues (Mizumoto 2023; Wurm and Fisicaro 2014). It is thus particularly suited for comparing the contributions of our diversity measures that are correlated with one another. We additionally built models, the results of which were broadly consistent and are reported in the Supporting Materials.

2.1 | Corpora Representing Language Environments

2.1.1 | Speech Environment

Child-directed speech was sourced from the Mandarin Chinese section of CHILDES (MacWhinney 2000) via the *childesr* library (Sanchez et al. 2019). The sample included transcripts of 1668 conversation sessions with 1385 typically developing children under 108 months ($M = 54.33$ months, SD = 18.14 months) from 17 databases. These data were collected during play sessions, home recordings, assessments, and shared book reading. Although some of the speech data were accompanied by book and/or media contexts, the speech was primarily conversational and not fully scripted (i.e., not directly reading book texts). For example, many of the shared reading sessions involved wordless picture books, designed to prompt caregiver-child spoken interactions. Therefore, we did not further exclude any databases. Utterances tagged with "Child" or "Target_Child" were excluded. Foreign language words, symbols, and pinyin transcriptions were removed. The final sample comprised 1,592,146 word tokens and 16,101 unique word types.

2.1.2 | Book Environment

To represent the average children's print language environment, we obtained 1497 picture books from the CCLOWW G2 subcorpus[3] (Li et al. 2023). To match word tokens with the other corpora, we additionally collected 1726 picture books using the children's book lists provided by a non-governmental organization iRead (http://www.iread.org.cn/project/shumu/shumu2022.html) as an orientation. The picture books covered a wide range and diversified child-friendly topics. We followed Li et al. (2023) for word segmentation and cleaning. The final sample contained 3223 picture books, 1,657,868 word tokens, and 42,877 unique word types.

2.1.3 | Media Environment

We utilized the CCLOOW corpus (Li et al. 2023), which includes transcripts of popular animated TV series and movies targeted at 3-to-9-year-old children in mainland China. We excluded 20 long episodes from the original 1,893,432 word-token corpus to match the size of the other corpora. The animated TV series and movies were all child-directed and were narratives. We also followed Li et al. (2023) for word segmentation and cleaning. The final sample contained 245 documents with 1,649,996 word tokens and 36,060 unique word types.

2.1.4 | Pseudo-Multimodal Environment

To create a pseudo-multimodal condition, we sampled documents equally from the three corpora (Figure 1). Given the different lengths of speech conversations, picture books, and animations, we first divided each corpus into short contexts of approximately 150 words. This produced 10,955 speech contexts (length $M = 145.34$, SD = 24.78), 11,667 book contexts (length $M = 142.10$, SD = 32.88), and 10,319 media contexts (length $M = 159.90$, SD = 15.45). These contexts also served as the basis for calculating connectivity, CD, and SD in each unimodal corpus, as described later. The 150-word context size was chosen to ensure reliable diversity measures, previously shown to explain unique variance in word naming efficiency in Chinese (Chang and Lee 2018). Each pseudo-multimodal sample included an equal third of speech, book, and media contexts. Since the exact makeup of children's natural language environment is unknown, we used

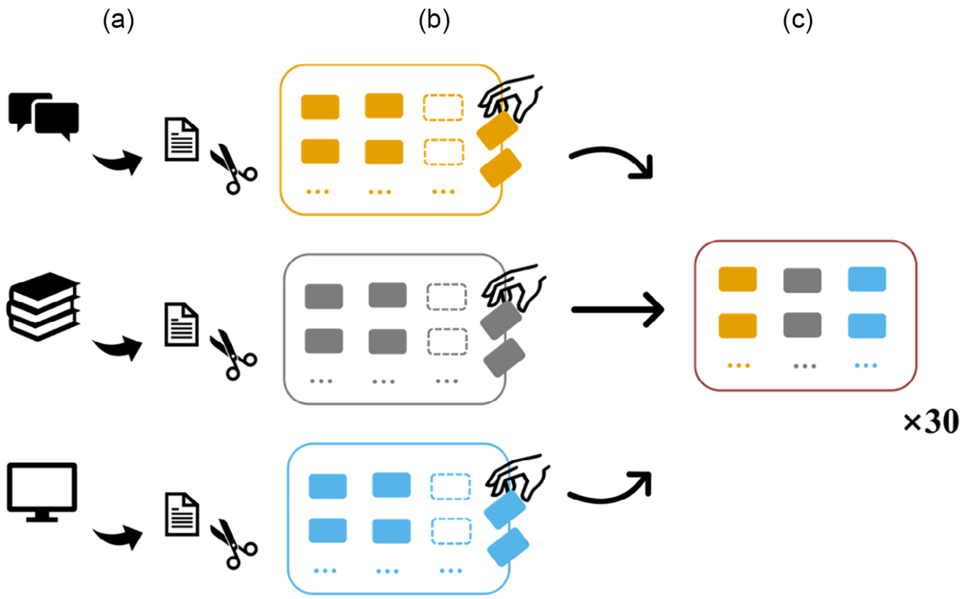

FIGURE 1 | Schematic representation of creating the pseudo-multimodal language environment.

Note: Each of the unimodal corpora was divided into multiple 150-word contexts (represented by the bricks in the figure) (a), which were used for computing CD and SD in each unimodal corpora, as well as for creating the pseudo-multimodal corpus; one-third of the contexts from each corpus were randomly sampled (b); the drawn samples were combined to create a pseudo-multimodal corpus (c). This sampling procedure was repeated 30 times, and the multimodal diversity metrics were computed as the averages of the 30 samples.

this balanced approach and generated 30 random samples, each averaging 1,633,348-word tokens and 24,698-word types.

The pseudo-multimodal sample was created to approximate lexical features that might exist in a multimodal language environment. Given the sampling method, we acknowledge that this sample is not an accurate representation of the textual or syntactic characteristics found in natural multimodal input, where speech, print, and media language often temporally co-occur and blend. Our aim with the pseudo-multimodal sample was to investigate whether a language sample derived from distinct unimodal corpora might provide a lexically diverse environment that reduces the inherent biases associated with each unimodal source. Therefore, caution is advised when interpreting findings related to the pseudo-multimodal sample's lexical diversity.

2.2 | Lexical Diversity Measures

For each corpus, we computed metrics to assess both global (type-token ratio) and local lexical diversity (connectivity, CD, and SD), as well as input frequency in the environment. Frequency, connectivity, and CD were computed for 8631 common words shared across corpora, while SD was computed for 2207 words common to all corpora. Input frequency and lexical diversity metrics of the pseudo-multimodal corpus were the average of the 30 samples. The main analyses were conducted on lemmas, or word forms, aligning with previous work (Dawson et al. 2021; Montag, Jones, and Smith 2015). However, the length of the documents in each corpus, that is, the conversation sessions in the speech corpus, picture books in the book corpus, and episodes or movies in the media corpus, differ drastically due to the nature of the material, which prevents a fair comparison. For this reason and for our purpose of creating the multimodal samples, as noted above, we divided each corpus into approximately 150-word contexts. Context-dependent measures, including connectivity, CD, and SD, were calculated based on 150-word contexts for all four corpora.

Additionally, to explore potential differences among lexical categories, input frequency, and CD were calculated for words with part of speech (POS) tags taken from the Penn Chinese Treebank (Xue et al. 2019). For this analysis, words were grouped into major lexical categories: nouns, lexical verbs, adjectives, adverbs, pronouns, and numerals. Other tags, including auxiliary verbs and other function words, were categorized as "other," resulting in a total of 9937 POS-tagged words.

2.2.1 | Input Frequency

A standardized frequency index for the common words was computed—*Zipf* value (van Heuven, Mandera, Keuleers, & Brysbaert, 2014), calculated as log10 (counts per million) +3. The Zipf value ranges from 1 to 7 with a Zipf = 1 (1 per 100 million) suggesting very low frequency; a Zipf = 6 (1 per 1000) representing very high frequency; and a Zipf between 3 (1 per million) and 4 (1 per 100,000) suggesting medium frequency. The standardized measure allows across-corpora comparison.

2.2.2 | Type-Token Ratio

Following Montag, Jones, and Smith (2015), we employed a bootstrapping approach by randomly sampling word sets of increasing sizes to generate a distribution of unique word type counts in each corpus. We started at 50,000 words, iteratively increased the sample size in increments of 50,000 words, and

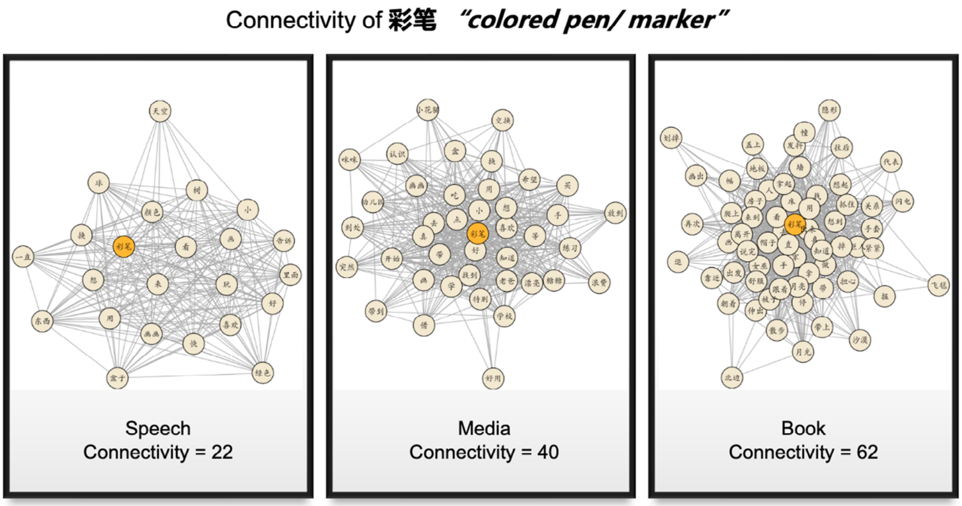

FIGURE 2 | An illustrative example of connectivity of the word 彩笔 (cai3 bi3, “colored pen/marker”) in children’s average speech, book, and media language environments.

Note: Connectivity values shown in the example were computed with the moving window size of 11 over the corpora. Function words were removed from the demonstration. When function words are included, the values are speech = 44, media = 81, and book = 98.

ended at the maximum tokens of the smallest corpus (i.e., 1.55 million words). This procedure was repeated 100 times at each sample size with replacement, resulting in 3100 random samples per corpus. Type-token ratio was the number of unique word types divided by the number of word tokens in each of the samples.

2.2.3 | Connectivity

A word’s local connectivity with other words were assessed by the number of unique words that likely co-occur with the word in the input, which reflects the direct proximity of two words in language exposure (Jiménez and Hills 2022b). A moving window of sizes of 5, 7, or 11 (i.e., 2, 3, and 5 words before and after the target) was slid through each corpus, resulting in a co-occurrence count matrix. The counts between words i and j were normalized into positive point-wise mutual information (PPMI) values using the formula (1). The matrix was then transformed into a binary matrix where 1 indicates presence and 0 indicates absence of connection between words i and j. Each word’s connectivity was calculated as the number of words with which it had a connection in the lower triangle of the matrix. Figure 2 shows the connectivity of an example word in the speech, book and media corpora. In the following, results for connectivity with the window size of 11 are reported. Analysis with the window sizes 5 and 7 have yielded same results and are reported in Supporting Materials.

$$\text{PPMI}(i, j) = \max\left(log\left(\frac{P(i \cap j)}{P(i)P(j)}\right), 0\right) \tag{1}$$

2.2.4 | Contextual Diversity

Traditionally, CD was the count of documents in which a word occurs, disregarding repetition within documents. As noted above, we divided each corpus into approximately 150-word contexts. The CD was then computed for each word by counting the number of unique contexts in which it occurs. We then log-transformed the raw counts.

2.2.5 | Semantic Diversity

A standard latent semantic analysis (LSA) procedure was performed to create a word × context high-dimensional matrix, where the same contexts for computing CD were used. The values in the matrix were log-transformed and divided by the word’s entropy (H) in the corpus using formula (2), where c indexes the contexts in which the word occurs, and p_c denotes the word’s frequency in the context divided by its total frequency in the corpus.

$$H = -\sum_{c} p_c \log(p_c) \tag{2}$$

Singular value decomposition was then used to reduce the high-dimensional matrix to 300 dimensions, producing a set of vectors for each word and a set of vectors for each context. Words that occurred in less than 10 contexts were excluded. For words with over 2000 contexts, following Hoffman, Lambon Ralph, and Rogers (2013), we randomly sampled 2000 for analysis to conserve computing resources. The average cosine similarity scores between any two contexts containing a word were log-transformed and sign reversed to represent the SD of the word. Because SD was not computed for words with very low frequency or that occurred in few contexts, there were fewer rarer words that had SD values compared to those with other diversity metrics. As a result, SD was only analyzed as an aspect of lexical diversity across modalities and not used to analyze the lexical development data.

2.3 | Lexical Development Data

Objective AoA for 435 words was obtained from Liu et al. (2011), of which 361 had lexical diversity measures and were analyzed.

In the original study, AoA was determined based on responses from 442 children's (age range: 2.4–11) performance in naming pictures. The children were assigned to seven age groups from preschoolers to Grade 3. Objective AoA for each word, which is the name of a picture, was calculated in the original research as AoA = age_m–4*(C%–0.75), where age_m refers to the mean age of the youngest group with more than 75% accuracy in naming the picture, and C% refers to the percentage of children who named the picture correctly in the age group. Accordingly, a picture name correctly identified by 75% of children with an average age of 8 years will have an AoA of 8 years, and a picture name correctly identified by 100% of the average 8-year-olds will have an AoA of 7 years. The second part of the equation, 4*(C%–0.75), was included to introduce variability in the AoA measure. This adjustment ensures that words named by 75% and 100% of children within the same age group are distinguished, allowing for finer-grained distinctions. By incorporating this term, we can differentiate between words with varying naming rates (e.g., 75% vs. 85%) and adjust the AoA values for items that were named correctly by less than 75% of the oldest group.

Based on the AoA computed by the original research, we further divided the words into three groups according to their AoA: 120 early- (AoA ≤ 4), 113 mid- (4< AoA ≤ 7), and 128 late-acquired (AoA > 7) words. Note that the words in these analyses are all concrete nouns because the AoA data was based on children's picture naming accuracies.

2.4 | Statistical Analysis

ANOVA was employed to examine the main effect of modality on frequency and local lexical diversity metrics. Effect sizes were estimated using partial eta squared ($\eta_p{}^2$). Subsequently, to test the difference between each pair of modalities, non-parametric Mann-Whitney *U* tests were performed, with alpha corrected for multiple comparisons using the Bonferroni method.

To investigate the effects of input frequency and lexical diversity (connectivity and CD) on word learning, we ran simple linear regression models with frequency and each diversity measure from each modality predicting AoA. SD was excluded from this analysis due to limited overlap between the words in the dataset and those having SD values in all modalities. Numeric variables were centered and scaled. Explained variance (R^2) and the Akaike information criteria (AIC) were used to compare the models. In these analyses, our primary focus is to determine which modality-specific diversity measures better explain lexical development.

Since we were also interested in exploring the relative contribution of lexical diversity across modalities, particularly at different developmental stages, we further built random forests regression models, where twelve predictors of interest (input frequency and 2 diversity metrics by 4 modalities) were simultaneously included. In random forest regressions, an ensemble of decision trees is created using bootstrapped samples and random predictor subsets. Out-of-bag samples were then used to estimate the error and to assess prediction accuracy. In our analysis, tree expansions were terminated at a pre-defined number of 500 trees, each with

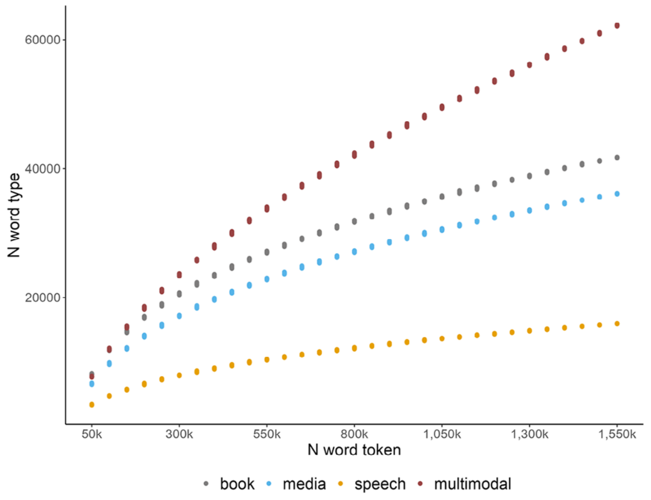

FIGURE 3 | Number of unique word types as a function of word tokens in samples across modalities.

Note: The graph shows the mean number of unique word types at different-sized samples of word tokens randomly selected from the four corpora (represented by different colors).

100 permutations and 3 predictors tried at each split. The relative importance of each predictor was estimated using increase in mean squared error (IncMSE%) when the predictor was permuted while keeping the other predictors constant, which indicates the impact of the given predictor on the predictive accuracy of the model. The significance of the predictors was obtained by permuting the response variables 100 times, using the *rfPermute* package (Archer 2023) in R. Additionally, to explore potential variations in the diversity effects across development, we divided the words into early-, mid-, and late-acquired categories according to their AoA and ran the random forest regressions separately for each of the three bins of words.

3 | Results

3.1 | Global Lexical Diversity Across Modalities

Figure 3 shows the numbers of unique word types as a function of word tokens in the samples in the four corpora. Type-token ratio values were the number of unique word types divided by the number of tokens in each sample (Table 1). Across samples of increasing tokens, the speech corpus contains the lowest type-token ratios, followed by media. The book corpus had higher type-token ratios than the other two unimodal corpora, but the pseudo-multimodal corpus contains the highest type-token ratios.

3.2 | Correlations of Frequency and Local Lexical Diversity Within and Between Modalities

Results of Pearson's correlation of input frequency (Zipf), connectivity (window size = 11), CD, and SD of common words are depicted in Figure 4. The highest correlations within the corpus were observed between frequency and CD, followed by correlations between frequency and connectivity, and then

TABLE 1 | Means and standard deviations (in parentheses) of frequency and lexical diversity metrics across modality.

	Speech	Print	Media	Pseudo-multimodal
***N* of word type**[a]	15,931 (13.6)	41,741 (31.1)	36,103 (25.9)	62,201 (51.1)
Input frequency	3.70 (0.83)	4.11 (0.77)	3.97 (0.78)	4.08 (0.71)
CD	0.77 (0.78)	1.21 (0.73)	1.07 (0.76)	1.16 (0.68)
SD	1.30 (0.34)	1.55 (0.28)	1.53 (0.31)	1.48 (0.26)
Connectivity (window size = 11)	84.8 (208)	203 (326)	153 (287)	302 (411)

[a]Number of unique word types in the 1,550,000-word-token samples.
Abbreviations: CD, contextual diversity; SD, semantic diversity.

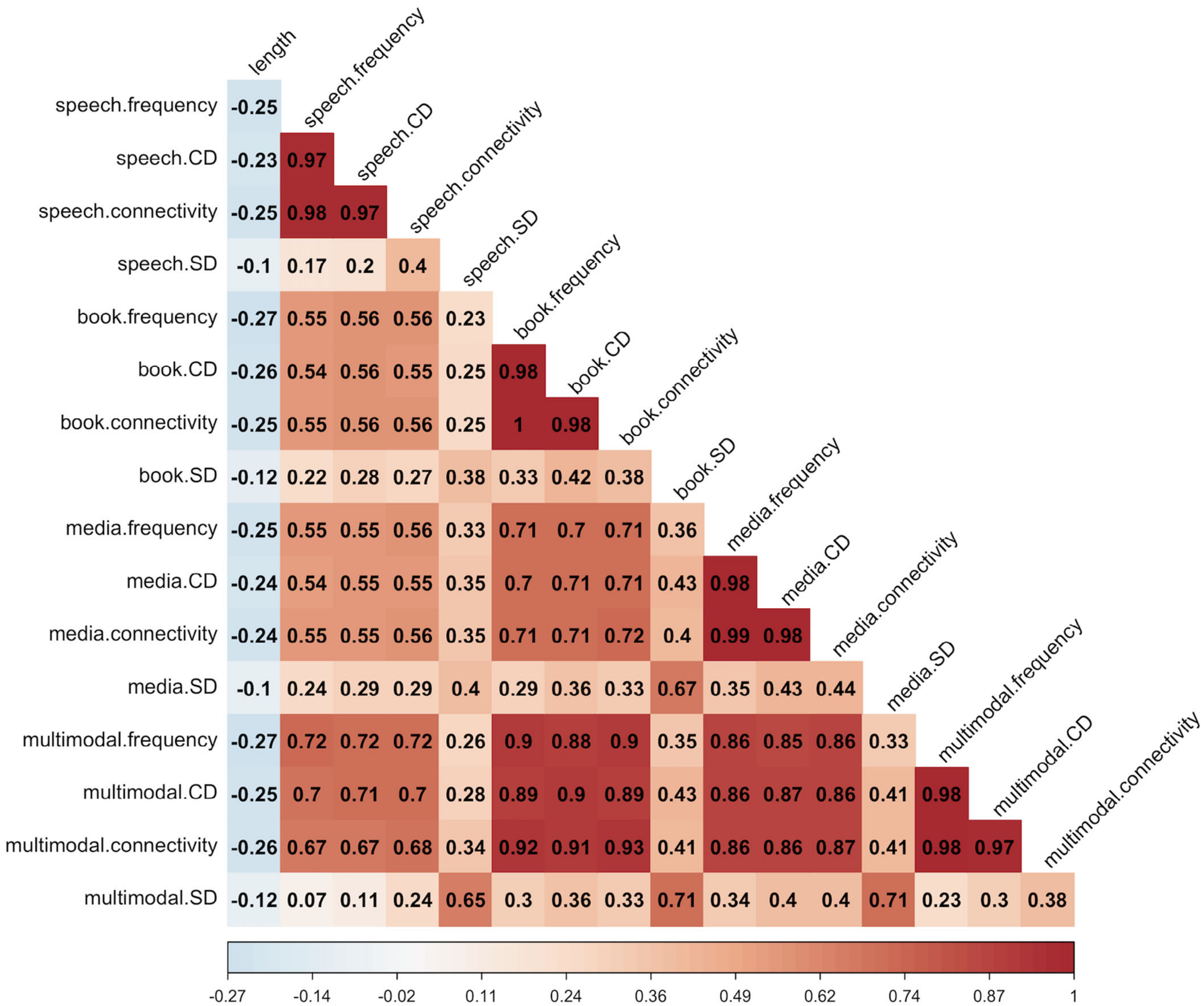

FIGURE 4 | Pearson's correlation coefficients among frequency and lexical diversity metrics.

Note: All *ps* < 0.001. Correlations among frequency, connectivity, and CD were computed on 8631 common words. Correlations among SD metrics and between SD and other metrics were computed on 2207 common words. CD indicates contextual diversity; SD, semantic diversity.

between connectivity and CD. SD exhibited the least strong correlations with other diversity metrics. Results for connectivity computed with window sizes of 5 and 7 are reported in Supporting Materials.

3.3 | Local Lexical Diversity Across Modalities

Distributions of input frequency and local lexical diversity across modalities are presented in Figure 5. Significant main effects of

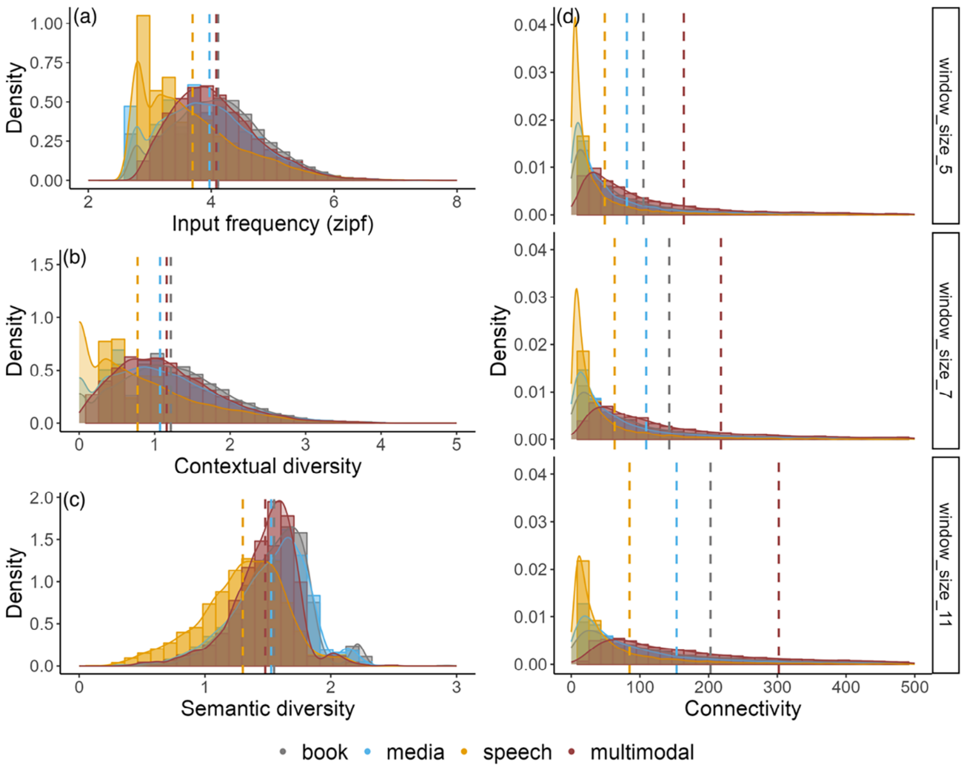

FIGURE 5 | Distributions of input frequency and local lexical diversity across corpora.

Note: (a) input frequency, (b) logarithmic CD, (c) SD, and (d) connectivity computed from moving window sizes 5, 7, and 11. Dashed lines indicate the mean. Frequency, connectivity and CD were computed for 8631 common words; SD was computed for 2207 common words. CD indicates contextual diversity; SD, semantic diversity.

corpus were found across all measures (frequency: F (3, 34,520) = 512.7, $p < 0.001$, $\eta_p^2 = 0.05$; CD: F (3, 34,520) = 609.7, $p < 0.001$, η_p^2 = 05; SD: F (3, 8824) = 319.9, $p < 0.001$, $\eta_p^2 = 0.10$; connectivity: F (3, 34,520) = 720.4, $p < 0.001$, $\eta_p^2 = 0.06$). Paired Mann-Whitney *U* tests revealed that for input frequency (Table 1, Figure 5a) and CD (Table 1, Figure 5b), the print modality exhibited the highest diversity, followed by multimodal, media, and speech environments. In SD (Table 1, Figure 5c), while book provided the most diverse meaningful variations for words, media followed very closely. Multimodal environment came next and speech environment was the least diverse. In contrast, the multimodal corpus had the highest average connectivity, followed by book, media, and speech (Table 1, Figure 5d). The adjusted alpha was set at 0.008 to correct for multiple comparisons, which was used for all following analyses comparing across modalities. All the above pairwise comparisons were significant ($ps < 0.001$).

Paired Mann-Whitney *U* tests were also conducted on frequency and CD for words by lexical categories. The results were largely the same as those of lemmas for content words (nouns, verbs, adjectives, and adverbs), numerals, and words classified as other: book had the highest frequency and CD, followed by multimodal, media, and then speech environments ($ps < 0.001$). The exception was that the difference between the media and multimodal corpora was not significant for adverbs (frequency: $p = 0.024$; CD: $p = 0.200$) or numerals (frequency: $p = 0.450$; CD: $p = 0.930$). Pronouns did not differ significantly in frequency or CD across modalities ($ps > 0.008$). Interjections had the highest frequency and CD in the multimodal environment, which differed significantly from speech and book ($ps < 0.001$) but not the media corpus (frequency: $p = 0.172$; CD: $p = 0.383$). Frequency and CD of interjections in the media input were also significantly higher than in the book ($p < 0.001$) but not in the speech environments ($p = 0.032$) (see Figure 6).

3.4 | Impacts of Input Frequency and Lexical Diversity on Lexical Development

All word-specific diversity metrics significantly predicted objective AoA. Speech measures explained the largest variance, followed by multimodal metrics, while book and media measures contributed similar explanations (Figure 7). Model comparison resulted in similar results. The model with speech CD had the lowest AIC among all the simple regression models, followed by speech frequency and then multimodal CD. Crucially, across all measures, higher diversity was always associated with earlier AoA.

3.5 | Lexical Diversity and Its Impact Across Development

The effect of developmental stage on frequency and lexical diversity was significant for the three diversity metrics (frequency: F (2, 1432) = 161.27, $p < 0.001$, $\eta_p^2 = 0.18$; CD: F (2, 1432) = 174.73, $p < 0.001$, $\eta_p^2 = 0.20$; connectivity: F (2, 1432) = 124.32, $p < 0.001$, $\eta_p^2 = 0.15$, Figure 8a), suggesting a general decreasing trend. Simple regression models were also run with frequency and each diversity metric predicted words' AoA by time bins

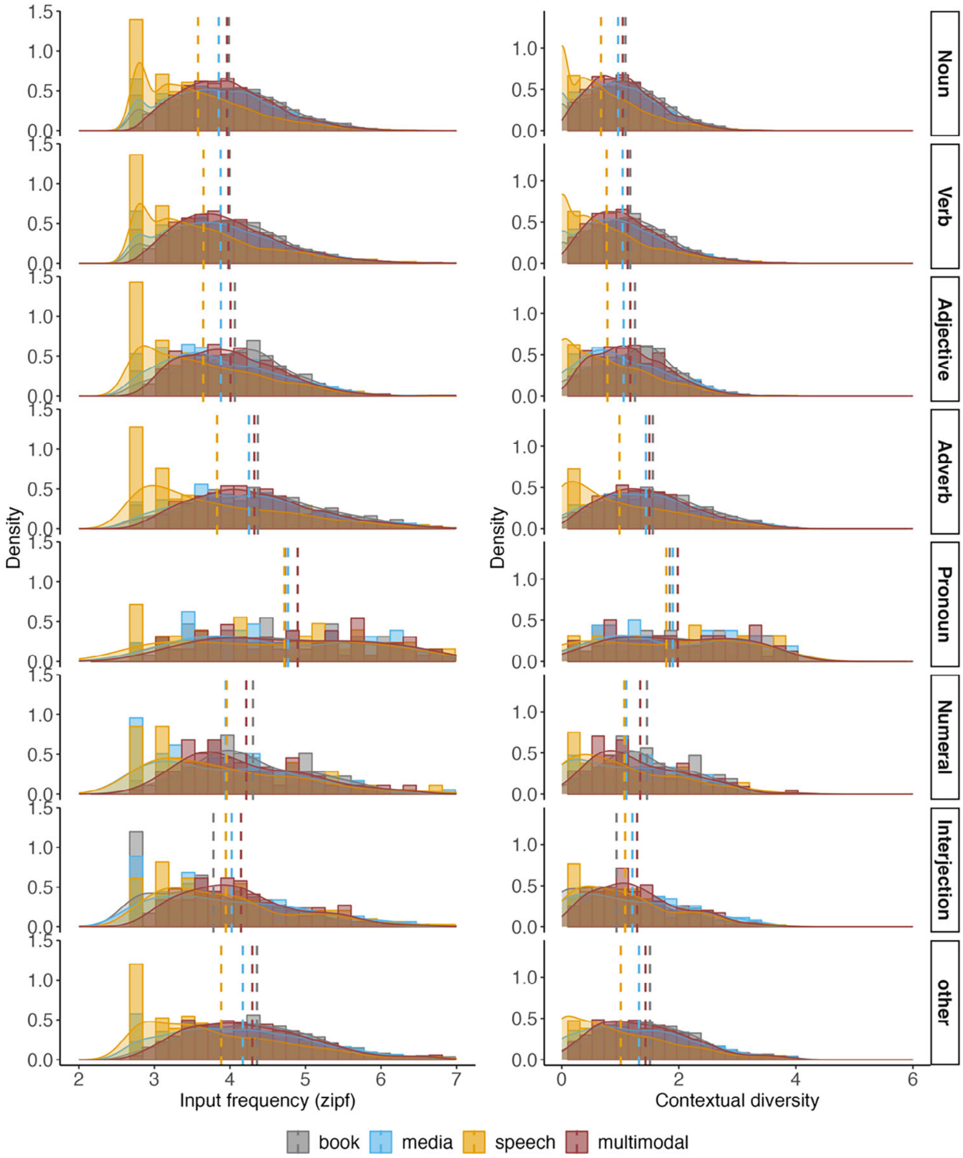

FIGURE 6 | Distributions of frequency and CD of common words by lexical category.

Note: The common words included 4147 nouns, 3320 lexical verbs, 722 adjectives, 606 adverbs, 77 pronouns, 103 number words, 170 interjections, and 791 other words. CD indicates contextual diversity.

(Figure 8b). For words acquired before age 4, speech diversity explained the largest variance (>20%), followed by multimodal and then book diversity. For words acquired between the ages of 4 and 7, speech diversity measures were still the strongest predictors, but the variances explained decreased to around 10%–15%. Variances explained by book and media diversity are reduced to lesser degrees for these mid-acquired words. In contrast, lexical diversity contributed much less to word learning after age 7, when none of the speech diversity measures was significant in predicting AoA, while CD of the other modalities, book frequency, and multimodal frequency remained significant predictors.

3.6 | Relative Importance of Diversity Measures on Lexical Development

In contrast to the simple regression analyses, we additionally ran random forest regressions to assess the relative importance of frequency and each diversity metric across modalities in predicting lexical development (Figure 9a). Speech diversity was the strongest predictor of objective AoA, with all measures in the speech modality having significant influences and their contributions ranked the most important. Book diversity was the least important and none of the metrics were significant.

This pattern held for words acquired in early childhood (Figure 9b). Words' connectivity in the speech modality had the highest impact not only for words acquired early but also for words acquired during the transition from early to middle childhood (i.e., between 4 and 7 years old). For the mid-acquired words (Figure 9c), multimodal frequency ranked as the second most important predictor, although its impact did not reach significance. In contrast, for words acquired after 7, the influence of speech diversity was no longer as profound, whereas multimodal CD had the strongest influence and was the only significant predictor (Figure 9d).

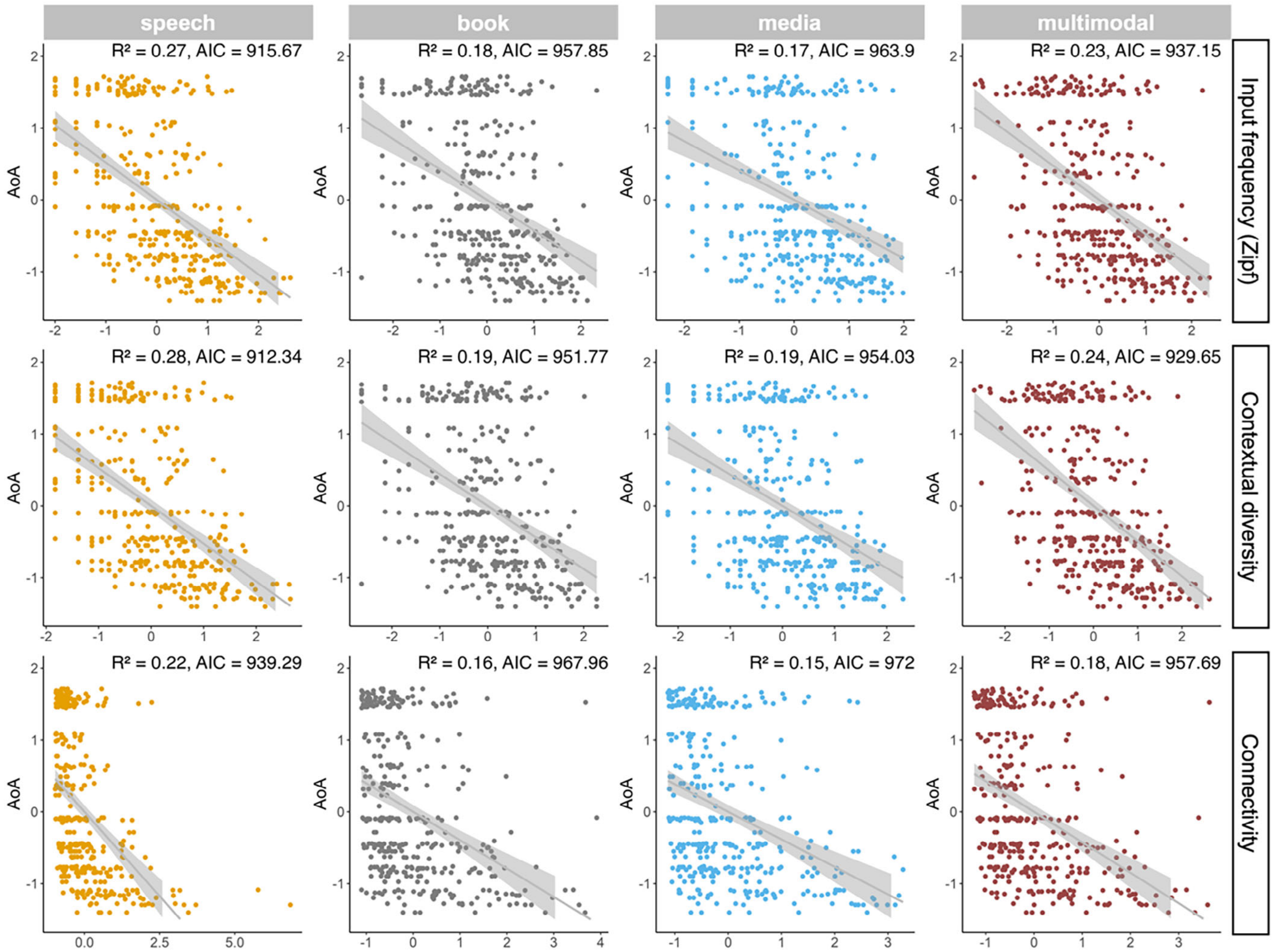

FIGURE 7 | Results of simple regression models with frequency, CD, and connectivity separately predicting words' AoA.

Note: Horizontal panel: input modality; vertical panel: lexical measures (predictors). Lower AIC indicates better model fit. CD indicates contextual diversity.

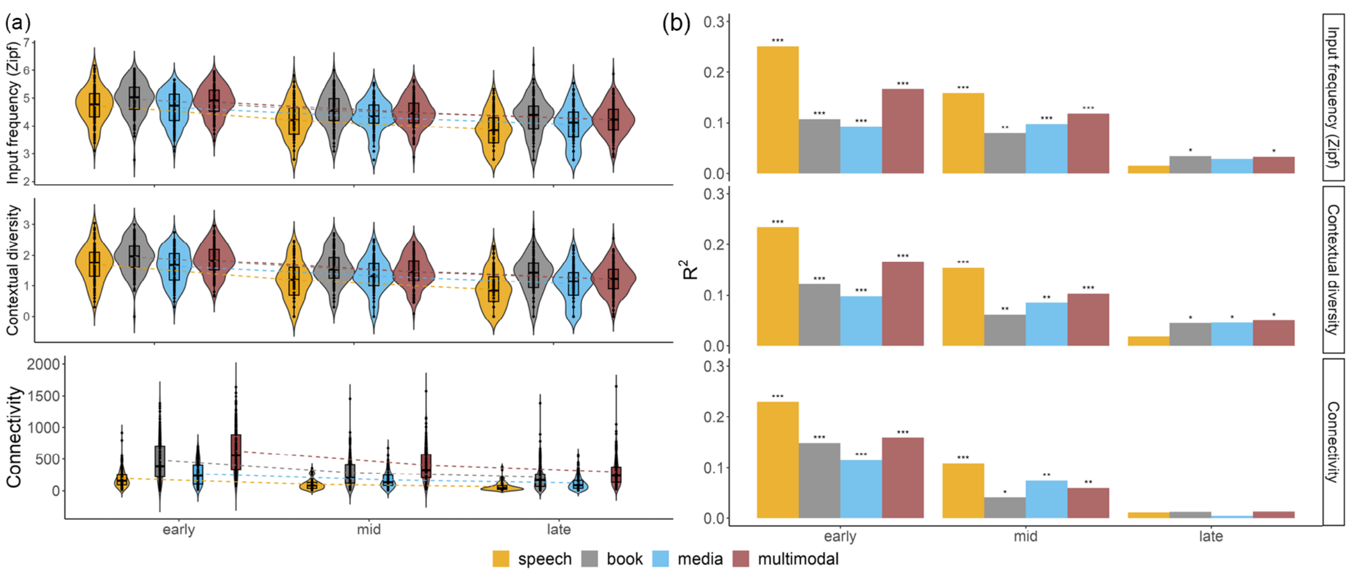

FIGURE 8 | Lexical diversity and their impacts on AoA by developmental time bins.

Note: (a) input frequency, CD, and connectivity of early- (AoA ≤ 4, $N = 120$); mid- (4 < AoA ≤ 7, $N = 113$); and late-acquired (AoA > 7, $N = 128$) words. (b) *R*-squared values of regression models with each diversity metric predicting words' AoA by developmental stage. *$p < 0.05$, **$p < 0.01$, and ***$p < 0.0$. CD indicates contextual diversity.

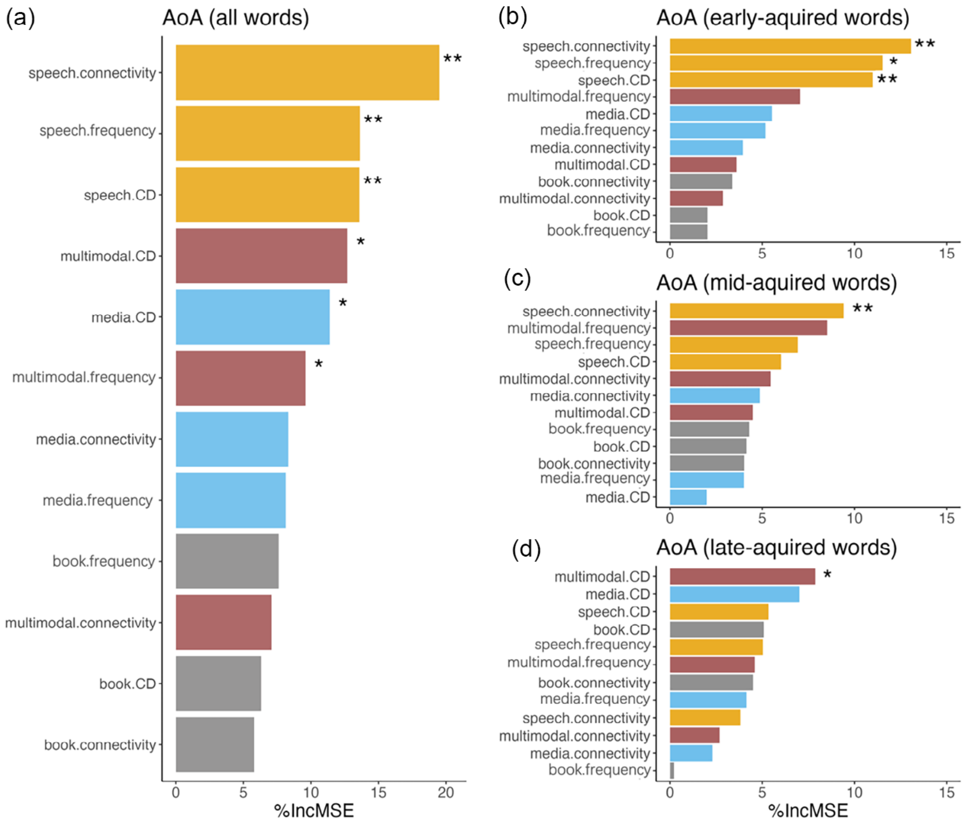

FIGURE 9 | Results of random forest regressions with lexical diversity metrics predicting AoA.

Note: %IncMSE: percentage increase in mean squared error. Larger %IncMSE indicates greater importance. *$p < 0.05$ and **$p < 0.01$.

4 | Discussion

The transition from early to middle childhood marks a crucial period characterized by exposure to diverse and multimodal language inputs. Our study reveals variations in the quality of language input across children's speech, print, and media language environments and their impacts on lexical development, which were compared with a pseudo-multimodal environment we compiled by combining the three input modalities. The pseudo-multimodal and the print environments demonstrated the highest global and local lexical diversity. Speech input, despite its lower diversity, significantly predicted lexical development. Our exploratory analysis showed that speech diversity was a robust predictor of word learning, particularly before the onset of middle childhood, after which other modalities exert increased influences. A crucial finding is the consistent association between higher lexical diversity and earlier word acquisition, underscoring the facilitative role of diverse linguistic exposure in lexical development.

The findings contribute to understanding how language input across modalities differentially influences language development. We extend previous research (Dawson et al. 2021; Montag, Jones, and Smith 2015) by suggesting that picture books not only contain a richer array of unique word types but also offer greater local lexical diversity, potentially facilitating lexical learning through words' broad connections with other words, varied contextual use and meanings. We also present novel evidence that children's animated media provide a lexically diverse experience, surpassing speech in overall diversity and featuring the highest diversity for interjections, which are pivotal in conversational turn-taking. This finding contributes to the ongoing debate on the impact of media exposure on child development (Jing et al. 2023; Madigan et al. 2019). The discrepancy may stem from the differential impacts of screen media on language versus other physical and cognitive development. Our findings suggest that child-targeted TV series and movies provide a lexical experience similar to print language while retaining conversational speech properties, like varied interjections. Additionally, interacting with screens has the potential to facilitate building links between what children see on the screen and their real-life experiences (Lauricella et al. 2010), thus enhancing the development of pragmatic language use. Nonetheless, it remains unknown whether watching or hearing language with high interactivity on screen (e.g., varied presence of interjections) could compensate for the lack of actual interactions that is otherwise abundant in real-world language interactions. Given the evidence for the facilitative effects of interactivity in child-directed speech on language development (Donnelly and Kidd 2021; Ramírez, Lytle, and Kuhl 2020), similar effects may occur with media language, highlighting critical questions for future research.

As noted, our method for creating the pseudo-multimodal corpus involved randomly drawing samples from the speech, media, and print corpora. The extent to which this sample approximates a true multimodal environment has not been validated. Therefore, findings related to the pseudo-multimodal corpus should be interpreted with caution. Despite its limitations, this approach allowed us to begin exploring lexical properties that may be present in a multimodal language environment. The high lexical diversity in the pseudo-multimodal condition implicates a potential benefit of multimodal input on lexical development, particularly in later stages. A combination of multiple sources of language input may mitigate the limited lexical diversity imposed by the homogeneity often inherent in single modalities. Furthermore, diversity of the input modality per se might introduce additional benefits.

The results thus underscore the value of a language environment that integrates varied modalities. Children's learning environment is inherently multimodal, not only in terms of the language input but also in the information they obtain through daily interactions that contribute to physical, language, and cognitive development. Recent research has utilized multimodal data to model individual infant's word learning (Roy et al. 2015; Vong et al. 2024). These models' success points directly to the necessity of considering multimodal sources that jointly influence language learning. However, such data from older children is currently lacking. Future studies could examine the effects of different input modalities by collecting longitudinal data on individual children's exposure. This would enable validation of the pseudo-multimodal approach and provide insight into how individual variation in multimodal input influences lexical development. Ongoing projects, such as the multimodal child-directed data project in our lab and Princeton's 1KD project, aim to collect child-centric data for modeling child development in real-world contexts. These initiatives are expected to enhance our understanding of how language, sensory, social, and environmental contacts interplay to shape language growth.

Our findings also contribute to addressing the debate over whether diversity facilitates or hinders language learning. Consistent with prior research (e.g., Alhama, Rowland, and Kidd 2023; Chang and Deák 2020; Goodman, Dale, and Li 2008; Hills 2013; Hills et al. 2010; Jones and Rowland 2017), we showed that high diversity in the lexical environment facilitates word learning beyond early childhood. This pattern was consistent across all metrics (input frequency, connectivity, and CD) and modalities. Studies reporting that diversity hinders word learning (Roy et al. 2015; Unger et al. 2024) have focused on toddlerhood (before 2.5 years), whereas our study examined a broader age range from early to beyond middle childhood (2.4–11 years). Our findings also add to a line of studies on the role of lexical diversity in written word acquisition in school-age children and adults, which have consistently shown facilitative effects (Hsiao and Nation 2018; Jones, Dye, and Johns 2017; Mak et al. 2021). This difference in age focus may explain the discrepant results. In an extensive review on the relationship between diversity and learning, Raviv, Lupyan, and Green (2022) hypothesized that high diversity makes learning difficult when learners are in the early stages of acquiring a target behavior; once a learner has surpassed the "novice" phase, high diversity would start to facilitate learning. This is, the impact of diversity likely changes with development. Early on, variability is detrimental to establishing a word's discriminative identity when young children are still acquiring word knowledge from scratch. As children's lexical knowledge matures, more variable learning contexts could lead to a broader coverage of the true environmental variation in words' usage, which accommodates future generalizations. Another potential drive for the change is that as children grow from infancy into middle childhood, the types of words they learn and the nature of words' variations are also changing. Importantly, this happens in tandem with their language and cognitive development, which might also make some word meanings more accessible than others.

It is worth mentioning that our analysis was limited to concrete nouns due to the limited availability of existing AoA data, while the impact of lexical diversity, as well as input frequency, on early lexical development, likely varies across word types. For example, the correlation between AoA and the input frequency of common nouns is higher than that between AoA and verbs, adjectives, or closed-class words in toddlerhood (Goodman, Dale, and Li 2008). Similarly, connectivity is more closely related to AoA for nouns than for verbs and other word classes (Alhama, Rowland, and Kidd 2023; Flores, Montag, and Willits 2023; Hills et al. 2010). Our study cannot determine whether the pattern of the relationship between lexical diversity and AoA generalizes to other word types. The acquisition of nouns, particularly concrete nouns with high imageability, is influenced by children's interactions with their surroundings, where the referents of the lexical items are perceivable. It is possible that not only the diversity of linguistic experience with words contributes to word learning, but also the diversified physical interactions further scaffold the formation of word-referent links. This is less applicable for other word types and low-imageability nouns. In comparison, the acquisition of other word classes might rely more on different types of diversity. For example, the diversity of syntactic frames in which verbs occur particularly predicts their order of acquisition, indicating a syntactic bootstrapping mechanism by which children use structural information in language to learn new words (Naigles and Hoff-Ginsberg 1998). Recent research also found that measures accounting for children's prior knowledge by computing the connectivity of a word within children's vocabulary relative to its overall connectivity provide a better fit to word learning for word types other than nouns (Flores, Montag, and Willits 2023). Future work should pursue how other lexical properties, such as word types and imageability/concreteness, modulate the impact of lexical diversity to better understand the mechanisms by which lexical experiences influence language development.

Our exploratory analysis revealed two developmental trends. First, the influence of input modalities on lexical learning changes as children grow. Diversity in the speech environment most strongly facilitated lexical learning, particularly in the early stages. However, its impact decreased, and other modalities became more prominent from middle childhood onward. This shift mirrors the transition in children's language exposure from being dominated by daily speech to a broader engagement with various forms of language input.

Although speech provided the least diversified lexical environment compared to other modalities, it is still most relevant to early language development. This may be because spoken language is often embedded in rich, interactive contexts, which helps children form strong associations between words and their referents, scaffolded by caregivers' interactive feedback. Such contextual grounding and interactive dynamics are crucial for early word learning, particularly for nouns, and may compensate for the lower lexical diversity in speech in the early years (Pexman 2019). As children's linguistic and cognitive capabilities develop, they are likely better equipped to benefit from the increased lexical diversity present in other input modalities, which might further enrich their lexical knowledge and support the acquisition of more complex and abstract words. Another possible explanation[4] for the finding is the different timeframes of the corpora. Although the media and book corpora were collected recently, both the speech corpus and the AoA data are over a decade old. This similarity in timing may account for why speech measures align more closely with AoA data. Although this possibility cannot be entirely ruled out, we note that the materials in the book and media corpora cover a broad timeframe, with some items published more than 10 years ago. Future research could improve comparability across modalities by using updated spoken language corpora that reflect contemporary language input.

The second developmental trend is that children's early word acquisition is closely tied to words' immediate linguistic contexts, shifting toward larger contextual windows to which children attend. During the early years, connectivity, which is based on words' co-occurrences within a window of a dozen words, was the primary predictor of words' AoA. A recent computational simulation with Skip-gram models has also found that the effect of semantic neighbors obtained with a window size of one word best predicted AoA in toddlers (Alhama, Rowland, and Kidd 2023). Conversely, in later stages, the most significant predictor of word learning was CD, computed with a much larger window size. These findings broadly indicate increasing memory and attention span characteristics of developing children (Elman 1993), which allows children to leverage a larger context for learning. It is important to note that the exploratory nature of our analysis and the potential limitations due to a restricted range of word items in each developmental phase means that the conclusions require cautious interpretation. There is a need for more robust, well-powered studies to confirm the findings and to fully understand the evolving relationship between lexical diversity and lexical development across childhood.

Finally, this study offers a unique contribution by exploring the lexical environment of non-English-speaking children exposed to the native Chinese language. The findings regarding lexical diversity and the impacts on lexical development align largely with previous results in English contexts, implicating universal patterns in lexical growth. However, there are known language-/cultural-specific distinctions in lexical diversity in children's language input, which may affect children's linguistic and cognitive development in unique ways. For instance, differences in the frequency and types of words in child-directed speech, such as a greater emphasis on verbs compared to nouns in Chinese than in English (Liu, Zhao, and Li 2008), could influence the acquisition trajectory of varying word classes. Extending this line of inquiry across different languages could further clarify the generalizability of our findings and enhance our understanding of language-specific developmental processes.

Although our study emphasizes the importance of multimodal lexical diversity in lexical development, it does not establish causality. Our analyses are based on child-directed corpora reflecting a snapshot of the average language environment but do not account for individual child-level factors or non-linguistic influences. Future research should incorporate child-centric data and computational modeling to delineate the relationship between lexical diversity in multimodal and unimodal environments and lexical development. We provide an initial exploration of the language environment of children moving from early to middle childhood from a cross-modal perspective and encourage future research to extend this line of inquiry across different languages to determine the generalizability of the findings.

Acknowledgments

This work was funded by the National Science and Technology Innovation 2030 (STI 2030—Major Projects 2021ZD0200500), the National Natural Science Foundation of China (31970987 to Qing Cai), and the National Social Science Fund of China (23CYY040 to Luan Li).

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data and the analysis scripts are publicly available on the Open Science Framework (https://osf.io/3tc5n/). The book and media corpora materials are not shared due to copyright restrictions.

Endnotes

[1] Other terms have been used to refer to what we call connectivity here, such as contextual diversity (Hills et al. 2010) and degree (Unger et al. 2024). To differentiate, we use connectivity to refer to words' connections with other words and contextual diversity to refer to words' variability in larger textual windows/documents.

[2] Again, other terms have been used to indicate contextual diversity, such as document diversity (Flores, Montag, and Willits 2023). Acknowledging the disagreement in term use, we use contextual diversity to refer to words' variability in documents.

[3] The G2 subcorpus of CCLOWW is a collection of picture books, Grades 1 and 2 textbooks and extracurricular books recommended for children under Grade 2. It comprises 1,554,243 word tokens and 37,516 word types.

[4] We thank an anonymous reviewer for suggesting this potential explanation.

References

Adelman, J., G. Brown, and J. Quesada. 2006. "Contextual Diversity, Not Word Frequency, Determines Word-Naming and Lexical Decision Times." *Psychological Science* 19, no. 9: 814–823. https://doi.org/10.1111/j.1467-9280.2006.01787.x.

Alhama, R. G., C. F. Rowland, and E. Kidd. 2023. "How Does Linguistic Context Influence Word Learning?." *Journal of Child Language* 50, no. 6: 1374–1393. https://doi.org/10.1017/S0305000923000302.

Archer, E. 2023. rfPermute: Estimate Permutation p-Values for Random Forest Importance Metrics. Retrieved from https://github.com/EricArcher/rfPermute.

Cameron-Faulkner, T., and C. Noble. 2013. "A Comparison of Book Text and Child Directed Speech." *First Language* 33, no. 3: 268–279. https://doi.org/10.1177/0142723713487613.

Chang, L. M., and G. O. Deák. 2020. "Adjacent Andnon-Adjacent Word Contexts Both Predict Age of Acquisition of English Words: A Distributional Corpus Analysis of Child-Directed Speech." *Cognitive Science* 44, no. 11: 1–24. https://doi.org/10.1111/cogs.12899.

Chang, Y. N., and C. Y. Lee. 2018. "Semantic Ambiguity Effects on Traditional Chinese Character Naming: A Corpus-Based Approach."

Behavior Research Methods 50, no. 6: 2292–2304. https://doi.org/10.3758/s13428-017-0993-4.

Dawson, N., Y. Hsiao, A. W. M. Tan, N. Banerji, and K. Nation. 2021. "Features of Lexical Richness in Children's Books: Comparisons With Child-Directed Speech." *Language Development Research* 1, no. 2021: 1–45. https://doi.org/10.34842/5we1-yk94.

Donnelly, S., and E. Kidd. 2021. "The Longitudinal Relationship Between Conversational Turn-Taking and Vocabulary Growth in Early Language Development." *Child Development* 92, no. 2: 609–625. https://doi.org/10.1111/cdev.13511.

Elman, J. L. 1993. "Learning and Development in Neural Networks: The Importance of Starting Small." *Cognition* 48, no. 1: 71–99. https://doi.org/10.1016/S0010-0277(02)00106-3.

Flores, A. Z., J. L. Montag, and J. A. Willits. 2023. "Using Known Words to Learn More Words: A Distributional Model of Child Vocabulary Acquisition." *Journal of Memory and Language* 132: 104446. https://doi.org/10.1016/j.jml.2023.104446.

Goodman, J. C., P. S. Dale, and P. Li. 2008. "Does Frequency Count? Parental Input and the Acquisition of Vocabulary." *Journal of Child Language* 35, no. 3: 515–531. https://doi.org/10.1017/S0305000907008641.

Hart, B., and T. Risley. 1995. *Meaningful Differences in the Everyday Experience of Young American Children*. Baltimore: Paul H. Brookes Publishing Co. https://doi.org/10.1007/s00431-005-0010-2.

Hills, T. 2013. "The Company That Words Keep: Comparing the Statistical Structure of Child- Versus Adult-Directed Language." *Journal of Child Language* 40, no. 3: 586–604. https://doi.org/10.1017/S0305000912000165.

Hills, T. T., J. Maouene, B. Riordan, and L. Smith. 2010. "The Associative Structure of Language: Contextual Diversity in Early Word Learning." *Journal of Memory and Language* 63: 259–273. https://doi.org/10.1016/j.jml.2010.06.002.

Hills, T. T., M. Maouene, J. Maouene, A. Sheya, and L. Smith. 2009. "Longitudinal Analysis of Early Semantic Networks: Preferential Attachment or Preferential Acquisition?." *Psychological Science* 20, no. 6: 729–739. https://doi.org/10.1111/j.1467-9280.2009.02365.x.

Hoff, E. 2003. "The Specificity of Environmental Influence: Socioeconomic Status Affects Early Vocabulary Development via Maternal Speech." *Child Development* 74, no. 5: 1368–1378. https://doi.org/10.1111/1467-8624.00612.

Hoff, E., and L. Naigles. 2002. "How Children Use Input to Acquire a Lexicon." *Child Development* 73, no. 2: 418–433. https://doi.org/10.1111/1467-8624.00415.

Hoffman, P., M. A. Lambon Ralph, and T. T. Rogers. 2013. "Semantic Diversity: A Measure of Semantic Ambiguity Based on Variability in the Contextual Usage of Words." *Behavior Research Methods* 45, no. 3: 718–730. https://doi.org/10.3758/s13428-012-0278-x.

Hsiao, Y., M. Bird, H. Norris, A. Pafan, and K. Nation. 2020. "The Influence of Item-Level Contextual History on Lexical and Semantic Judgments by Children and Adults." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 46, no. 12: 2367–2383. https://doi.org/10.1037/xlm0000795.

Hsiao, Y., N. J. Dawson, N. Banerji, and K. Nation. 2023. "The Nature and Frequency of Relative Clauses in the Language Children Hear and the Language Children Read: A Developmental Cross-Corpus Analysis of English Complex Grammar." *Journal of Child Language* 50, no. 3: 555–580. https://doi.org/10.1017/S0305000921000957.

Hsiao, Y., and K. Nation. 2018. "Semantic Diversity, Frequency and the Development of Lexical Quality in Children's Word Reading." *Journal of Memory and Language* 103: 114–126. https://doi.org/10.1016/j.jml.2018.08.005.

Huttenlocher, J., H. Waterfall, M. Vasilyeva, J. Vevea, and L. V. Hedges. 2010. "Sources of Variability in Children's Language Growth." *Cognitive Psychology* 61, no. 4: 343–365. https://doi.org/10.1016/j.cogpsych.2010.08.002.

Jiménez, E., and T. T. Hills. 2022a. "Differences in the Semantic Structure of the Speech Experienced by Late Talkers, Late Bloomers, and Typical Talkers." *Developmental Psychology* 59, no. 1: 141–160. https://doi.org/10.1037/dev0001446.

Jiménez, E., and T. T. Hills. 2022b. "Semantic Maturation During the Comprehension-Expression Gap in Late and Typical Talkers." *Child Development* 93, no. 6: 1727–1743. https://doi.org/10.1111/cdev.13815.

Jing, M., T. Ye, H. L. Kirkorian, and M. L. Mares. 2023. "Screen Media Exposure and Young Children's Vocabulary Learning and Development: A Meta-Analysis." *Child Development* 94, no. 5: 1398–1418. https://doi.org/10.1111/cdev.13927.

Johns, B. T., and M. N. Jones. 2022. "Content Matters: Measures of Contextual Diversity Must Consider Semantic Content." *Journal of Memory and Language* 123: 104313. https://doi.org/10.1016/J.JML.2021.104313.

Jones, G., and C. F. Rowland. 2017. "Diversity Not Quantity in Caregiver Speech: Using Computational Modeling to Isolate the Effects of the Quantity and the Diversity of the Input on Vocabulary Growth." *Cognitive Psychology* 98: 1–21. https://doi.org/10.1016/j.cogpsych.2017.07.002.

Jones, M. N., M. Dye, and B. T. Johns. 2017. "Context as an Organizing Principle of the Lexicon." *Psychology of Learning and Motivation* 67: 239–283. https://doi.org/10.1016/bs.plm.2017.03.008.

Jones, M. N., B. T. Johns, and G. Recchia. 2012. "The Role of Semantic Diversity in Lexical Organization." *Canadian Journal of Experimental Psychology* 66, no. 2: 115–124. https://doi.org/10.1037/a0026727.

Lauricella, A. R., T. A. Pempek, R. Barr, and S. L. Calvert. 2010. "Contingent Computer Interactions for Young Children's Object Retrieval Success." *Journal of Applied Developmental Psychology* 31, no. 5: 362–369. https://doi.org/10.1016/j.appdev.2010.06.002.

Li, L., Y. Yang, M. Song, M. Y. Zhang, Q. R. Chen, and Q. Cai. 2023. "CCLOWW: A Grade-Level Chinese Children's Lexicon of Written Words." *Behavior Research Methods* 55: 1874–1889. https://doi.org/10.3758/s13428-022-01890-9.

Li, L., W. Zhao, M. Song, J. Wang, and Q. Cai. 2023. "CCLOOW: Chinese Children's Lexicon of Oral Words." *Behavior Research Methods* 55: 1874–1889. https://doi.org/10.3758/s13428-023-02077-6.

Liu, S., X. Zhao, and P. Li. 2008. "Early Lexical Development: A Corpus-Based Study of Three Languages." In *Proceedings of the 30th Annual Conference of the Cognitive Science Society*, edited by B. C. Love, K. McRae, and V. M. Sloutsky, 1343–1348. Austin, TX: Cognitive Science Society.

Liu, Y., M. Hao, P. Li, and H. Shu. 2011. "Timed Picture Naming Norms for Mandarin Chinese." *PLoS ONE* 6, no. 1: e16505. https://doi.org/10.1371/JOURNAL.PONE.0016505.

MacWhinney, B. 2000. *The CHILDES Project: Tools for Analyzing Talk*. 3rd edition., Mahwah, NJ: Lawrence Erlbaum Associates.

Madigan, S., D. Browne, N. Racine, C. Mori, and S. Tough. 2019. "Association Between Screen Time and Children's Performance on a Developmental Screening Test." *JAMA Pediatrics* 173, no. 3: 244–250. https://doi.org/10.1001/jamapediatrics.2018.5056.

Mak, M. H. C., Y. Hsiao, and K. Nation. 2021. "Anchoring and Contextual Variation in the Early Stages of Incidental Word Learning During Reading." *Journal of Memory and Language* 118: 104203. https://doi.org/10.1016/j.jml.2020.104203.

Mar, R. A., J. L. Tackett, and C. Moore. 2010. "Exposure to Media and Theory-of-Mind Development in Preschoolers." *Cognitive Development* 25, no. 1: 69–78. https://doi.org/10.1016/j.cogdev.2009.11.002.

Mintz, T. H. 2003. "Frequent Frames as a Cue for Grammatical Categories in Child Directed Speech." *Cognition* 90, no. 1: 91–117. https://doi.org/10.1016/S0010-0277(03)00140-9.

Mizumoto, A. 2023. "Calculating the Relative Importance of Multiple Regression Predictor Variables Using Dominance Analysis and Random Forests." *Language Learning* 73, no. 1: 161–196. https://doi.org/10.1111/lang.12518.

Mol, S., and A. Bus. 2011. "To Read or Not to Read: A Meta-Analysis of Print Exposure From Infancy to Early Adulthood." *Psychological Bulletin* 137, no. 2: 267–296. https://doi.org/10.1037/a0021890.

Montag, J. L. 2019. "Differences in Sentence Complexity in the Text of Children's Picture Books and Child-Directed Speech." *First Language* 39, no. 5: 527–546. https://doi.org/10.1177/0142723719849996.

Montag, J. L., M. N. Jones, and L. B. Smith. 2015. "The Words Children Hear: Picture Books and the Statistics for Language Learning." *Psychological Science* 26, no. 9: 1489–1496. https://doi.org/10.1177/0956797615594361.

Montag, J. L., M. N. Jones, and L. B. Smith. 2018. "Quantity and Diversity: Simulating Early Word Learning Environments." *Cognitive Science* 42, no. 2: 375–412. https://doi.org/10.1111/COGS.12592.

Naigles, L. R., and E. Hoff-Ginsberg. 1998. "Why Are Some Verbs Learned Before Other Verbs? Effects of Input Frequency and Structure on Children's Early Verb Use." *Journal of Child Language* 25, no. 1: 95–120. https://doi.org/10.1017/S0305000997003358.

Nation, K. 2017. "Nurturing a Lexical Legacy: Reading Experience Is Critical for the Development of Word Reading Skill." *Npj Science of Learning* 2, no. 1: 1–4. https://doi.org/10.1038/s41539-017-0004-7.

Pagani, L. S., C. Fitzpatrick, and T. A. Barnett. 2013. "Early Childhood Television Viewing and Kindergarten Entry Readiness." *Pediatric Research* 74, no. 3: 350–355. https://doi.org/10.1038/pr.2013.105.

Pexman, P. M. 2019. "The Role of Embodiment in Conceptual Development." *Language, Cognition and Neuroscience* 34, no. 10: 1274–1283. https://doi.org/10.1080/23273798.2017.1303522.

Ramírez, N. F., S. R. Lytle, and P. K. Kuhl. 2020. "Parent Coaching Increases Conversational Turns and Advances Infant Language Development." *Proceedings of the National Academy of Sciences of the United States of America* 117, no. 7: 3484–3491. https://doi.org/10.1073/PNAS.1921653117/ASSET/0C93EA02-66C4-4564-B336-22B2B16CC762/ASSETS/GRAPHIC/PNAS.1921653117FIG03.JPEG.

Raviv, L., G. Lupyan, and S. C. Green. 2022. "How Variability Shapes Learning and Generalization." *Trends in Cognitive Sciences* 26, no. 6: 462–483. https://doi.org/10.1016/J.TICS.2022.03.007.

Rowe, M. L. 2012. "A Longitudinal Investigation of the Role of Quantity and Quality of Child-Directed Speech Vocabulary Development." *Child Development* 83, no. 5: 1762–1774. https://doi.org/10.1111/j.1467-8624.2012.01805.x.

Rowe, M. L., and A. Weisleder. 2020. "Language Development in Context." *Annual Review of Developmental Psychology* 2, no. 1: 201–223. https://doi.org/10.1146/ANNUREV-DEVPSYCH-042220-121816.

Roy, B. C., M. C. Frank, P. DeCamp, M. Miller, and D. Roy. 2015. "Predicting the Birth of a Spoken Word." *Proceedings of the National Academy of Sciences of the United States of America* 112, no. 41: 12663–12668. https://doi.org/10.1073/pnas.1419773112.

Sanchez, A., S. C. Meylan, M. Braginsky, K. E. MacDonald, D. Yurovsky, and M. C. Frank. 2019. "childes-db: A Flexible and Reproducible Interface to the Child Language Data Exchange System." *Behavior Research Methods* 51, no. 4: 1928–1941. https://doi.org/10.3758/s13428-018-1176-7.

Scheele, A., P. Leseman, and A. Mayo. 2010. "The Home Language Environment of Monolingual and Bilingual Children and Their Language Proficiency." *Applied Psycholinguistics* 31, no. 1: 117–140. https://doi.org/10.1017/S0142716409990191.

Unger, L., T. Chang, O. Savic, B. K. Bergen, and V. M. Sloutsky. 2024. "When Is a Word in Good Company for Learning?." *Developmental Science* 27, no. 5: e13510. https://doi.org/10.1111/desc.13510.

Unger, L., and A. V. Fisher. 2021. "The Emergence of Richly Organized Semantic Knowledge From Simple Statistics: A Synthetic Review." *Developmental Review* 60: 100949. https://doi.org/10.1016/j.dr.2021.100949.

Vong, W. K., W. Wang, A. E. Orhan, and B. M. Lake. 2024. "Grounded Language Acquisition Through the Eyes and Ears of a Single Child." *Science* 383, no. 6682: 504–511. https://doi.org/10.1126/SCIENCE.ADI1374.

Weizman, Z. O., and C. E. Snow. 2001. "Lexical Input as Related to Children's Vocabulary Acquisition: Effects of Sophisticated Exposure and Support for Meaning." *Developmental Psychology* 37, no. 2: 265–279. https://doi.org/10.1037/0012-1649.37.2.265.

Wurm, L. H., and S. A. Fisicaro. 2014. "What Residualizing Predictors in Regression Analyses Does (and What it Does not do)." *Journal of Memory and Language* 72, no. 1: 37–48. https://doi.org/10.1016/j.jml.2013.12.003.

Xue, N., F. Xia, F. Chiou, and M. Palmer. 2019. "The Penn Chinese Treebank: Phrase Structure Annotation of a Large Corpus." *Natural Language Engineering* 11, no. 2: 207–238. https://doi.org/10.1017/S135132490400364X.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.