Article

FIRST LANGUAGE

First Language
1–20

DOI: 10.1177/0142723719849996
journals.sagepub.com/home/fla
SAGE

Differences in sentence complexity in the text of children's picture books and child-directed speech

Jessica L. Montag
University of Illinois at Urbana-Champaign, USA

Abstract

Reading picture books to pre-literate children is associated with improved language outcomes, but the causal pathways of this relationship are not well understood. The present analyses focus on several syntactic differences between the text of children's picture books and typical child-directed speech, with the aim of understanding ways in which picture book text may systematically differ from typical child-directed speech. The analyses show that picture books contain more rare and complex sentence types, including passive sentences and sentences containing relative clauses, than does child-directed speech. These differences in the patterns of language contained in picture books and typical child-directed speech suggest that one important means by which picture book reading may come to be associated with improved language outcomes is by providing children with types of complex language that might be otherwise rare in their input.

Keywords

Corpus analysis, language development, passives, picture books, relative clauses

Parents are often advised to read to young children. This advice is ubiquitous and is given not only by educators and medical professionals, but also by celebrities and other public figures. Despite the folk origins of this advice, it is not without an empirical basis. Reading to young children is positively associated with language outcomes (Arterberry,

Corresponding author:
Jessica L. Montag, Department of Psychology, University of Illinois at Urbana-Champaign, 603 E. Daniel Street, Champaign, IL 61820, USA.
Email: jmontag@illinois.edu

Bornstein, Midgett, & Putnick, 2007; Demir-Lira, Applebaum, Goldin-Meadow, & Levine, 2018; Farrant & Zubrick, 2012; Fletcher & Reese, 2005; Karrass & Braungart-Rieker, 2005; Payne, Whitehurst, & Angell, 1994; Sénéchal & LeFevre, 2002) and literacy skills (Bus, Van Ijzendoorn, & Pellegrini, 1995; Deckner, Adamson, & Bakeman, 2006; Dickinson & Tabors, 1991; Lonigan, Burgess, & Anthony, 2000; Scarborough, Dobrich, & Hager, 1991; Shahaeian et al., 2018). However, the causal pathway by which reading comes to be associated with positive language and reading outcomes is not well understood. Picture book reading is currently and will likely continue to be an avenue by which large-scale interventions aim to improve language outcomes (Dickinson, Griffith, Golinkoff, & Hirsh-Pasek, 2012; Sharif, Rieber, & Ozuah, 2002; Zuckerman, 2009). However, for these interventions to be maximally effective, it is important to understand exactly *why* reading to children is associated with better language skills, to better adapt interventions to be the state of the language development art.

Much of the work investigating pathways by which shared book reading predicts better language and literacy outcomes has focused on the language generated when caregivers read to children. For example, picture books contain more unique words than child-directed speech (Hayes & Ahrens, 1988; Massaro, 2015; Montag, Jones, & Smith, 2015), and this lexical diversity does indeed seem to become a part of the spoken language input during book reading. Recordings of caregivers and children interacting in book reading contexts suggest that picture book reading provides children with more speech input and more lexically sophisticated speech than other caregiver–child activities (Crain-Thoreson, Dahlin, & Powell, 2001; Salo, Rowe, Leech, & Cabrera, 2016; Sosa, 2016; Weizman & Snow, 2001). These findings suggest that one means by which picture book reading may contribute to language outcomes is by exposing children to words that they might not otherwise encounter or might encounter less frequently.

The emphasis on the lexical contribution of picture books to the language environment is consistent with the tendency for early language outcomes and outcome disparities to be most often discussed in terms of lexical knowledge (e.g., Bates et al., 1994; Huttenlocher, Haight, Bryk, Seltzer, & Lyons, 1991; Rowe, 2012; Shneidman, Arroyo, Levine, & Goldin-Meadow, 2013; Weisleder & Fernald, 2014). Relatively less work has focused on the processing of multi-word utterances. However, early disparities in sentence processing are well-documented (Borovsky, Elman, & Fernald, 2012; Golinkoff, Ma, Song, & Hirsh-Pasek, 2013; Hoff-Ginsberg, 1986; Huang, Leech, & Rowe, 2017; Huttenlocher, Vasilyeva, Cymerman, & Levine, 2002). Further, much like how early vocabulary disparities are often attributed to differences in language input (Hoff, 2003; Hoff & Naigles, 2002; Hurtado, Marchman, & Fernald, 2008; Huttenlocher, Waterfall, Vasilyeva, Vevea, & Hedges, 2010; Weisleder & Fernald, 2014), early sentence processing disparities are associated with differences in language input as well (Hoff-Ginsberg, 1986; Huttenlocher et al., 2002; Lieven, 2010). The present investigation aims to extend analyses of the linguistic contribution of children's picture books into the syntactic domain. Differences in the sentence structures present in picture books and typical child-directed speech would suggest that picture books may be an important source of certain types of complex language, with possible consequences for early sentence processing.

Unlike with children, the contribution of texts to adults' sentence processing skills is well-established. A documented feature of speech and texts aimed at adults and young

readers is that texts contain far more syntactic complexity than speech (Biber, 1988; Montag & MacDonald, 2015; Roland, Dick, & Elman, 2007). Given the relative infrequency of complex sentences in speech, reading is an important source of experience with complex language for adults and young readers. There is mounting evidence that performance in laboratory-based sentence comprehension and production tasks are influenced by an individual's reading experience (Acheson, Wells, & MacDonald, 2008; Farmer, Fine, Misyak, & Christiansen, 2017; Mishra, Singh, Pandey, & Huettig, 2012; Montag & MacDonald, 2015; Payne, Gao, Noh, Anderson, & Stine-Morrow, 2012; Street & Dabrowska, 2010), suggesting that reading may account for a disproportionate amount of adults' experience with complex sentences. In young pre-readers, picture books may be an analogous source of complex language.

For many reasons, texts intended to be read aloud to young children may be a particularly important source of complex sentences. One reason is the potential parallels with adult-directed texts, which are generally more syntactically complex than adult-directed speech. A second reason is that caregivers appear to have minimal explicit control over the syntactic complexity of their child-directed utterances. Caregivers often increase the *lexical* diversity of their speech as their child's expressive vocabulary grows, but caregivers do not seem to increase the *syntactic* complexity of their speech in accordance with their child's own productive language skills (Huttenlocher et al., 2010). In other words, caregiver sentence complexity seems to unidirectionally drive children's learning of complex sentences and does not reflect an accommodation of the child's own abilities. Though syntactic complexity of caregiver utterances overall increases with child age and caregiver educational attainment (Huttenlocher, Vasilyeva, Waterfall, Vevea, & Hedges, 2007), caregivers may not have the same degree of explicit control over the type and complexity of sentence structures they use with children, as seems to be the case with word choice. Given both (1) overall differences between the spoken and written domain in terms of overall sentence complexity, and (2) the lack of strategic control that caregivers have over the syntactic complexity of their utterances, if the text of picture books is indeed more syntactically complex than typical child-directed speech, text may be an important means by which children encounter certain sentence types. If picture books are an important source of some rare or complex sentence forms, exposure to picture books may be an important factor to consider when investigating early sentence processing skills, with implications for individual differences.

Previous work has identified some syntactic differences between the text of picture books and child-directed speech. Cameron-Faulkner and Noble (2013) found that while some sentence types, such as questions, were more frequent in child-directed speech, complex utterances – defined as any utterance with two lexical verbs – were more frequent in picture book text. The present analyses investigate a larger corpus of child-directed speech and a larger set of picture books to expand on these findings and describe in greater detail the types and frequency of complex language in picture books and child-directed speech.

The present analyses investigate the frequencies of six of the most commonly investigated rare and complex sentence types: passive sentences and sentences containing one of five types of relative clauses in both child-directed speech and in the text of children's picture books. Examples of these sentence types are in Table 1, along with the corpus

Table 1. Sample sentences from the CHILDES corpus.

Passive	The milk has already been poured. (Brent Corpus, Miranda, 1;1.24)
Subject Relative (transitive verb)	Here's the kitty who likes that toy. (Brent Corpus, Henry, 0;9.24)
Subject Relative (intransitive verb)	This is the one that squeaks. (Brent Corpus, Alexander, 1;0.27)
Object Relative	What's that animal we saw at the zoo yesterday? (Bernstein Corpus, Gail, 2;1)
Oblique Relative	It's not a crayon you draw with. (Bloom-1970 Corpus, Peter, 2;8.12)
Passive Relative	That's stale old candy left over from Halloween. (Bloom-1970 Corpus, Peter, 2;0.10)

In parentheses, the corpus from which the example was found, the addressee child to whom the utterance was spoken, and that child's age. In all instances, the child's mother produced the utterance.

within CHILDES (MacWhinney, 2000) in which the example was found. These sentence types are among the most commonly studied sentence types in investigations of both language learning and adult psycholinguistics. Passive sentences are among the most commonly elicited or comprehended sentences in language studies with adults (e.g., Bock, 1987; Christianson & Ferreira, 2005; Dapretto & Bookheimer, 1999; Ferreira, 1994; Street & Dabrowska, 2010; Tanaka, Branigan, McLean, & Pickering, 2011) and are a key sentence type used in investigations of children's early sentence production and comprehension skills (Bever, 1970; Brooks & Tomasello, 1999; Huang et al., 2017; Huang, Zheng, Meng, & Snedeker, 2013; Maratsos, Fox, Becker, & Chalkley, 1985; Savage, Lieven, Theakston, & Tomasello, 2003). Finding that passive sentences were more common in picture books would have both theoretical and practical implications for understanding language learning trajectories.

Sentences containing relative clauses have also been the topic of thousands of research articles, in investigations of both adult sentence processing and in developmental trajectories of language acquisition. For decades, relative clauses have been widely seen as key test of human language abilities because they contain sentence embeddings (Chomsky & Miller, 1963). In addition to being a historically significant arena for debates and hypotheses about language use, relative clauses have also been an arena where the experiential bases of language processing is commonly studied. The role of experience via language input in relative clause processing skill is well attested in both children (Diessel & Tomasello, 2005; Kidd, Brandt, Lieven, & Tomasello, 2007; Roth, 1984) and adults (Desmet, De Baecke, Drieghe, Brysbaert, & Vonk, 2006; Gennari & MacDonald, 2009; Hsiao & MacDonald, 2016; Reali & Christiansen, 2007; Wells, Christiansen, Race, Acheson, & MacDonald, 2009). Consequently, there has been extensive work documenting the frequency and type of relative clauses that adults encounter (Gordon & Hendrick, 2005; Reali & Christiansen, 2007; Roland et al., 2007). There has been less work investigating the frequencies and types of relative clauses in children's language input. One exception is Diessel (2004), who investigated relative clause frequencies in four different mothers' speech to their children with the goal of tracking the relation between mothers'

and children's patterns of relative clause use. Diessel found that relative clauses were overall rare, but that there were substantial differences in the relative frequencies of different types of relative clauses in mothers' speech which were indeed reflected in their children's speech. Exposure to rare and complex sentence types via text may be an important aspect of the experiential bases by which complex language is learned and may shed insight into the sources of individual differences in the processing and production of complex language. The present analyses aim to better understand the role that texts may play in the development of complex language skills by investigating the frequencies of various rare and complex sentence types in a corpus of children's picture books and a corpus of child-directed speech.

Method

The sample of picture book text is from the corpus of 100 picture books (about 68,000 words), described in Montag et al. (2015). These books were selected to be representative of the books that caregivers might read to young children, *not* books that beginning readers might read to themselves. The titles were selected from recommended book lists, bestseller lists, and circulation statistics from the local public library. Given the small size of the picture book corpus, all relevant sentence types were identified manually.

The sample of child-directed speech was taken from the CHILDES corpus (MacWhinney, 2000). Specifically, the sample consisted of a subset of the North American CHILDES corpus: Bates (Bates, Bretherton, & Snyder, 1988), Bernstein (Bernstein, 1982), Bloom-1970 (Bloom, 1970; Bloom, Hood, & Lightbown, 1974), Bloom-1973 (Bloom, 1973), Bohannon (Bohannon & Marquis, 1977), and Brent (Brent & Siskind, 2001), and consisted of about 763,000 words. The target children in these corpora were between the ages of 5 and 37 months. These speech samples consist of transcripts of parents and other adults interacting with young children in a variety of contexts, including play-time and snack-time in the lab, or naturalistic recordings made in the home, including many longitudinal home recordings. Given that the goal of the present work was to estimate the syntactic complexity in child-directed speech, a range of speech contexts and speakers were selected to be included in these analyses, rather than focus on a single context or a smaller set of speakers. In analyses such as these, an important question is whether the recorded speech was indeed child-directed. In the recordings made in the lab, it is reasonably clear that the speech was child-directed. In the home recordings, it is possible that some of the speech was child-available but not specifically child-directed. However, the fact that the home recordings are not day-long recordings, but rather between a few minutes and a few hours long (many are about one hour), it may be less likely than with the day-long recordings that audio was being recorded while the caregiver was not interacting with the child.

To identify the relevant sentences containing relative clauses, the CLAN program was used to extract all complement modifications. Then, the target relative clause types were manually identified from the set of all sentences containing complement modifications. To identify simple passive sentences, the CLAN program was used to extract all sentences containing a past participle, and then all passives were manually identified from that set.

Given the extreme difficulty associated with identifying and coding relative clauses, to verify the identification and identity of relative clauses, a second coder independently coded a subset of the CHILDES sentences identified by CLAN as containing complement modifications. This individual coded 1810 sentences (33% of the total number of sentences extracted by CLAN) and coded for both the presence of a relative clause and the type of relative clause. Overall, there was 94.6% agreement in the identification of relative clauses ($\kappa = 0.54$). However, because the overall rate of relative clauses among the complement clauses identified by CLAN was very low (< 10%), there was some disagreement between coders despite this inter-rater reliability. About 60% of relative clauses identified by at least one coder were only identified by one coder. Upon further investigation, there were some systematic sources of disagreement between coders. Among those relative clauses that were only identified by one coder, 45% of those items contained an indefinite or otherwise 'nonspecific' head noun such as *the one you did, everything we need*, or *someone you know*, suggesting that these types of relative clauses, in particular, are easy to miss. Of the relative clauses identified by both coders, there was 87.3% agreement in relative clause type ($\kappa = 0.82$). Again, a pattern of items more likely to contain disagreement emerged. Of the items that contained disagreement, the overwhelming majority were discrepancies between object versus oblique relative clauses. The remaining discrepancies were errors in classifying the verb of subject relative clauses as either transitive or intransitive.

In response to these inter-coder discrepancies and the identification of items that were most likely to contain errors, the first coder (the author) reviewed all the sentences containing complement clauses from the CHILDES corpus and all sentences in the picture book corpus to double check for (1) relative clauses with indefinite or 'nonspecific' head nouns that might have been missed, (2) object relative clauses that should have been coded as oblique relative clauses, and vice versa, and (3) subject relative clauses miscoded for verb transitivity. All utterances included in these analyses, the CHILDES corpus from which each utterance came, and the age of the child to whom the utterance was addressed are available online at https://osf.io/tjs2e/.

Proportions of the six sentence types were computed per 1000 words, to provide an overall measure of the frequencies of these sentence types in language. In addition, frequencies of main clause passive sentences were computed per 1000 verb phrases and frequencies of the five relative clause types were computed per 1000 noun phrases. Calculating a proportion over noun or verb phrases yields a measure of how often a verb was passivized or a noun was relativized, given all opportunities for relativizing or passivizing. The numbers of noun phrases and verb phrases in CHILDES was calculated by using CLAN to count the number of nouns (including pronouns) and verbs (excluding auxiliary or modal verbs) in the child-directed speech. The number of noun and verb phrases in the picture books was calculated using the Python module NLTK (Bird, Loper, & Klein, 2009) to count the number of nouns (including pronouns) and verbs (excluding auxiliary or modal verbs) in the picture book text.

Results

All six sentence types were more frequent in the picture book text than in the child-directed speech. Figure 1 shows the frequency per 1000 words of each of the six sentence

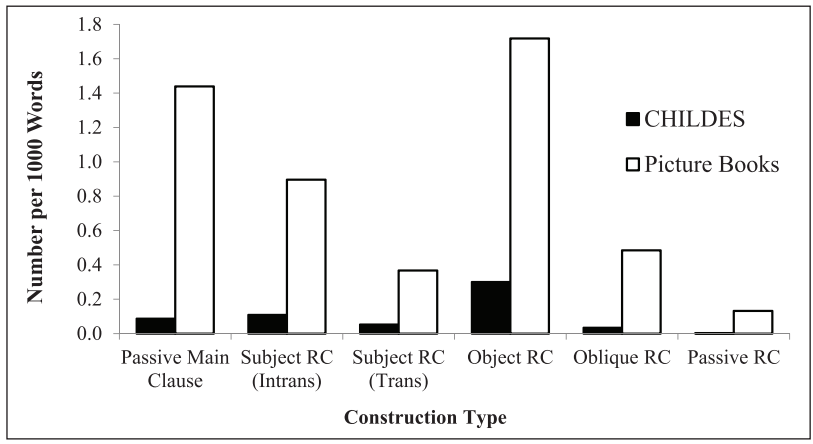

Figure 1. Frequency per 1000 words of the passive sentences and five relative clause types in the picture book text and child-directed speech.

Table 2. Frequency per 1000 words of the six sentence types, and the ratio of the frequencies in picture book text and child-directed speech.

	Child-directed speech	Picture books	Ratio
Passive	0.087	1.439	16.64
Subject RC (Intransitive)	0.109	0.896	8.24
Subject RC (Transitive)	0.052	0.367	7.00
Object RC	0.300	1.718	5.73
Oblique RC	0.034	0.485	14.22
Passive RC	0.003	0.132	50.43

types and Table 2 shows the frequencies and ratio of the frequencies in the two domains. These findings are consistent with previous analyses that compared a much smaller sample of picture book text to child-directed speech and found that picture books contain both more complex sentences, broadly defined, and more subject-predicate ('complete') sentences than speech (Cameron-Faulkner & Noble, 2013). These findings also show clear parallels to the adult literature, which suggests that passive sentences and various relative clause (RC) types are more common in written language than in spoken language directed at adult speakers (Biber, 1988; Roland et al., 2007).

The number of passives per 1000 noun phrases and relative clauses per 1000 verb phrases are given in Table 3. These frequency counts permit the calculation of the proportion and numbers of times a verb was passivized, or a noun was relativized, given all opportunities to do so, and whether these rates vary by domain (speech vs. text), via chi-square tests of independence. The results of these chi-square tests are also presented in Table 3, and show that passives and relative clauses were both more frequent in the picture book text than in the child-directed speech.

Table 3. Frequency of passivized verbs per 1000 verb phrases, relative clauses per 1000 noun phrases, ratio of those frequencies in speech and text, and count of each construction type.

	Child-directed speech	Picture books	Ratio	Raw count picture books	Raw count speech	Chi-square
Passive	0.575	7.067	12.280	64	98	$\chi^2(1) = 396.69$
Subject RC (Intransitive)	0.362	2.643	7.311	82	61	$\chi^2(1) = 186.71$
Subject RC (Transitive)	0.176	1.083	6.143	40	25	$\chi^2(1) = 62.81$
Object RC	1.005	5.070	5.043	228	117	$\chi^2(1) = 248.06$
Oblique RC	0.115	1.430	12.474	26	33	$\chi^2(1) = 148.00$
Passive RC	0.009	0.390	44.227	2	9	$\chi^2(1) = 60.75$

For the purposes of computing the chi-square values, child-directed speech contained 111,212 verb phrases and 226,805 noun phrases, and picture books contained 13,867 verb phrases and 23,077 noun phrases. All $p < .001$.

The relative frequencies of the different sentence types are also consistent with previous findings. Diessel (2004) investigated the frequencies of various relative clause types in the child-directed speech of four mothers. Overall, as in the present analyses, the object relative clauses were the most frequent relative clause type. In Diessel's analyses, object relative clauses were 1.6 times more frequent than subject relative clauses and 7.5 times more frequent than oblique relative clauses. In the present analyses object relative clauses were 1.9 times more frequent than subject relative clauses (merging together the transitive and intransitive verbs) and 8.8 times more frequent than oblique relative clauses. However, Diessel's overall frequencies were higher than those in the present analyses, with estimates of subject, object, and oblique relative clauses 2.8, 2.5, and 2.9 times more frequent than the estimates in the present analyses. For reference, these frequencies are in the ballpark of the frequencies for adult-directed subject and object relative clauses in the Switchboard corpus, as reported by Roland et al. (2007), which are still 2 to 3 times less frequent than the same sentences in the picture book text. It is unclear whether a methodological difference in how relative clauses were identified may account for the overall difference in frequency, or if Diessel's sample contained speakers or contexts more likely to produce relative clauses. These differences in estimated frequencies illustrate the potential sources of noise and true variability in these measures, and that more work is necessary to extract accurate frequency counts of various sentence types, in this and other corpora. Despite discrepancies in the present analyses and Diessel (2004), both analyses produced frequencies substantially lower than those in the children's picture books, so similar conclusions about the relative rates of complex sentences in speech versus text can be drawn from either analysis.

Crucially, the present findings do not suggest minor variations in the margins of a child's language experience, but rather, large and important contributions of the text of picture books, for those children who encounter this type of language experience. To provide a rough mathematical demonstration of scale, the average child hears very roughly 20,000 words per day (Hart & Risley, 1995; Shneidman et al., 2013), which would mean that this average child would hear about two passive sentences a day in spoken language.

The mean book length in the Montag et al. (2015) picture book corpus was about 680 words, which would on average contain one passive sentence. An increase of a single passive sentence per day may seem trivial, but given the overall low rate of passive sentences, it is equivalent to a 50% increase in the number of passive utterances that the average child would encounter. A single book a day is well within the expected experience of many but not all children. In nationally representative samples of U.S. caregivers, about half of all caregivers of children over 12 months reported reading to their child at least daily (Raikes et al., 2006; Yarosz & Barnett, 2001; Young, Davis, Schoen, & Parker, 1998). However, not all children will get this 50% boost in passive sentences; about a quarter of caregivers report rarely or never reading to their children (Bradley, Corwyn, McAdoo, & García Coll, 2001; Raikes et al., 2006; Yarosz & Barnett, 2001; Young et al., 1998). At the other extreme, caregivers in laboratory samples, who are generally older and more likely to hold a college degree, report reading to 27-month-old children an *average* of twice a day (Deckner et al., 2006), suggesting a very high upper-limit to the frequency with which children are read to. There exists enormous variability in the amount of picture book text a child might hear, but even a few books a week may account for a statistically important proportion of the complex sentences a child might hear.

An additional source of variability may be the degree to which the language contained in the picture books included in the present analyses are or are not representative of the language generated when caregivers read to children. Caregivers may not read all the text contained in the books, and of course, many picture books contain no text at all. Further, Hudson Kam and Matthewson's (2017) survey of picture book selections across households illustrates the enormous variability across families in the picture books that caregivers read to children. A 'representative' sample of picture books may be hard to define – over 70% of the picture book titles listed by survey respondents were listed by only one respondent and the most frequently listed book, *Goodnight Moon*, was listed by fewer than 20% of respondents. Given the variability across families in both the amount of reading and the selection of picture books that might be read, the above statistics may not be representative of any single family, but rather might be better interpreted as a useful central tendency at a population level.

Differences in lexico-syntactic combinations

Counts and ratios of sentence types may not tell a complete story because certain sentence types are systematically more likely to occur with certain words or classes of words. These lexico-syntactic combinations are an important feature of the syntactic properties of a given language sample, with important implications for language processing. Sentences of the same syntactic structure can nonetheless have very different patterns of comprehension and production as a consequence of the words used in those sentences, a phenomenon that is particularly well studied in complex sentences (Mak, Vonk, & Schriefers, 2002; Montag & MacDonald, 2015; Reali & Christiansen, 2007; Traxler, Morris, & Seely, 2002; Warren & Gibson, 2002). The lexico-syntactic combinations that exist in the language environment is an important aspect of that environment.

In addition to overall differences in the frequencies of passives and multiple types of relative clauses, there are substantial differences in the types of words contained in these

sentences across the two domains. Appendices A, B, and C contain information about the most relevant lexico-syntactic patterns of these complex sentences.

Appendix A contains raw counts and overall percentages of object relative clauses by embedded subject noun type (Full Noun Phrase: *The book that the teacher read*, Proper Noun: *The book that Alicia read*, or Pronoun: *The book that she read*), head noun animacy (*The woman that you saw; The tree that you saw*), and whether or not the relative clause was preceded by a relative pronoun (*The book that I read; The book Ø I read*). Consistent with previous corpus analyses, object relative clauses most frequently occur with embedded pronoun subjects, without relative pronouns, and when modifying inanimate head nouns (Montag & MacDonald, 2015; Roland et al., 2007; Temperley, 2003). In fact, the most commonly investigated type of object relative clause in the sentence processing literature, animate head nouns modified by full-NP subjects, as in the classic sentence *The reporter that the senator attacked admitted the error* (King & Just, 1991), does not appear in either the picture book text or the child-directed speech, and object relative clauses with any embedded full noun phrases are overall rare, though more frequent in the picture book sample (1.3% vs. 6.0%). These lexico-syntactic frequencies highlight the forms that relative clauses take in written and spoken language, and how these forms may or may not be like those used in studies of language processing and language development.

Appendix B contains raw counts and percentages of the remaining relative clause types by head noun animacy and the presence or absence of a relative pronoun. Again, patterns emerge such that certain lexico-syntactic combinations are far more frequent than others. Unlike the object relative clauses, subject relative clauses overwhelmingly appeared with relative pronouns, as did over half of the oblique relatives. This relative pronoun by relative clause type interaction suggests that corpus analyses that aim to identify relative clauses by the presence of a relative pronoun may systematically mis-count certain types of relative clauses and should be avoided. All relative clause types were more likely to modify inanimate head nouns. Again, there are clear lexico-syntactic patterns that emerge such that overall rates of various relative clause types may tell an incomplete story about the sorts of complex sentences that appear in children's language environments.

Finally, Appendix C contains raw counts and percentages of passive main clause sentences and relative clause sentences by presence or absence of an optional agentive *by*-phrase (main clause: *The picture was drawn by the child*; relative clause: *The picture that was drawn by the child*) and whether the passives were *get* or *be* passives (*The paper got torn*/*the paper was torn*). Both main clause passive sentences as well as in passive relative clauses overwhelmingly appeared without *by*-phrases, consistent with other analyses suggesting that agent omission is a common feature of passive use in both child-directed and adult-directed English (Biber, 1988; Gordon & Chafetz, 1990; Roland et al., 2007). Next, *get*-passives were more frequent in child-directed speech than in the picture book text (36% vs. 11% of passives), again consistent with previous findings that *get*-passives are more frequent in adult-directed spoken than written language (Collins, 1996).

Researchers interested in investigating patterns of comprehension or production of these complex sentences may want to refer to these statistics to understand lexico-syntactic patterns that children have the most experience with, to understand finer-grained detail about the complex sentences that appear in children's input, or to evaluate the consistency or inconsistency of experimental items with these patterns of experience.

General discussion

The present corpus analyses showed that the text of children's picture books contained more passive sentences and sentences containing relative clauses than did typical child-directed speech. This work adds to a growing body of work (Cameron-Faulkner & Noble, 2013; Hayes & Ahrens, 1988; Massaro, 2015; Montag et al., 2015) that describes ways in which the language of picture books varies from that of typical speech, and the potential consequences of those differences for the observed benefits of reading to young children. There are strong links between exposure to rare or complex sentence types, and subsequent comprehension and production of those sentence types in children and adults (Clark, 2003; Diessel & Tomasello, 2000; Huttenlocher et al., 2002; Montag & MacDonald, 2015). These links include evidence of a causal relationship where experimentally manipulated exposure to more of a particular sentence type improves comprehension of that sentence type (Roth, 1984; Vasilyeva, Huttenlocher, & Waterfall, 2006; Wells et al., 2009). Consistent with this experience-based approach to language learning, language input via picture books may be an important driver of individual differences in early language skills because these texts contain a much higher proportion of various rare and complex sentence types.

Children's knowledge of sentence structures can have additional, cascading, effects on other aspects of language development. Children use knowledge of sentence structure to aid the learning of new words that appear in those sentences (Landau & Gleitman, 1985; Naigles, 1996; Yuan & Fisher, 2009), suggesting a reciprocal relationship between word and sentence knowledge. Further, syntactic diversity itself – of sentences as well as the words that appear in particular sentence frames – may contribute to better learning of those words and sentences (Blackwell, 2005; Hsu, Hadley, & Rispoli, 2017; Naigles & Hoff-Ginsberg, 1998). Better, earlier, comprehension of a wider range of sentence structures may facilitate future learning of both words and sentences and syntactic knowledge may have reciprocal consequences for many different aspects of language learning, beyond simply the comprehension of those individual sentence types.

Beyond early oral language development, asymmetries in complex syntax between written and spoken language may have profound consequences when children learn to read. Children who were often read to before the onset of reading instruction will have more experience with complex sentences when they begin to read on their own. Consistent with the idea that early reading skills is a product of two separate skills: print knowledge and language knowledge (Dickinson, Golinkoff, & Hirsh-Pasek, 2010; Hoover & Gough, 1990; Kendeou, Van den Broek, White, & Lynch, 2009), young beginning readers must learn to master both sound–letter mappings as well as novel sentence structures that are common in the written domain. Children who were often exposed to complex sentences prior to reading instruction may be at an advantage because they will already be familiar with some of the complex sentence types that are common in written language but rare in typical oral language.

Despite the finding that there are differences between typical child-directed speech and picture book text, future work must aim to better understand how much of the linguistic complexity of the picture book text actually becomes part of a child's linguistic environment. There is evidence that complex sentences in picture book text do indeed

become part of the child's language environment (Noble, Cameron-Faulkner, & Lieven, 2018). However, picture book reading is a complex, multifaceted activity. When reading books to children, caregivers read the text of the book out loud, but they also point to and label pictures, paraphrase the text, expand upon or comment on the text, ask and answer questions, and engage in a range of other extra-text speech (Deckner et al., 2006; Fletcher, Cross, Tanney, Schneider, & Finch, 2008; Hudson Kam & Matthewson, 2017; Ninio & Bruner, 1978; Whitehurst et al., 1988). To understand how and why picture book reading benefits young children, we have to understand the contribution of these multiple factors – how the book text, extra-text utterances, and the pictures all contribute to the learning environment. The goal of the present work is to provide information about how one of these factors, the text of the book, may contribute to the language environment.

If it is the case that sentence structure, specifically, is an important feature of picture books that contributes to language outcomes, there are important consequences for the use of picture books as a language intervention. If one of the reasons that picture book reading is associated with positive language outcomes is that it exposes children to complex sentences, then reading the text in a manner that preserves this complex language may be important for the intervention's maximum efficacy. The goal of keeping complex language intact may, but need not, compete with the goals of other intervention strategies, such as 'dialogic reading,' which emphasizes caregiver–child interaction and conversation during picture book reading (Arnold, Lonigan, Whitehurst, & Epstein, 1994; Whitehurst et al., 1988). De-emphasizing the importance of the picture book text may discourage caregivers from reading complex sentences as they appear in the text even if it encourages other types of caregiver–child interaction and language, which provide a different set of valuable learning opportunities (Blewitt, Rump, Shealy, & Cook, 2009; Mol, Bus, de Jong, & Smeets, 2008; Snow & Ninio, 1986). The types of language and conversation that are generated by different picture book reading techniques, the advantages and disadvantages of each, and their role in predicting individual variability in language outcomes, are important remaining questions.

Future work investigating the mechanisms by which picture book reading aids in language learning will likely focus on many aspects of shared book reading, but an important part of the story will likely be the text of the books, and the language that those books provide that is rare or absent in spontaneous spoken language. Understanding the pathways by which reading picture books to young children leads to better language outcomes has both basic and applied implications. This knowledge will help us understand, broadly, what children learn from the language they experience, and how the words and sentences contained in picture books may or may not contribute to those experiences. Then, by understanding how aspects of the language environment contribute to language learning, we can leverage this knowledge to construct effective interventions and empirically-sound advice for parents and caregivers.

Acknowledgements

I thank Lindy Comstock for assistance coding the relative clauses.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article: This work was supported by NIH Grant R03 HD096157, NSF grant BCS-1749594 and a James S. McDonnell Foundation Scholar Award to JLM.

ORCID iD


Jessica L. Montag https://orcid.org/0000-0001-9446-1016

References

Acheson, D. J., Wells, J. B., & MacDonald, M. C. (2008). New and updated tests of print exposure and reading abilities in college students. *Behavior Research Methods*, *40*, 278–289.

Arnold, D. H., Lonigan, C. J., Whitehurst, G. J., & Epstein, J. N. (1994). Accelerating language development through picture book reading: Replication and extension to a videotape training format. *Journal of Educational Psychology*, *86*, 235–243.

Arterberry, M. E., Bornstein, M. H., Midgett, C., & Putnick, D. L. (2007). Early attention and literacy experiences predict adaptive communication. *First Language*, *27*, 175–189.

Bates, E., Bretherton, I., & Snyder, L. (1988). *From first words to grammar: Individual differences and dissociable mechanisms*. Cambridge, MA: Cambridge University Press.

Bates, E., Marchman, V., Thal, D., Fenson, L., Dale, P., Reznick, J. S., . . . Hartung, J. (1994). Developmental and stylistic variation in the composition of early vocabulary. *Journal of Child Language*, *21*, 85–123.

Bernstein, N. (1982). *Acoustic study of mothers' speech to language-learning children: An analysis of vowel articulatory characteristics* (Unpublished doctoral dissertation). Boston University, Boston, MA.

Bever, T. G. (1970) The cognitive basis for linguistic structures. In J. R. Hayes (Ed.), *Cognition and the development of language* (pp. 279–352). New York, NY: John Wiley.

Biber, D. (1988). *Variation across speech and writing*. New York, NY: Cambridge University Press.

Bird, S., Klein, E., & Loper, E. (2009). *Natural language processing with Python: Analyzing text with the natural language toolkit*. Sebastopol, CA: O'Reilly Media, Inc.

Blackwell, A. A. (2005). Acquiring the English adjective lexicon: Relationships with input properties and adjectival semantic typology. *Journal of Child Language*, *32*, 535–562.

Blewitt, P., Rump, K. M., Shealy, S. E., & Cook, S. A. (2009). Shared book reading: When and how questions affect young children's word learning. *Journal of Educational Psychology*, *101*, 294–304.

Bloom, L. (1970). *Language development: Form and function in emerging grammars*. Cambridge, MA: MIT Press.

Bloom, L. (1973). *One word at a time: The use of single-word utterances before syntax*. The Hague, The Netherlands: Mouton.

Bloom, L., Hood, L., & Lightbown, P. (1974). Imitation in language development: If, when and why. *Cognitive Psychology*, *6*, 380–420.

Bock, K. (1987). An effect of the accessibility of word forms on sentence structures. *Journal of Memory and Language*, *26*, 119–137.

Bohannon, J. N., & Marquis, A. L. (1977). Children's control of adult speech. *Child Development*, *48*, 1002–1008.

Borovsky, A., Elman, J. L., & Fernald, A. (2012). Knowing a lot for one's age: Vocabulary skill and not age is associated with anticipatory incremental sentence interpretation in children and adults. *Journal of Experimental Child Psychology*, *112*, 417–436.

Bradley, R. H., Corwyn, R. F., McAdoo, H. P., & García Coll, C. (2001). The home environments of children in the United States part I: Variations by age, ethnicity, and poverty status. *Child Development*, *72*, 1844–1867.

Brent, M. R., & Siskind, J. M. (2001). The role of exposure to isolated words in early vocabulary development. *Cognition*, *81*, 31–44.

Brooks, P. J., & Tomasello, M. (1999). Young children learn to produce passives with nonce verbs. *Developmental Psychology*, *35*, 29–44.

Bus, A. G., Van Ijzendoorn, M. H., & Pellegrini, A. D. (1995). Joint book reading makes for success in learning to read: A meta-analysis on intergenerational transmission of literacy. *Review of Educational Research*, *65*, 1–21.

Cameron-Faulkner, T., & Noble, C. (2013). A comparison of book text and child directed speech. *First Language*, *33*, 268–279.

Chomsky, N., & Miller, G. A. (1963). Introduction to the formal analysis of natural language. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Mathematical psychology* (pp. 269–320). New York, NY: John Wiley.

Christianson, K., & Ferreira, F. (2005). Conceptual accessibility and sentence production in a free word order language (Odawa). *Cognition*, *98*, 105–135.

Clark, E. V. (2003). *First language acquisition*. Cambridge, UK: Cambridge University Press.

Collins, P. C. (1996). Get-passives in English. *World Englishes*, *15*, 43–56.

Crain-Thoreson, C., Dahlin, M. P., & Powell, T. A. (2001). Parent–child interaction in three conversational contexts: Variations in style and strategy. *New Directions for Child and Adolescent Development*, *2001*, 23–38.

Dapretto, M., & Bookheimer, S. Y. (1999). Form and content: Dissociating syntax and semantics in sentence comprehension. *Neuron*, *24*, 427–432.

Deckner, D. F., Adamson, L. B., & Bakeman, R. (2006). Child and maternal contributions to shared reading: Effects on language and literacy development. *Journal of Applied Developmental Psychology*, *27*, 31–41.

Demir-Lira, Ö E., Applebaum, L. R., Goldin-Meadow, S., & Levine, S. C. (2018). Parents' early book reading to children: Relation to children's later language and literacy outcomes controlling for other parent language input. *Developmental Science*, *22*, e12764.

Desmet, T., De Baecke, C., Drieghe, D., Brysbaert, M., & Vonk, W. (2006). Relative clause attachment in Dutch: On-line comprehension corresponds to corpus frequencies when lexical variables are taken into account. *Language and Cognitive Processes*, *21*, 453–485.

Dickinson, D. K., Golinkoff, R. M., & Hirsh-Pasek, K. (2010). Speaking out for language why language is central to reading development. *Educational Researcher*, *39*, 305–310.

Dickinson, D. K., Griffith, J. A., Golinkoff, R. M., & Hirsh-Pasek, K. (2012). How reading books fosters language development around the world. *Child Development Research*, 2012, Article 602807.

Dickinson, D. K., & Tabors, P. O. (1991). Early literacy: Linkages between home, school and literacy achievement at age five. *Journal of Research in Childhood Education*, *6*, 30–46.

Diessel, H. (2004). *The acquisition of complex sentences*. Cambridge, UK: Cambridge University Press.

Diessel, H., & Tomasello, M. (2000). The development of relative clauses in spontaneous child speech. *Cognitive Linguistics*, *11*, 131–152.

Diessel, H., & Tomasello, M. (2005). A new look at the acquisition of relative clauses. *Language*, *81*, 882–906.

Farmer, T. A., Fine, A. B., Misyak, J. B., & Christiansen, M. H. (2017). Reading span task performance, linguistic experience, and the processing of unexpected syntactic events. *The Quarterly Journal of Experimental Psychology*, *70*, 413–433.

Farrant, B. M., & Zubrick, S. R. (2012). Early vocabulary development: The importance of joint attention and parent-child book reading. *First Language*, *32*, 343–364.

Ferreira, F. (1994). Choice of passive voice is affected by verb type and animacy. *Journal of Memory and Language*, *33*, 715–736.

Fletcher, K. L., Cross, J. R., Tanney, A. L., Schneider, M., & Finch, W. H. (2008). Predicting language development in children at risk: The effects of quality and frequency of caregiver reading. *Early Education and Development*, *19*, 89–111.

Fletcher, K. L., & Reese, E. (2005). Picture book reading with young children: A conceptual framework. *Developmental Review*, *25*, 64–103.

Gennari, S. P., & MacDonald, M. C. (2009). Linking production and comprehension processes: The case of relative clauses. *Cognition*, *111*, 1–23.

Golinkoff, R. M., Ma, W., Song, L., & Hirsh-Pasek, K. (2013). Twenty-five years using the intermodal preferential looking paradigm to study language acquisition: What have we learned? *Perspectives on Psychological Science*, *8*, 316–339.

Gordon, P., & Chafetz, J. (1990). Verb-based versus class-based accounts of actionality effects in children's comprehension of passives. *Cognition*, *36*, 227–254.

Gordon, P. C., & Hendrick, R. (2005). Relativization, ergativity, and corpus frequency. *Linguistic Inquiry*, *36*, 456–463.

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experiences of young American children*. Baltimore, MD: Brookes.

Hayes, D. P., & Ahrens, M. G. (1988). Vocabulary simplification for children: A special case of 'motherese'? *Journal of Child Language*, *15*, 395–410.

Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development*, *74*, 1368–1378.

Hoff, E., & Naigles, L. (2002). How children use input to acquire a lexicon. *Child Development*, *73*, 418–433.

Hoff-Ginsberg, E. (1986). Function and structure in maternal speech: Their relation to the child's development of syntax. *Developmental Psychology*, *22*, 155–163.

Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing*, *2*, 127–160.

Hsiao, Y., & MacDonald, M. C. (2016). Production predicts comprehension: Animacy effects in Mandarin relative clause processing. *Journal of Memory and Language*, *89*, 87–109.

Hsu, N., Hadley, P. A., & Rispoli, M. (2017). Diversity matters: Parent input predicts toddler verb production. *Journal of Child Language*, *44*, 63–86.

Huang, Y. T., Leech, K., & Rowe, M. L. (2017). Exploring socioeconomic differences in syntactic development through the lens of real-time processing. *Cognition*, *159*, 61–75.

Huang, Y. T., Zheng, X., Meng, X., & Snedeker, J. (2013). Children's assignment of grammatical roles in the online processing of Mandarin passive sentences. *Journal of Memory and Language*, *69*, 589–606.

Hudson Kam, C. L., & Matthewson, L. (2017). Introducing the Infant Bookreading Database (IBDb). *Journal of Child Language*, *44*, 1289–1308.

Hurtado, N., Marchman, V. A., & Fernald, A. (2008). Does input influence uptake? Links between maternal talk, processing speed and vocabulary size in Spanish-learning children. *Developmental Science*, *11*, F31–F39.

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology*, *27*, 236–248.

Huttenlocher, J., Vasilyeva, M., Cymerman, E., & Levine, S. (2002). Language input and child syntax. *Cognitive Psychology*, *45*, 337–374.

Huttenlocher, J., Vasilyeva, M., Waterfall, H. R., Vevea, J. L., & Hedges, L. V. (2007). The varieties of speech to young children. *Developmental Psychology*, *43*, 1062–1083.

Huttenlocher, J., Waterfall, H., Vasilyeva, M., Vevea, J., & Hedges, L. V. (2010). Sources of variability in children's language growth. *Cognitive Psychology*, *61*, 343–365.

Karrass, J., & Braungart-Rieker, J. M. (2005). Effects of shared parent–infant book reading on early language acquisition. *Journal of Applied Developmental Psychology*, *26*, 133–148.

Kendeou, P., Van den Broek, P., White, M. J., & Lynch, J. S. (2009). Predicting reading comprehension in early elementary school: The independent contributions of oral language and decoding skills. *Journal of Educational Psychology*, *101*, 765–778.

Kidd, E., Brandt, S., Lieven, E., & Tomasello, M. (2007). Object relatives made easy: A cross-linguistic comparison of the constraints influencing young children's processing of relative clauses. *Language and Cognitive Processes*, *22*, 860–897.

King, J., & Just, M. A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language*, *30*, 580–602.

Landau, B., & Gleitman, L. (1985). *Language and experience*. Cambridge, MA: Harvard University Press.

Lieven, E. (2010). Input and first language acquisition: Evaluating the role of frequency. *Lingua*, *120*, 2546–2556.

Lonigan, C. J., Burgess, S. R., & Anthony, J. L. (2000). Development of emergent literacy and early reading skills in preschool children: Evidence from a latent-variable longitudinal study. *Developmental Psychology*, *36*, 596–613.

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum.

Mak, W. M., Vonk, W., & Schriefers, H. (2002). The influence of animacy on relative clause processing. *Journal of Memory and Language*, *47*, 50–68.

Maratsos, M., Fox, D. E., Becker, J. A., & Chalkley, M. A. (1985). Semantic restrictions on children's passives. *Cognition*, *19*, 167–191.

Massaro, D. W. (2015). Two different communication genres and implications for vocabulary development and learning to read. *Journal of Literacy Research*, *47*, 505–527.

Mishra, R. K., Singh, N., Pandey, A., & Huettig, F. (2012). Spoken language-mediated anticipatory eye-movements are modulated by reading ability: Evidence from Indian low and high literates. *Journal of Eye Movement Research*, *5*, 1–10.

Mol, S. E., Bus, A. G., de Jong, M. T., & Smeets, D. J. (2008). Added value of dialogic parent–child book readings: A meta-analysis. *Early Education and Development*, *19*, 7–26.

Montag, J. L., Jones, M. N., & Smith, L. B. (2015). The words children hear picture books and the statistics for language learning. *Psychological Science*, *26*, 1489–1496.

Montag, J. L., & MacDonald, M. C. (2015). Text exposure predicts spoken production of complex sentences in 8-and 12-year-old children and adults. *Journal of Experimental Psychology: General*, *144*, 447–468.

Naigles, L. R. (1996). The use of multiple frames in verb learning via syntactic bootstrapping. *Cognition*, *58*, 221–251.

Naigles, L. R., & Hoff-Ginsberg, E. (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*, 95–120.

Ninio, A., & Bruner, J. (1978). The achievement and antecedents of labeling. *Journal of Child Language*, *5*, 1–15.

Noble, C. H., Cameron-Faulkner, T., & Lieven, E. (2018). Keeping it simple: The grammatical properties of shared book reading. *Journal of Child Language*, *45*, 753–766.

Payne, A. C., Whitehurst, G. J., & Angell, A. L. (1994). The role of home literacy environment in the development of language ability in preschool children from low-income families. *Early Childhood Research Quarterly*, *9*, 427–440.

Payne, B. R., Gao, X., Noh, S. R., Anderson, C. J., & Stine-Morrow, E. A. (2012). The effects of print exposure on sentence processing and memory in older adults: Evidence for efficiency and reserve. *Aging, Neuropsychology, and Cognition*, *19*, 122–149.

Raikes, H., Luze, G., Brooks-Gunn, J., Raikes, H. A., Alexander Pan, B., Tamis-LeMonda, C. S., & Rodriguez, E. T. (2006). Mother–child bookreading in low-income families: Correlates and outcomes during the first three years of life. *Child Development*, *77*, 924–953.

Reali, F., & Christiansen, M. H. (2007). Processing of relative clauses is made easier by frequency of occurrence. *Journal of Memory and Language*, *57*, 1–23.

Roland, D., Dick, F., & Elman, J. L. (2007). Frequency of basic English grammatical structures: A corpus analysis. *Journal of Memory and Language*, *57*, 348–379.

Roth, F. P. (1984). Accelerating language learning in young children. *Journal of Child Language*, *11*, 89–107.

Rowe, M. L. (2012). A longitudinal investigation of the role of quantity and quality of child directed speech in vocabulary development. *Child Development*, *83*, 1762–1774.

Salo, V. C., Rowe, M. L., Leech, K. A., & Cabrera, N. J. (2016). Low-income fathers' speech to toddlers during book reading versus toy play. *Journal of Child Language*, *43*, 1385–1399.

Savage, C., Lieven, E., Theakston, A., & Tomasello, M. (2003). Testing the abstractness of children's linguistic representations: Lexical and structural priming of syntactic constructions in young children. *Developmental Science*, *6*, 557–567.

Scarborough, H. S., Dobrich, W., & Hager, M. (1991). Preschool literacy experience and later reading achievement. *Journal of Learning Disabilities*, *24*, 508–511.

Sénéchal, M., & LeFevre, J. A. (2002). Parental involvement in the development of children's reading skill: A five-year longitudinal study. *Child Development*, *73*, 445–460.

Shahaeian, A., Wang, C., Tucker-Drob, E., Geiger, V., Bus, A. G., & Harrison, L. J. (2018). Early shared reading, socioeconomic status, and children's cognitive and school competencies: Six years of longitudinal evidence. *Scientific Studies of Reading*, *22*, 485–502.

Sharif, I., Rieber, S., & Ozuah, P. O. (2002). Exposure to reach out and read and vocabulary outcomes in inner city preschoolers. *Journal of the National Medical Association*, *94*, 171–177.

Shneidman, L. A., Arroyo, M. E., Levine, S. C., & Goldin-Meadow, S. (2013). What counts as effective input for word learning? *Journal of Child Language*, *40*, 672–686.

Snow, C. E., & Ninio, A. (1986). The contracts of literacy: What children learn from learning to read books. In W. H. Teale & E. Sulzby (Eds.), *Emergent literacy: Writing and reading* (pp. 116–138). Norwood, NJ: Ablex.

Sosa, A. V. (2016). Association of the type of toy used during play with the quantity and quality of parent-infant communication. *JAMA Pediatrics*, *170*, 132–137.

Street, J. A., & Dabrowska, E. (2010). More individual differences in language attainment: How much do adult native speakers of English know about passives and quantifiers? *Lingua*, *120*, 2080–2094.

Tanaka, M. N., Branigan, H. P., McLean, J. F., & Pickering, M. J. (2011). Conceptual influences on word order and voice in sentence production: Evidence from Japanese. *Journal of Memory and Language*, *65*, 318–330.

Temperley, D. (2003). Ambiguity avoidance in English relative clauses. *Language*, *79*, 464–484.

Traxler, M. J., Morris, R. K., & Seely, R. E. (2002). Processing subject and object relative clauses: Evidence from eye movements. *Journal of Memory and Language*, *47*, 69–90.

Vasilyeva, M., Huttenlocher, J., & Waterfall, H. (2006). Effects of language intervention on syntactic skill levels in preschoolers. *Developmental Psychology*, *42*, 164–174.

Warren, T., & Gibson, E. (2002). The influence of referential processing on sentence complexity. *Cognition*, *85*, 79–112.

Weisleder, A., & Fernald, A. (2014). Talking to children matters early language experience strengthens processing and builds vocabulary. *Psychological Science*, *24*, 2143–2152.

Weizman, Z. O., & Snow, C. E. (2001). Lexical output as related to children's vocabulary acquisition: Effects of sophisticated exposure and support for meaning. *Developmental Psychology*, *37*, 265–279.

Wells, J. B., Christiansen, M. H., Race, D. S., Acheson, D. J., & MacDonald, M. C. (2009). Experience and sentence processing: Statistical learning and relative clause comprehension. *Cognitive Psychology*, *58C*, 250–271.

Whitehurst, G. J., Falco, F. L., Lonigan, C. J., Fischel, J. E., DeBaryshe, B. D., Valdez-Menchaca, M. C., & Caulfield, M. (1988). Accelerating language development through picture book reading. *Developmental Psychology*, *24*, 552–559.

Yarosz, D. J., & Barnett, W. S. (2001). Who reads to young children? Identifying predictors of family reading activities. *Reading Psychology*, *22*, 67–81.

Young, K. T., Davis, K., Schoen, C., & Parker, S. (1998). Listening to parents: A national survey of parents with young children. *Archives of Pediatrics & Adolescent Medicine*, *152*, 255–262.

Yuan, S., & Fisher, C. (2009). 'Really? She blicked the baby?' Two-year-olds learn combinatorial facts about verbs by listening. *Psychological Science*, *20*, 619–626.

Zuckerman, B. (2009). Promoting early literacy in pediatric practice: Twenty years of reach out and read. *Pediatrics*, *124*, 1660–1665.

Appendix A. Percent of all object relative clauses and raw counts (in parentheses) broken down by embedded noun type and presence/absence of a relative pronoun in child-directed speech.

	With relative pronoun			Without relative pronoun			Grand total
	Animate	Inanimate	Total	Animate	Inanimate	Total	
Full NP	0% (0)	0.9% (2)	0.9% (2)	0% (0)	0.4% (1)	0.4% (1)	1.3% (3)
Proper noun	0% (0)	1.7% (4)	1.8% (4)	0% (0)	6.6% (15)	6.6% (15)	8.3% (19)
Pronoun	0% (0)	7.9% (18)	7.9% (18)	4.4% (10)	78.1% (178)	82.5% (189)	90.4% (206)
Grand total	0% (0)	10.5% (24)	10.5% (24)	4.4% (10)	85.1% (194)	89.5% (205)	100% (228)

Percent of all object relative clauses and raw counts (in parentheses) broken down by embedded noun type and presence/absence of a relative pronoun in picture book text.

	With relative pronoun			Without relative pronoun			Grand total
	Animate	Inanimate	Total	Animate	Inanimate	Total	
Full NP	0% (0)	0.9% (1)	0.9% (1)	0% (0)	5.1% (6)	5.1% (6)	6.0% (7)
Proper noun	0% (0)	4.3% (5)	4.3% (5)	1.7% (2)	3.4% (4)	5.1% (6)	9.4% (11)
Pronoun	0% (0)	8.5% (10)	8.5% (10)	5.1% (6)	70.9% (83)	76.1% (89)	84.6% (99)
Grand total	0% (0)	13.7% (16)	13.7% (16)	6.8% (8)	79.5% (93)	86.3% (101)	100% (117)

Appendix B. Percent of all subject, oblique, and passive relative clauses, and raw counts (in parentheses) broken down by presence/absence of a relative pronoun in child-directed speech.

	With relative pronoun			Without relative pronoun			Grand total
	Animate	Inanimate	Total	Animate	Inanimate	Total	
SRC (Intrans)	29.3% (25)	65.9% (54)	95.1% (78)	2.4% (2)	2.4% (2)	4.9% (7)	100% (82)
SRC (Trans)	80% (32)	17.5% (7)	97.5% (39)	2.5% (1)	0% (0)	2.5% (1)	100% (40)
OblRC	0% (0)	57.7% (15)	57.7% (15)	0% (0)	42.3% (11)	42.3% (9)	100% (26)
PassRC	0% (0)	0% (0)	0% (0)	0% (0)	100% (2)	100% (2)	100% (2)

Percent of all subject, oblique, and passive relative clauses, and raw counts (in parentheses) broken down by presence/absence of a relative pronoun in picture book text.

	With relative pronoun			Without relative pronoun			Grand Total
	Animate	Inanimate	Total	Animate	Inanimate	Total	
SRC (Intrans)	44.3% (27)	42.6% (26)	86.9% (53)	9.8% (6)	3.3% (2)	13.1% (8)	100% (61)
SRC (Trans)	52.0% (13)	48.0% (12)	100% (25)	0% (0)	0% (0)	0% (0)	100% (25)
OblRC	3.0% (1)	75.8% (25)	78.8% (26)	0% (0)	21.2% (7)	21.2% (7)	100% (33)
PassRC	0% (0)	44.4% (4)	44.4% (4)	0% (0)	55.6% (5)	55.6% (5)	100% (9)

Appendix C. Percent of all main clause passives and passive relative clauses, and raw counts (in parentheses) broken down by presence/absence of a *by*-phrase and *get* vs. *be* passivization in child-directed speech.

	Passive		
	By-phrase	No *by*-phrase	Total
Get	0% (0)	35.9% (23)	35.9% (23)
Be	12.5% (8)	45.3% (29)	57.8% (37)
None/Other	0% (0)	6.3% (4)	6.25% (4)
Total	12.5% (8)	87.5% (56)	100% (64)
	Passive relative clause		
	By-phrase	No *by*-phrase	Total
Get	0% (0)	0% (0)	0% (0)
Be	0% (0)	0% (0)	0% (0)
None/Other	0% (0)	100% (2)	100% (2)
Total	0% (0)	100% (2)	100% (2)

Percent of all main clause passives and passive relative clauses and raw counts (in parentheses) broken down by presence/absence of a *by*-phrase and get vs. *be* passivization in picture book text.

	Passive		
	By-phrase	No *by*-phrase	Total
Get	2.0% (2)	9.2% (9)	11.2% (11)
Be	5.1% (5)	7.6% (74)	80.6% (79)
None/Other	0% (0)	8.2% (8)	8.2% (8)
Total	7.1% (7)	92.9% (91)	100% (98)
	Passive relative clause		
	By-phrase	No *by*-phrase	Total
Get	0% (0)	0% (0)	0% (0)
Be	0% (0)	44.4% (4)	44.4% (4)
None/Other	0% (0)	55.6% (5)	55.6% (5)
Total	0% (0)	100% (9)	100% (9)