

Two Different Communication Genres and Implications for Vocabulary Development and Learning to Read

Journal of Literacy Research
1–23
© The Author(s) 2016
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/1086296X15627528
jlr.sagepub.com


Dominic W. Massaro¹

Abstract

This study examined potential differences in vocabulary found in picture books and adult's speech to children and to other adults. Using a small sample of various sources of speech and print, Hayes observed that print had a more extensive vocabulary than speech. The current analyses of two different spoken language databases and an assembled picture book corpus replicated and extended these findings. The vocabulary in picture books was more extensive than that found in child-directed speech (CDS) and even adult-directed speech (ADS). The likelihood of observing a rare word not contained in the most common 5,000 words in English was more likely in a corpus of picture books than in two different corpora of CDS. The likelihood of a rare word in the picture books was even greater than that found in ADS. It is proposed that these differences are more indicative of informal versus formal language rather than the spoken versus written modalities per se. Nonetheless, these results highlight the value of rich read-aloud experiences for vocabulary development and potentially for reading comprehension once written language is acquired. These findings are described in terms of a distinction between formal and informal language, which has implications for views of literacy, cognitive and linguistic development, and learning to read.

Keywords

child-directed speech, picture books, reading, print, speech, text analysis, vocabulary, reading comprehension, genre of language use, literacy

¹University of California, Santa Cruz, USA

Corresponding Author:

Dominic W. Massaro, Department of Psychology, Social Sciences 2, University of California, Santa Cruz, Santa Cruz, CA 95064, USA.
Email: massaro@ucsc.edu

Most educators, regardless of theoretical orientation, recommend reading books to children to nourish language development (Dickinson, Griffith, Golinkoff, & Hirsh-Pasek, 2012; Duursma, Augustyn, & Zuckerman, 2008; Sénéchal & LeFevre, 2014; Wells, 1985). Most child care and preschools have reading time devoted to a teacher reading aloud picture books to children. It is not always obvious, however, why the recommendation of this type of shared reading is so ubiquitous. There might be many good reasons but the goal of this article is to assess a specific one. The research determines whether popular picture books read to children extend their linguistic and cognitive experience beyond what usually occurs in day-to-day spoken language exchanges. As will be addressed in the “Discussion” section, it is important to acknowledge that children might also acquire the same linguistic and cognitive experience from oral speech alone depending on genre of the language experience.

Picture Books

Picture books are usually described in terms of possessing two forms of communication: iconic and symbolic (Levin, Anglin, & Carney, 1987; Nikolajeva & Scott, 2001). Pictures are iconic because they resemble objects or scenes depicted, and the text is symbolic because it narrates a linguistic description related to the pictures. There is an art to composing picture books beyond the scope of this contribution. At the minimum, successful picture books require a seamless link between pictures and narrative text. Not surprisingly, children being read to look at pictures and not text (van Kleeck, 2003). Evans and Saint-Aubin (2005) measured eye fixations of 4-year-old children as they were read to. Results revealed that the children spent about 95% more time looking at pictures than text. Thus, during shared picture book reading, the child is actually listening to the story while looking at pictures in the book. Thus, although children are not obtaining direct experience with written letters and words, they are benefiting from the language of the story.

Vocabulary Knowledge

Words and language provide a “cognitive niche,” which expands a person’s cognitive capacities (Clark, 2013), and the earlier children are equipped with this linguistic savvy, the more they will benefit from everyday experiences. Few people would argue with the importance of language in achieving success in society today (Ritchie & Bates, 2013). A direct measure of language quality is the amount and nature of the vocabulary it contains. Research has shown that a child’s vocabulary is fundamental to his or her perceptual and cognitive development (LaSasso & Davey, 1987; Massaro, 2006a, 2006b; Waxman, 2002).

Vocabulary knowledge is positively correlated with both listening and reading comprehension (Anderson & Freebody, 1981; National Reading Panel, National Institute of Child Health and Human Development, 2000). It is not only critical for language competence, but it also plays an important role in understanding the world. Very young children more easily form conceptual categories when category labels are

available than when they are not (Waxman, 2002). Even children experiencing language delays because of specific language impairment benefit once this level of word knowledge is obtained (Massaro, 2006b). Given this evidence,

. . . it follows that increasing the pervasiveness and effectiveness of vocabulary learning offers a timely opportunity for improving conceptual knowledge and language competence for all individuals, whether or not they are disadvantaged because of sensory limitations, learning disabilities, or social condition. (Massaro, 2006a, p. 228)

The present research reveals a vocabulary advantage of picture books over spoken language directed to children and therefore the possibility of picture books accelerating and enhancing vocabulary development.

The widely cited research by Hart and Risley (1995) discovered that before the age of 3, children heard many more words in high than in low socioeconomic status (SES) families. In an important follow-up study, Hoff (2003) demonstrated that “quantity, lexical richness, and sentence complexity of mother’s speech to their children” of the language also differed between these two types of families. Thus, it appears that quality as well as quantity of language is important in children’s language development. Analogously, I make the argument that reading picture books to children provides a similar advantage over the usual speech parents and others direct at children.

Although researchers and educators agree that children should acquire strong vocabularies and use them often, there is less consideration of when and how they should acquire them. Vocabulary should be acquired early in life because a word appears to be better learned the earlier a child learns it. As expected, competence in word recognition is also a function of how frequently the word is encountered (Petrova, Gaskell, & Ferrand, 2011). Age of acquisition is positively correlated with frequency of occurrence, however, making it difficult to isolate their effects.

In a large-scale project, age of acquisition and word frequency had very large and roughly equivalent correlations with reaction time and accuracy in a lexical decision task (Kuperman, Stadthagen-Gonzalez, & Brysbaert, 2012; Kuperman & Van Dyke, 2013). In addition, age-of-acquisition ratings accounted for a substantial percentage of variance after effects of log word frequency, word length, and similarity to other words were accounted for (Kuperman et al., 2012). In another informative study, there was an advantage of high-frequency over low-frequency words acquired fairly late, but early acquisition of words eliminated any frequency advantage (Stadthagen-Gonzalez, Bowers, & Damian, 2004). Thus, experiencing broader vocabulary early in life is beneficial for conceptual and linguistic learning and performance.

Research Goals

Nature of input is important for vocabulary and language development. A child is exposed to a variety of inputs, including exchanges with caregivers, television and related media on mobile devices, and shared picture book reading in which a book is read aloud to the child. The goal of this research is to assess the nature of language the

child experiences in most spoken and written contexts. Significant language and vocabulary differences between these two modalities would be informative for both research inquiry and development of educational applications.

A layperson might think of written language as simply spoken language written down, especially with increasing popularity of automated dictation programs and instant text messaging in which speech is accurately converted to text (e.g., Dragon Dictation, 2015). Similarly, the seamless conversion of written text to speech blurs the distinction between these two modalities. Upon analysis, however, important differences between spoken conversations and written media have been found.

Donald P. Hayes's Study

Hayes (1988) took on at that time a momentous task of collecting and analyzing language occurrences from printed magazines, newspapers, and books, television and natural conversations. Given these samples of speech and writing, Hayes analyzed the variety of words that occurred. Hayes (1988) analyzed what he called the *lexical pitch* of the vocabulary used to describe word usage as in "The speaker was able to pitch her remarks to the level of her audience." The words occurring in these contexts were evaluated against frequency of occurrence in the *American Heritage Dictionary* study (AHD; Carroll, Davis, & Richman, 1971). The AHD was at that time the most comprehensive digital database of written language, with a corpus of more than 5 million tokens.

The question Hayes addressed was the degree to which word tokens found in his specific set of samples overlapped with the most common 10,000 word types in the AHD. (When language occurrences are analyzed, a distinction is made between word types and word tokens. A type is a unique word category, whereas a token indicates an occurrence of a particular type. A count of word types gives the number of unique words that occur, whereas a count of word tokens gives the total number of word occurrences.) As might be expected, the results showed a direct correspondence between frequency of word occurrences in magazines, newspapers, and books and rank order frequency in the AHD, derived from written language. Occurrence of words in samples of these three forms of written media roughly replicated the frequency of occurrence of written language from the much larger AHD. Numerically, about 90% of word occurrences in the samples of newspapers and magazines overlapped with the 10,000 most frequent word types in the AHD.

In contrast to the written language samples, however, the spoken language databases showed a significantly different pattern of results. The most frequent 10,000 word types from the AHD corpus occurred much more often in the spoken language samples than in the written language samples. This was particularly true for words ranked between 300 and 3,000 in the AHD. The 1,000 highest frequency words made up about 85% of the spoken language samples and only about 60% of the written language samples.

These results make the salient point that word occurrences in spoken language appear to differ from those in written language, with a bias for occurrence of high-frequency words in spoken language. Hayes concluded that lexical pitch (richness) of spoken

language is aimed significantly lower than lexical pitch of written language (see also Halliday, 1987). Another way to view the results is that spoken language, with its heavy use of high-frequency words, tends to preclude occurrence of novel, less frequent words.

Current Replication Study

Hayes's seminal work sets the stage for additional analyses with larger and more accessible databases. To address language use and vocabulary acquisition, contrasting picture books with conversational speech is a valid inquiry into their properties and differences. Hayes (1988) found that lexical pitch of his sample of books to be read to preschool children was somewhat intermediate between the spoken and written language samples. However, Hayes and Ahrens (1988, Table 2) reported roughly equivalent results for these books and college graduates speaking to spouses and friends. This difference is most likely due to Hayes's use of 10,000 versus Hayes and Ahrens's use of 5,000 most frequent word types in the AHD as the normative baseline. To address this apparent discrepancy and the question of the richness of the two genres of language more thoroughly, I compared word occurrences in picture books, child-directed speech (CDS), and adult-directed speech (ADS).

An extension of the Hayes study is also called for given that his analyses were based on very small samples, especially given today's availability of databases of spoken and written language. The different categories in his spoken language corpora consisted of only about 1,000 words each, made up of 10 subsamples of 100 or more words. Only 21 books were recorded for the sample of books to be read to preschool children. In the present constructive or conceptual replication (Makel & Plucker, 2014; Schmidt, 2009) and analysis, the written language sample consisted of the complete text of 112 picture books for a total of more than 57,000 words. There were two spoken language databases used in the present study. The first had more than 15,000 words for CDS and more than 65,000 words for ADS. The second had more than 2.5 million words of parental speech in the company of children.

Method

Databases

Picture book database. The picture book database was obtained from a shared picture book reading application called Read With Me! (2012). The text from 112 popular picture books was transcribed for the application, although the text is hidden and not accessible by the user. The choice of picture books involved selections by the author, three experienced librarians working in the children's section of two regional libraries, and several undergraduates. The goal was to choose their favorite picture books as well as those frequently read. As can be seen in the appendix, many of the picture books are well known and consist mostly of "storybooks" rather than "information books." Thus, any advantage in the vocabulary of picture books would be expected to be even larger if "information books" were used. The genre of these books, however,

makes a reasonable comparison with the speech corpora, which were collected in an informal conversational setting.

CDS. For the first corpus of CDS, I analyzed a database collected by Dr. Patricia K. Kuhl at the Institute for Learning and Brain Sciences (formerly the Center for Mind, Brain and Learning) at the University of Washington (Kirchhoff & Schimmel, 2005). The corpus was obtained to evaluate possible differences between acoustic characteristics of adult speech directed to children and to other adults. It consists of 64 conversations that included 32 different mothers solicited from neighboring communities. Each mother had two conversations, one with an adult experimenter and the other with her infant (age 2-5 months). Conversations were recorded using a far-field microphone (suspended from the ceiling near speakers' heads). The recorded speech thus reflects natural conversational conditions. The recorded speech was time segmented and transcribed orthographically by phonetically experienced transcribers at the Department of Electrical Engineering at the University of Washington. These transcriptions were double-checked and serve as the corpus for the present analysis.

The CDS was recorded while the mother was interacting and playing with her child in a playroom situation with a variety of toys. The mother was instructed to simply play with her child as she would normally in a 20- to 30-min play session. (The child's speech was not transcribed.) This produced an average of 520 CDS words from each of the 32 mothers. The topics of conversation were very broad, including discussions of animals, body parts, cartoons, colors, cooking, clothes, eating, emotions, family members, family activities, moods, places to visit, shapes, toys, and similar subjects.

ADS. The ADS came from the same collection as the CDS and consisted of both the mother's and experimenter's speech during a casual conversation. Much of the speech for the ADS corpus included discussions about the mother's child and their various day-to-day activities. Similar to the CDS, the ADS conversation was open-ended although prompts were given by the experimenter when necessary to continue the conversation. The ADS therefore came from both mother and experimenter in each of the 32 sessions. These sessions produced an average of 2,182 words per session. The topics of conversation were similar to those observed in the CDS sessions. The data collection was also designed to elicit nine cue words, containing the most distant points in the vowel triangle, either by toys being played within the CDS session or prompts by the experimenter in the ADS session. The words were key, sheep, bead, boot, shoe, spoon, pot, top, and sock.

Corpus of Contemporary American English (COCA). The COCA (1990-2012) was used to obtain a measure of common and rare words. This corpus is evenly balanced between spoken and written English. The corpus contains more than 450 million words from unscripted radio and TV shows, books of fiction, short stories, and movie scripts, and more than 100 popular magazines, 10 newspapers, and 100 academic journals. More than 150,000 samples contribute to the complete database.

Child Language Data Exchange System (CHILDES) Corpora. A subset of the CHILDES corpora (MacWhinney, 2000) was also used in the analyses. This CHILDES Parental Corpus database consists of contributions of 27 individual corpora provided by 27 different investigators (ParentFreq, 2015). The total number of lexical items in this parental corpus is 2,579,966 word tokens, with 24,156 word types (counting all inflected forms of a word as separate types). It consists of spoken utterances from parents, caregivers, and experimenters in the presence of children (age = 0;7-7;5; *M* age = 36 months). This corpus provides a huge sample of speech that children are exposed to (e.g., dinner table talk, talk during free play, and storytelling), even though not all of the utterances are strictly child-directed.

Procedures/Analyses

Using Unix utility programs `sort` (2015) and `uniq` (2015), I analyzed transcriptions of the spoken language corpora and the text in the picture books to count total number of words (tokens) and number of unique words (types). Given that the size of the databases differed, I took a subset of the picture books to create a sample to give a comparable contrast between the picture books and the CDS corpora. To increase reliability of the analyses, I repeated the sampling of words from two additional subsets of books in the picture book database. The three subsets were created by randomly sampling from the books making up the complete picture book database. Two samples had 32 books, and the remaining sample had 38 books to create samples comparable with the CDS database. All 112 books were used in the comparison with the ADS database. The appendix lists titles of the 112 books and an indication of which books were contained in each of the three samples.

Results

Picture Books Versus CDS

The current databases were analyzed to determine whether picture book and spoken language vocabulary differed significantly. Table 1 lists the number of types and tokens for the CDS database and three samples of the picture book database along with number of unique words in each database. As seen in the column of the Number of Types in Table 1, there were significantly more word types in the three samples of the picture book database (average = 2,504) than in the CDS database (1,017), even though their token counts were fairly equal. This result means that the sample of picture books had a larger vocabulary than found in the sample of CDS.

The type/token ratio gives a measure of how often word types occur relative to the total number of word tokens. A larger ratio means that the database has a larger number of different words relative to the total number of word occurrences. Table 1 shows that the type/token ratios for the three samples of picture books were 0.147, 0.156, and 0.162, respectively, much larger than the CDS corpus ratio of 0.061. Thus, there are about 2½ times as many word types in the picture book corpus relative to the CDS corpus. This initial result, however, is not sufficient to demonstrate

Table 1. Number of Tokens, Number of Types, Type/Token Ratio, Number of Rare Types, and Ratio Advantage of Books Over the CDS Corpus.

Corpus	Number of tokens	Number of types	Type/token ratio ^a	Number of rare types ^b	Ratio advantage of books over CDS ^c
CDS	16,659	1,017	0.061	451	—
Book Sample A	16,431	2,421	0.147	1,342	2.98
Book Sample B	15,473	2,410	0.156	1,367	3.03
Book Sample C	16,556	2,681	0.162	1,650	3.66
Average book samples	16,153	2,504	0.155	1,453	3.22

Note. CDS = child-directed speech; COCA = Corpus of Contemporary American English.

^aThe type/token ratio is defined as the number of word types divided by the number of word tokens.

^bThe number of rare types is defined as the number of words that occur in the corpus but are not among the 5,000 most frequent words in the COCA (1990-2012) database.

^cThe ratio advantage of books over CDS is defined as the number of rare types for books divided by the number of rare types for CDS.

that picture books have a more widespread and challenging vocabulary than CDS. It might be the case, for example, that the picture books might simply contain more *common* word types than those in the CDS sample. Following the procedure used by Hayes (1988), it is necessary to determine how vocabulary in the current samples compares with the most frequently used words in English. Given Hayes's findings, picture books should have a more extensive vocabulary over and above frequently used words.

The picture book, CDS, and ADS samples were evaluated with respect to the 5,000 most frequently used spoken and written words in the COCA (1990-2012). The COCA list of the 5,000 most frequent words distinguishes among identically spelled words corresponding to different parts of speech. Given that part of speech is not available in the picture book, CDS, and ADS databases, it had to be ignored in the analysis. Combining words that had the same spelling in the COCA list reduced it from 5,000 to 4,353 uniquely spelled words. Given the enormous COCA database, this list provides a robust measure of the most *common* or frequently used words in the English language. In the present analysis, a *rare* word is defined as one not occurring in the list of common words.

The column of the Number of Rare Types in Table 1 eliminated words in the samples that overlapped with the most frequent words in the COCA database. This column shows that there were roughly 3 times as many rare word types (average 1,453 vs. 451) in the picture book word corpus than in the CDS corpus. This result means that children listening to a reading of a picture book are roughly 3 times more likely to experience a new word type that is not among the most frequent words relative to the situation of listening to their caregiver's speech. Thus, children experience a greater number of rare words and supposedly a more extensive and challenging vocabulary with picture books than with CDS.

Table 2. Number of Tokens, Number of Types, Type/Token Ratio, Number of Rare Types, and Ratio Advantage of Books Over the ADS Corpus.

Corpus	Number of tokens	Number of types	Type/token ratio ^a	Number of rare types ^b	Ratio advantage of books over ADS ^c
ADS	69,838	3,866	0.055	2,227	
All picture books	57,311	5,362	0.094	3,672	1.65

Note. ADS = adult-directed speech; COCA = Corpus of Contemporary American English.

^aThe type/token ratio is defined as the number of word types divided by the number of word tokens.

^bThe number of rare types is defined as the number of words that occur in the corpus but are not among the 5,000 most frequent words in the COCA (1990-2012) database.

^cThe ratio advantage of books over CDS is defined as the number of rare types for books divided by the number of rare types for CDS.

Picture Books Versus ADS

In the next analysis, vocabulary in the picture books was contrasted with ADS. Given that the number of words in the ADS database was similar to the complete picture book database, the full corpus of picture books was compared with ADS. Results in Table 2 show that picture books have a more extensive vocabulary than that observed when two adults are talking to each other. The column of the Number of Rare Types in Table 2 shows that there were roughly 1 and 1½ times as many rare word types (3,672 vs. 2,227) in the picture book word corpus than in the CDS corpus. Thus, the advantage of picture books over spoken language was about half the size for ADS than for CDS. Thus, the analysis shown in Tables 1 and 2 confirms the conclusion by Hayes (1988) that lexical pitch of written language is aimed higher than lexical pitch of the spoken language of conversations. Given that text complexity is correlated with word difficulty, it is reasonable to expect that a more challenging text exists when there is a larger number of word types not found in the most frequent words in the language.

CHILDES Parental Corpus

It might be argued that the current CDS speech samples might not be representative of spoken language to children in general. To address this issue, I determined whether my conclusions would hold up with an extremely large database of CDS by analyzing a subset of the CHILDES corpora (MacWhinney, 2000). To test whether the picture books have a more widespread and challenging vocabulary than the CHILDES parental corpus, I took three independent samples from the CHILDES corpus and compared these with the current picture book sample. Following the previous analyses, it is necessary to determine how vocabulary in these samples compares with the 5,000 most frequently used words (COCA, 1990-2012). The column of the Number of Rare Types in Table 3 shows that there were roughly 1 and 1½ times as many rare word types (3,672 vs. 2,118) in the picture book word corpus than in the CHILDES corpus. The

Table 3. Number of Tokens, Number of Types, Type/Token Ratio, Number of Rare Types, and Ratio Advantage of Picture Books Over Three Samples of the CHILDES Parental Speech Corpus.

Corpus	Number of tokens	Number of types	Type/token ratio ^a	Number of rare types ^b	Ratio advantage of picture books over CHILDES ^c
CHILDES Sample 1	57,311	3,499	0.061	2,120	1.73
CHILDES Sample 2	57,311	3,523	0.061	2,149	1.70
CHILDES Sample 3	57,311	3,467	0.060	2,085	1.76
CHILDES Average	57,311	3,496	0.061	2,118	1.73
All picture books	57,311	5,362	0.094	3,672	

Note. CHILDES = Child Language Data Exchange System; COCA = Corpus of Contemporary American English.

^aThe type/token ratio is defined as the number of word types divided by the number of word tokens.

^bThe number of rare types is defined as the number of words that occur in the corpus but are not among the 5,000 most frequent words in the COCA (1990-2012) database.

^cThe ratio advantage of picture books over CHILDES is defined as the number of rare types for books divided by the number of rare types for CHILDES.

results show an advantage of picture books over the CHILDES parental corpus. Eliminating words in the samples that overlap with the most frequent words in the COCA database, the Ratio Advantage column shows that there were roughly 1.73 times as many rare word types in the picture book word corpus than in the CHILDES corpus. There are three likely explanations why this result shows a smaller advantage than the comparison in the first CDS corpus. First, the children were younger in the first CDS corpus, and second, the CHILDES corpus also contained parental speech not directed at children. However, Hayes and Ahrens (1988, Figure 1) found that the parental speech to children did not differ across their first 12 years. A third possibility is that CHILDES corpus is highly correlated with ADS. Evidence for the third possibility is that the CHILDES corpus and the ADS corpus gave essentially identical results (see Tables 2 and 3 of the present article).

To provide another test of the hypothesis that the type/token ratio is larger in print than in spoken language to children, the CHILDES parental spoken corpus (MacWhinney, 2000) was contrasted with the AHD print corpus (Carroll et al., 1971). (The AHD corpus is the appropriate comparison because the larger COCA corpus contains both spoken and written languages.) The type/token ratio in the CHILDES parental corpus is 0.0093 compared with a type/token ratio of 0.017 in the AHD print corpus. Assuming that the number of word types will remain fairly constant with increases in sample size from around 2.5 million to 5 million, a new word type would occur about 3.6 times more often in print than in spoken language to children (the ratio

of 86,741/24,156 word types). This result based on much larger databases replicates the differences found in Tables 1 to 3 with the derived CDS, ADS, CHILDES, and picture book corpora.

Another measure of vocabulary size in picture books relative to CDS is to determine number of word types that occurred in the picture book database that did not occur in the CHILDES parental corpus. There were 1,274 words that occurred in the picture book database that did not occur in the CHILDES parental corpus. This is an impressive result given that there were 2,579,966 words in the CHILDES parental corpus and only 57,311 words in the picture book database, a factor of 45 in corpus size. It is also useful to compare these results with those from ADS. In this case, there were only 541 word types that occurred in the ADS database that did not occur in the CHILDES parental corpus. This result is much lower than the 1,274 words that occurred in the picture book database that did not occur in the CHILDES parental corpus. It is also consistent with the result that the likelihood of a unique word was significantly larger in picture books than in ADS (see Table 2).

In summary, several measures from several databases indicate that picture books are generally composed of less frequent and more complex vocabulary relative to the speech of caregivers to children.

Discussion

In the present research, the vocabulary found in three corpora of spoken language and an assembled picture book corpus was evaluated. Rare vocabulary words were more characteristic of print in picture books than in spoken language. There was a larger likelihood of observing a rare word (not one of the most common words) in picture books than in spoken language. These results extend and refine the earlier finding of Hayes (1988) who was limited by much smaller samples of speech and print.

An analysis of words that occurred in the picture books that did not occur in speech indicated that the child is exposed to challenging words in picture books that are not commonly found in speech. It has also been shown that the earlier the age of acquisition of words, the better their memory and processing in adulthood (Kuperman et al., 2012). Thus, with picture books, the child benefits not only from being exposed to infrequent and challenging words but also from being exposed to them early in life.

Differences Between Speech and Writing

It is not surprising that the linguistic content of speech and writing differs from one another because usually their goals differ. Most speech occurs in contexts in which interlocutors are cooperating with one another to construct meaningful conversation. Spoken dialogues usually require rapid exchanges, and the interlocutors find high-frequency words more accessible. Some of Paul Grice's (1975) maxims for good conversations help highlight how spoken word choice is somewhat constrained. The interlocutor's contributions should be efficient, informative, relevant, and unambiguous. Compared with

Table 4. Twelve Most Frequently Occurring Words and Their Proportion of Occurrence in the Google Books Database, Which Consisted of 97,565 Distinct Words That Occurred a Total of 743,842,922,321 Times (After Norvik, 2012).

Word	% occurrence ^a
The	7.14
Of	4.16
And	3.04
To	2.60
In	2.27
Of	2.06
Is	1.13
That	1.08
For	0.88
It	0.77
As	0.77
Was	0.74

^aThe percentage of occurrence of each of the words corresponds to the percentage of times that word occurred in the Google Books database.

traditional writing, which requires rewriting and even external editing, speaking must be initiated fairly rapidly and is in most cases off the cuff.

Evidence for the different requirements of spoken and written language can also be seen in the occurrence of the dozen most frequent ranked words in English. These occurred more often in written language than in spoken language. Table 4, showing the 12 most frequent words in the Google Books database (Norvik, 2012), reveals that these words are mostly articles, propositions, and conjunctions. These words occur more often in written than spoken language, most likely because they are currently considered more essential for written language than for spoken language, whereas these words can be dropped in casual speech without harm. In contrast to most written language, spoken language is usually “face-to-face” so that gestures and context can substitute for these words. These differences between use of spoken and written language indicate why the most frequent dozen or so words in the language will occur more often in print than in speech.

In contrast to the more frequent use of the most common articles, propositions, and conjunctions in written language, common nouns and verbs tend to be used more often in spoken language. One reason is that speech production in dialogues is partially constrained by how quickly conversationalists can access words. It is well documented that time required for lexical access is dependent on both familiarity and age of acquisition (Kuperman et al., 2012; Kuperman & Van Dyke, 2013). These constraints can bias word choice to more frequent and early-acquired words. The context of conversations also limits word choice; for example, interlocutors tend to repeat words that have been recently uttered and use deictic gestures in place of content words. All of these

constraints bias spoken conversations to use more common words than would be found in written language that does not normally have these constraints.

Genre of Language Use

The present comparisons between written and spoken language were necessarily confounded with genre of the linguistic content. The language in picture books is qualitatively different from most speech that parents address to their children, but the difference is not because it is written versus spoken language. In fact, children are actually hearing speech when books are read to them. Most importantly, well-scripted spoken language can be rich in word choice and grammar, as in prepared talks (Stand Up to Bullying, 2015; Talks to Watch With Kids, 2015). If parents and caregivers give a formal lecture to their children or narrate true or fictional stories, properties of the utterances in these talks should be more similar to those in picture books than to CDS.

The goal of this article was not to argue in favor of one language modality versus another but to simply assess the nature of language that is currently found in picture books versus speech. It is understood that different contexts require varying intentions, purposes, and manners of speaking. Thus, some conversational speech might be more sophisticated than conversation in storybooks (Hayes & Ahrens, 1988, Table 2). However, the current statistical analyses indicate that, as expected, picture books will usually offer more extensive language than conversational speech.

Rather than considering any observed difference between speech and writing as a difference due to sensory/motor modality, a distinction between informal and formal language might be more germane. Traditionally, most speech has been viewed as informal and most writing as formal. However, recent rapid changes in technology and social media are blurring this long-accepted difference between speech and writing. With current editing and recording technologies, the impermanence of speech and the permanence of writing are not any longer a distinguishing factor. The Internet has co-opted differences in time and space as either media can reach across the world (Socialnomics, 2015). With texting and Skype, immediacy can be a feature of either medium. Complexity can vary in either medium with nonfiction books and oral lectures as examples of complex instances. Speech is usually a dynamic interaction between two or more people but there are frequent occurrences of teenagers and even adults texting to each other across a shared table. Even emotional content and paralinguistic cues found in speech can be approximated in writing with emoticons and other embellishments of text.

Given that both spoken and written languages can be naturally used in most situations, it might be reasoned that it is the formal versus informal genre of linguistic occurrence rather than its oral or written medium that is important. Table 5 provides a taxonomy revealing the potential independence of language modality (spoken vs. written) and formal (nonconversational) versus informal (conversational) dialogue. Categorizing linguistic occurrences as formal versus informal provides a framework that gives some predictability of complexity of vocabulary that would be found. More rare and complex vocabulary should be found in the formal relative to the informal

Table 5. A Taxonomy Revealing the Potential Independence of Language Modality (Spoken vs. Written) and Formal (Nonconversational) Versus Informal (Conversational) Dialogue.

	Spoken language	Written language
Formal language	TED (Technology, Entertainment, Design) talk Lecture MOOC (Massive Open Online Course)	Non-Fiction book Scholarly article Newspaper
Informal language	Face-to-face conversation TV dialogue Fiction films	Texting Instant messaging Light fiction writing

instances of language use in Table 5. Although there is currently no direct evidence to substantiate this categorization, future research could sample vocabulary found in these instances. For example, frequent occurrence of the 12 most frequent words from the Google Books Database shown in Table 4 should occur much less in texting and instant messaging. In addition, there should be more rare and complex vocabulary in more formal instances of language use regardless of whether they are delivered via speech or writing.

Implications of the Research

Views of literacy. There has been continuing discussion since Plato's (1925) *Phaedrus* about differences between spoken and written language. Socrates bemoaned loss of the human skill of memory with the advent and use of print. Beginning with Plato, scholars have delineated the many values that print provides. Havelock (1986) proposed a qualitative change in the form of thought beginning in ancient Greece with the transition from orality to literacy. Ong (2002) described writing as a technology that must be laboriously learned, and which effects the first transformation of human thought from the world of sound to the world of sight. Ong sought to identify distinguishing characteristics of thought and its verbal expression in societies with and without literacy in most of the population (see also Olson & Torrance, 1991).

Considering the genre of language being used, however, differences observed between speech and print might be more adequately explained by a distinction between formal and informal language. Scribner and Cole's (1978, 1981) research in Liberia was able to distinguish between school effects and literacy effects. They evaluated three different scripts and literacy traditions: school literacy in English, religious literacy in Arabic script and, third, an indigenous script used by individuals for letter writing and record keeping in the local language. They found that cognition, reasoning, and memory benefited most from school experience not from literacy itself. Although literacy correlates positively with income, wealth, and health at both individual and societal levels (Prinsloo & Baynham, 2013; Ritchie & Bates, 2013), there may be nothing special about written language itself as much as various societal

practices such as schooling. Additional discussion and evidence for the bridging of a divide between oral language and literacy can be found in Prinsloo and Baynham (2013). Given electronic media, for example, reading and writing can be easily combined with speech to communicate with friends and teachers, engage in learning sessions, and create scholarly projects. Shared reading of picture books can be viewed as embodying a component of school practice not found in typical spoken language. If so, this experience not only contributes to children's language but also prepares them for school's didactic practices.

Cognitive and linguistic development. The current analysis of spoken and written language corpora indicates that the picture books contain a more extensive vocabulary than speech. A primary supposition of this article is that the more extensive language found in picture books relative to spoken language directed at children has important consequences for linguistic and cognitive development. It is well documented that word learning is positively correlated with word frequency and age of acquisition. In turn, oral language competence and vocabulary knowledge are positively correlated with learning to read and reading ability (National Early Literacy Panel, 2008; National Reading Panel, National Institute of Child Health and Human Development, 2000). Just recently, results have shown that breadth and depth of vocabulary knowledge as measured by the amount of information known about a word are correlated with the reader's ability to make global coherence inferences from text (Oakhill, Cain, & McCarthy, 2015).

More generally, vocabulary growth is positively correlated with understanding more complex language. Children's understanding of and use of grammar is positively correlated with their vocabulary (Bates, Bretherton, & Snyder, 1988). There is also good evidence that children tend to produce language that they hear (Cameron-Faulkner, Lieven, & Tomasello, 2003). To test the relationship between the words children produce and the words they hear, I correlated the log₁₀ frequency of 2,797 words from the CHILDES parental corpus with the log₁₀ frequency of these words from another CHILDES corpus of children's speech (Baath, 2014). This set of words was chosen because they had both concreteness and imagery ratings from another database (MRC Psycholinguistic Database, 2015). The correlations were extremely large: .812, .900, and .869 for children aged 18, 36, and 60 months, respectively. Thus, young children's language reflects the language they hear from their parents. Given the higher lexical pitch of picture books, children hearing these books read by their parents will promote language use to a greater degree than their experience with most other spoken language.

There has been renewed interest in enriching a growing child's language by increasing the amount of language the child hears (Talbot, 2015). The present research, however, questions the value of simply talking more to children without any concern with what is communicated (Massaro, 2015). A recent project by the City of Providence, called "Providence Talks," is designed to automatically record words being spoken to children with the goal of increasing the amount of language children hear (Providence Talks, 2014). As indicated by research reviewed in this article, language quality is

critical in providing vocabulary necessary for children to develop into good readers and competent learners. Given the fact that word mastery in adulthood is correlated with early acquisition of words, shared picture book reading offers a potentially powerful strategy to prepare children for competent literacy skills.

Learning to read. One controversial line of thinking in the psychology and pedagogy of reading either implicitly or explicitly conceptualizes successful reading as an act in two stages (Gough & Tunmer, 1986). When schooling begins, the child owns an impressive facility in spoken language making it valuable to have reading further build on that facility. To exploit this facility, the child's task is to "decode" printed language into spoken language (Gough, Juel, & Griffith, 1992). Once decoded, the child's ability to comprehend written language is roughly equivalent to their listening comprehension. This traditional conceptualization of learning to read expects the child to associate or bridge written text to spoken language. The present analyses indicate, however, that much of written language is not found in day-to-day spoken language. When children begin reading independently, they will encounter many new words that they have not experienced in spoken language. Thus, spoken language alone does not prepare the child for written language of books. It might be appropriate to rethink the standard model that makes reading and learning to read parasitic on speech.

I believe that learning to read requires two fundamental skills: the mechanics of processing written language (e.g., recognizing written letters and words) and language understanding. Research by Sénéchal and colleagues supports this distinction. They have shown that shared storybook reading predicts vocabulary acquisition, whereas teaching written language predicts alphabet knowledge and even future reading fluency. One study (Sénéchal, 2006) found that parents teaching the alphabet, reading words, and printing words predicted kindergarten alphabet knowledge and Grade 4 reading fluency. Shared storybook reading, however, predicted kindergarten vocabulary and frequency with which children reported reading for pleasure in Grade 4.

Sénéchal and LeFevre (2014) were able to study children who were schooled in French but spoke English at home. This assessment tested how English home literacy activities influenced their progress in English literacy independent of school-based English language instruction. They found that an informal literacy environment (e.g., parents reading to their children at home) predicted growth in English receptive vocabulary from kindergarten to Grade 1, whereas parent reports of a formal literacy environment (consisting of parents teaching the alphabet, printing words, and reading words) predicted growth in children's English early literacy between kindergarten and Grade 1 and growth in English word reading during Grade 1. Other contributing factors to literacy acquisition include parents' didactic activities, their expectations about learning early literacy at home, as well as their child's interest in learning literacy (Martini & Sénéchal, 2012). Although parental teaching activities help guarantee the child's direct experience of written language, it is important that shared storybook reading contributes to growth of the child's receptive vocabulary.

A relatively recent U.S. national research review panel on reading was aimed at children from birth through age 5 (National Early Literacy Panel, 2008). Predictors of

reading acquisition included written language knowledge, cognitive skills, and memory. Five more moderately correlated factors also revealed the importance of written language. Factors included concepts about print and print knowledge, early decoding, reading readiness, oral language, and visual processing. Most interesting for our purposes, however, was the overall correlation of early oral language development with later reading achievement. When oral language was measured with complex measures (such as using multicomponent tests that included knowledge of syntax, ability to define word meanings, and listening comprehension), this variable explained roughly 50% of the variation in later reading comprehension. A possible implication of this finding is that the richer language of storybook reading can provide young children a richer oral language that will make a positive contribution to their future reading skill.

The present research documented that picture books will necessarily have greater lexical pitch than found in most conversational spoken language. An important caveat is whether children will be able to understand and learn challenging vocabulary and concepts of picture books without guidance (Schleppegrell, 2012; Snow, 2002). Children will usually receive appropriate guidance in shared picture book reading. When reading independently, children necessarily benefit when given guidance via print, speech, or other media such as pictures, animations, or videos. There are several interventions that provide this type of support (Glenberg, 2008; Glenberg & Gallese, 2012; Tecwave, 2013), and there should be more of these as technology and reading science evolve.

In conclusion, very little inquiry has addressed the issue of how picture books and conversational spoken language differ from one another. In this article, I show that picture books embody a lexical pitch that would support learning of new and challenging vocabulary. This vocabulary is particularly valuable because an extensive vocabulary is positively correlated with reading comprehension (National Reading Panel, National Institute of Child Health and Human Development, 2000). Most parents are advised that they should read to their children. As far as I know, no one previously has substantiated an important reason why this advice is particularly valid: The books read to children are usually more linguistically and cognitively enriching than everyday speech.

Appendix

Names of the books in the picture book database taken from transcriptions of the books used in Read With Me! (2012). The letters A, B, and C refer to the samples used in the child-directed speech (CDS) comparisons (as shown in Table 1).

17 Kings And 42 Elephants (C)	Amazing Grace (A)
A Place For Ben (C)	Animalia (C)
A Pocket For Corduroy	Are You My Mother? (A)
All The Colors Of The Earth (C)	Arthur Babysits (C)
All The Places To Love (A)	Arthur Writes A Story (A)

Arthur's Tooth (C)	Jumanji
Arthur's Underwear (C)	Kitten's First Full Moon
Arthur's Valentine (A)	Leaf Man
Ask Mr. Bear (A)	Library Lion
Barbapapa Zoo (C)	Little Bear's Little Boat (A)
Bark, George (C)	Little Cloud
Below (C)	Little Gorilla (B)
Boats (C)	Little Pea (B)
Brown Bear, Brown Bear, What Do You See? (C)	Llama Llama Red Pajama (B)
Charlie Needs A New Cloak (C)	Love You Forever (B)
Counting Kisses (C)	Madeline (A,B)
Curious George (C)	Martha Speaks (A,B)
Dear Zoo (C)	Matthew's Dream (A,B)
Doctor De Soto (C)	Mike Mulligan And His Steam Shovel (B)
Don't Frighten The Lion (C)	Milk And Cookies (A,B)
Duck For President (A)	Miss Nelson Is Missing (A,B)
Duck On A Bike (A)	Miss Rumphius (B)
Earth Dance (A)	Mr. Gumpy's Motor Car (B)
Everyone Poops	Oh, The Places You'll Go! (B)
Fancy Nancy And The Posh Puppy (A)	Olivia And The Missing Toy (B)
Five Little Monkeys Jumping On The Bed	Over In The Meadow (B)
Frederick	Owl Moon (B)
Freight Train	Pancakes, Pancakes! (A,B)
George And Martha (A)	Panda Bear, Panda Bear, What Do You See? (B)
George And Martha One Fine Day (A)	Pig On The Titanic (A,B)
George Shrinks	Polar Bear (B)
Good Night Gorilla	Qwerty The Computer (B)
Goodnight Moon	Round Trip (B)
Grabby Pup	Shadow Play (B)
Grandfather's Journey (A)	Shake My Sillies Out (B)
Green Eggs And Ham	Sheep In A Jeep (B)
Guess How Much I Love You	Somewhere In Africa (B)
Harold And The Purple Crayon (A)	Stellaluna (B)
Hattie And The Fox	Swimmy (B)
Have You Filled A Bucket Today? (A)	Tar Beach (A,B)
Hop On Pop	Teeny Tiny (B)
How Do Dinosaurs Say Good Night? (A)	Ten Nine Eight (B)
How I Became A Pirate	Ten Oni Drummers (B)
If You Give A Mouse A Cookie	The Carrot Seed (C)
In My Heart (A)	The Cat In The Hat (C)
Inch By Inch	The Giving Tree (A,C)
Jamberry (A)	The Grandad Tree (C)
	The Little Engine That Could (C)

The Napping House (C)	What Do Ducks Dream? (C)
The Polar Express (C)	What Happened To Patrick's Dinosaurs? (A,C)
The Story Of Ferdinand (C)	Where Are You Going? (C)
The Very Busy Spider (C)	Where Once There Was A Wood (A,C)
The Very Hungry Caterpillar (C)	Where The Forest Meets The Sea (C)
There's A Wocket In My Pocket (C)	Where The Wild Things Are (C)
Time To Sleep (A,C)	Whistle For Willie
Tough Boris (C)	Wombat Walkabout (A,C)
Two Little Trains	

Acknowledgments

The author would like to thank Dr. Patricia Kuhl of I-Labs and Katrin Kirchhoff at the University of Washington for making the child-directed speech available to him. He also thanks the editors and the reviewers for their very helpful comments and guidance on the research.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: A portion of the data entry described in this article was supported by a grant from the Academic Senate of the University of California, Santa Cruz.

References

- Anderson, R. C., & Freebody, P. (1981). Vocabulary knowledge. In J. T. Guthrie (Ed.), *Comprehension and teaching: Research perspectives* (pp. 71-117). Newark, DE: International Reading Association.
- Baath, R. (2014). *ChildFreq: An online tool to explore word frequencies in child language*. Available from <http://childfreq.sumsar.net/>
- Bates, E., Bretherton, I., & Snyder, L. (1988). *From first words to grammar: Individual differences and dissociable mechanisms*. New York, NY: Cambridge University Press.
- Cameron-Faulkner, T., Lieven, E., & Tomasello, T. (2003). A construction based analysis of child directed speech. *Cognitive Science*, 27, 843-873. doi:10.1207/s15516709cog2706_2
- Carroll, J. B., Davis, P., & Richman, H. (1971). *Word frequency book*. Boston, MA: Houghton Mifflin.
- Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral & Brain Sciences*, 36, 181-253. doi:10.1017/S0140525X12000477
- Corpus of Contemporary American English (COCA). (1990-2012). *Words added between 1990 and 2012*. Retrieved from <http://corpus.byu.edu/coca/>
- Dickinson, D., Griffith, J., Golinkoff, R., & Hirsh-Pasek, K. (2012). How reading books fosters language development around the world. *Child Development Research*, 2012, Article 602807. doi:10.1155/2012/602807

- Dragon Dictation. (2015). Retrieved from <http://www.nuance.com/for-individuals/mobile-applications/dragon-dictation/index.htm>
- Duursma, E., Augustyn, M., & Zuckerman, B. (2008). Reading aloud to children: The evidence. *Archives of Disease in Childhood*, *93*, 554-557. Retrieved from <http://adc.bmj.com/content/93/7/554.full.pdf+html>
- Evans, M. A., & Saint-Aubin, J. (2005). What children are looking at during shared story-book reading? Evidence from eye movement monitoring. *Psychonomic Science*, *116*, 913-920.
- Glenberg, A. M. (2008). Toward the integration of bodily states, grounding: Social, cognitive and neuroscientific approaches. In G. R. Semin & E. R. Smith (Eds.), *Embodied grounding: Social, cognitive, affective and neuroscientific approaches* (pp. 43-70). Cambridge, UK: Cambridge University Press.
- Glenberg, A. M., & Gallese, V. (2012). Action-based language: A theory of language acquisition, comprehension, and production. *Cortex*, *48*, 905-922. doi:10.1016/j.cortex.2011.04.010
- Gough, P. B., Juel, C., & Griffith, P. (1992). Reading, spelling and the orthographic cipher. In P. B. Gough, L. C. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 35-48). Hillsdale, NJ: Lawrence Erlbaum.
- Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education*, *7*, 6-10.
- Grice, H. P. (1975). Logic and conversation. In A. P. Martinich (Ed.), *Philosophy of language* (pp. 165-175). New York, NY: Oxford University Press.
- Halliday, M. A. K. (1987). Spoken and written modes of meaning. In R. Horowitz & S. J. Samuels (Eds.), *Comprehending oral and written language* (pp. 55-82). San Diego, CA: Academic Press.
- Hart, B., & Risley, T. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore, CA: Paul H. Brookes.
- Havelock, E. A. (1986). *The muse learns to write: Reflections on orality and literacy from antiquity to the present*. New Haven, CT: Yale University Press.
- Hayes, D. P. (1988). Speaking and writing: Distinct patterns of word choice. *Journal of Memory and Language*, *27*, 572-585.
- Hayes, D. P., & Ahrens, M. (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 language development via maternal speech. *Child Development*, *74*, 1368-1378. doi:10.1111/1467-8624.00612
- Kirchhoff, K., & Schimmel, S. (2005). Statistical properties of infant-directed vs. adult-directed speech: Insights from speech recognition. *Journal of the Acoustical Society of America*, *117*, 2224-2237.
- Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30 thousand English words. *Behavior Research Methods*, *44*, 978-990.
- Kuperman, V., & Van Dyke, J. A. (2013). Reassessing word frequency as a determinant of word recognition for skilled and unskilled readers. *Journal of Experimental Psychology: Human Perception and Performance*, *39*, 802-823.
- LaSasso, C., & Davey, B. (1987). The relationship between lexical knowledge and reading comprehension for prelingually, profoundly hearing impaired students. *The Volta Review*, *89*, 211-220.

- Levin, J. R., Anglin, G. J., & Carney, R. N. (1987). On empirically validating functions of pictures in prose. In D. M. Willows & H. A. Houghton (Eds.), *The psychology of illustration: Volume I, Basic research* (pp. 52-85). New York, NY: Springer.
- MacWhinney, B. (2000). *The CHILDES project* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum.
- Makel, M. C., & Plucker, J. A. (2014). Facts are more important than novelty: Replication in the education sciences. *Educational Researcher*, 43, 304-316.
- Martini, F., & Sénéchal, M. (2012). Learning literacy skills at home: Parent teaching, expectations and child interest. *Canadian Journal of Behavioral Sciences*, 44, 210-221. doi:10.1037/a0026758
- Massaro, D. W. (2006a). A computer-animated tutor for language learning: Research and applications. In P. E. Spencer & M. Marshark (Eds.), *Advances in the spoken language development of deaf and hard-of-hearing children* (pp. 212-243). New York, NY: Oxford University Press.
- Massaro, D. W. (2006b). Embodied agents in language learning for children with language challenges. In K. Miesenberger, J. Klaus, W. Zagler, & A. Karshmer (Eds.), *Proceedings of the 10th International Conference on Computers Helping People With Special Needs (ICHP)*, pp. 809-816). Berlin, Germany: Springer.
- Massaro, D. W. (2015, February 2). Baby talk. *The New Yorker*, p. 3.
- MRC Psycholinguistic Database. (2015). Retrieved from http://websites.psychology.uwa.edu.au/school/MRCDatabase/uwa_mrc.htm
- National Early Literacy Panel. (2008). *Developing early literacy: Report of the National Early Literacy Panel*. Washington, DC: National Institute for Literacy.
- National Reading Panel, National Institute of Child Health and Human Development. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction* (NIH Publication No. 00-4769). Washington, DC: U.S. Government Printing Office.
- Nikolajeva, M., & Scott, C. (2001). *How picture books work*. New York, NY: Psychology Press.
- Norvik, P. (2012). *English letter frequency counts: Mayzner revisited or ETAOIN SRHLDCU*. Retrieved from <http://norvig.com/mayzner.html>
- Oakhill, J., Cain, K., & McCarthy, D. (2015). Inference processing in children: The contributions of depth and breath of vocabulary knowledge. In E. J. O'Brien, A. E. Cooke, & R. F. Lorch Jr. (Eds.), *Inferences during reading* (pp. 140-159). Cambridge, UK: Cambridge University Press.
- Olson, D. R., & Torrance, N. (1991). *Literacy and orality*. New York, NY: Cambridge University Press.
- Ong, W. J. (2002). *Orality and literacy: The technologizing of the word* (2nd ed.). New York, NY: Routledge.
- ParentFreq. (2015). Retrieved from <http://chilides.psy.cmu.edu/derived/parentfreq.cdc>
- Petrova, A., Gaskell, M. G., & Ferrand, L. (2011). Orthographic consistency and word-frequency effects in auditory word recognition: New evidence from lexical decision and rime detection. *Frontiers in Psychology*, 2, 263.
- Plato. (1925). *Phaedrus* (G. Crane, Ed. & H. N. Fowler, Trans.). London, England: The Perseus Digital Library. Retrieved from <http://data.perseus.org/citations/urn:cts:greekLit:tlg0059.tlg012.perseus-eng1:227a>
- Prinsloo, M., & Baynham, M. (2013). Editors' introduction: Literacy studies. In M. Prinsloo & M. Baynham (Eds.), *Volume I: Great divides and situated literacies. SAGE benchmarks in language and linguistics* (pp. xxiii-xxxv). Los Angeles, CA: SAGE.

- Providence Talks. (2014). Available from www.providencetalks.org/
- Read With Me! (2012). Retrieved from <http://psycenticmind.com/wp-content/uploads/2012/06/ReadWithMeBookList.pdf>
- Ritchie, S. J., & Bates, T. C. (2013). Enduring links from childhood mathematics and reading achievement to adult socioeconomic status. *Psychological Science, 24*, 1301-1308. doi:10.1177/0956797612466268
- Schleppegrell, M. (2012). Academic language in teaching and learning: Introduction to the special issue. *The Elementary School Journal, 112*, 409-418.
- Schmidt, S. (2009). Shall we really do it again? The powerful concept of replication is neglected in the social sciences. *General Review of Psychology, 13*, 90-100.
- Scribner, S., & Cole, M. (1978). Literacy without schooling: Testing for intellectual effects. *Harvard Educational Review, 48*, 448-461.
- Scribner, S., & Cole, M. (1981). *The psychology of literacy*. Cambridge, MA: Harvard University Press.
- Sénéchal, M. (2006). Testing the home literacy model: Parent involvement in kindergarten is differentially related to grade 4 reading comprehension, fluency, spelling, and reading for pleasure. *Journal for the Scientific Study of Reading, 10*, 59-87.
- Sénéchal, M., & LeFevre, J. (2014). Continuity and change in the home literacy environment as predictors of growth in vocabulary and reading. *Child Development, 85*, 1535-1551. doi:10.1111/cdev.12222
- Snow, C., & RAND Reading Study Group (Snow, C., Chair). (2002). *Reading for understanding, toward an R&D program in reading comprehension*. Santa Monica, CA: RAND.
- Socialnomics. (2015). Retrieved from <http://www.socialnomics.net/2014/07/28/25-years-of-the-internet-how-it-has-changed-the-way-we-interact/>
- sort. (2015). Retrieved from <http://man7.org/linux/man-pages/man1/sort.1.html>
- Stadthagen-Gonzalez, H., Bowers, J. S., & Damian, M. F. (2004). Age-of-acquisition effects in visual word recognition: Evidence from expert vocabularies. *Cognition, 93*, B11-B26.
- Stand Up to Bullying. (2015). Retrieved from https://www.ted.com/playlists/191/stand_up_to_bullying
- Talbot, M. (2015, January 12). The talking cure. *The New Yorker*. Retrieved from <http://www.newyorker.com/magazine/2015/01/12/talking-cure>
- Talks to Watch With Kids. (2015). Retrieved from https://www.ted.com/playlists/86/talks_to_watch_with_kids
- Tecwave. (2013). Available from <http://tecwave.soe.ucsc.edu/>
- uniq. (2015). Retrieved from <http://man7.org/linux/man-pages/man1/uniq.1.html>
- van Kleeck, A. (2003). Research on book sharing: Another critical look. In A. van Kleeck, S. Stahl, & E. Bauer (Eds.), *On reading books to children: Parents and teachers* (pp. 271-319). Mahwah, NJ: Lawrence Erlbaum.
- Waxman, S. R. (2002). Early word-learning and conceptual development: Everything had a name, and each name gave birth to a new thought. In U. Goswami (Ed.), *Blackwell handbook of childhood cognitive development* (pp. 102-126). Malden, MA: Blackwell.
- Wells, G. (1985). Preschool literacy-related activities and success in school. In D. R. Olson, N. Torrence, & A. Hildyard (Eds.), *Literacy, language and learning: The nature and consequences of reading and writing* (pp. 229-255). Cambridge, UK: Cambridge University Press.

Author Biography

Dominic W. Massaro is emeritus research professor of psychology at the University of California, Santa Cruz, and president of Psycientific Mind. His research interests include an experimental and theoretical analysis of spoken and written language processing. Applying technology and behavioral science to help solve real-world problems, he invented Kid Klok, an educational easy-to-read analog clock, developed successful products for language learning for language-challenged children such as those with hearing loss and autism, and is currently implementing a method for acquiring literacy informally without direct instruction. His research and applications have been published in a variety of journals and books.