

THE THEODOR GEISEL AWARD COLLECTION:
AN EXPLORATION OF EARLY READERS

A Dissertation Presented to the
Faculty of the College of Education
University of Houston

In Partial Fulfillment
of the Requirements for the Degree

Doctor of Philosophy

by

Eve Zehavi

May 2016

THE THEODOR GEISEL AWARD COLLECTION:
AN EXPLORATION OF EARLY READERS

A Dissertation for the Degree
Doctor of Philosophy

by

Eve Zehavi

Approved by Dissertation Committee:

Dr. Lee Mountain, Chairperson

Dr. Laveria Hutchison, Committee Member

Dr. Cheryl Craig, Committee Member

Dr. Margaret Hale, Committee Member

Dr. Robert McPherson, Dean
College of Education

May 2016

Acknowledgement

THE THEODOR GEISEL AWARD COLLECTION:

AN EXPLORATION OF EARLY READERS

An Abstract of
A Dissertation Presented to the
Faculty of the College of Education
University of Houston

In Partial Fulfillment
of the Requirements for the Degree

Doctor of Philosophy

by

Eve Zehavi

May 2016

Zehavi, Eve “The Theodor Geisel Collection: An Exploration of Early Readers.”
Unpublished Doctor of Philosophy Dissertation, University of Houston, May 2016.

Abstract

The importance of beginner books in the lives of nascent readers cannot be understated. In an effort to recognize these critical tools for early reading, the American Library Association created the Theodor Geisel Award in 2004. To date, no one has looked at this collection with a critical eye. The goal of this research is to define and analyze a group of texts meant to exemplify this genre. This work is particularly important because trade books are now a vital supplement to basals and workbooks in use in the nation’s PK–2 classrooms and function as a pivotal bridge between decoding and fluency necessary for comprehension of text. The research questions were addressed through multiple access points: a comparison of readability scales and word counts, syntactic parsing, linguistic coding, and, finally, narrative inquiry.

The results of this study underline the intricacies of even the simplest texts children read. The isolation of lexical words highlighted the importance and the vastness of vocabulary necessary for beginning readers. Syntactic analysis suggested a level of complexity not generally considered in books for young children. Equally revealing was the diversity of discourse registers evident in these trade books. Finally, visual analysis showed that comprehension skills and reading prosody are significantly enhanced in both text and image. It is hoped that this research will result in a meaningful and holistic view of beginning-to-read books that will assist teachers in their pedagogy and support of early literacy.

Table of Contents

Chapter	Page
Table of Contents	vi
Chapter 1 Introduction	1
Need for the Study	2
Statement of the Problem.....	3
Research Questions.....	5
Terms and Definitions.....	5
Organizational Overview	8
Chapter 2 Review of Literature.....	11
Reading Foundational Skills.....	12
Readability	16
Linguistic Measures	18
Cultural Considerations in Image and Text	23
Corpus Linguistics	27
Reader and Task Considerations.....	28
Conclusion	31
Chapter 3 Methodology	32
Purpose and Importance of Study.....	32
Study Design.....	34

Sample and Data Sources.....	35
Structure of Analysis.....	36
Analysis of Text at the Word Level.....	37
Analysis of Text at the Sentence Level.....	37
Analysis of Text at the Discourse Level.....	38
Analysis of the Reader–Text Relationship	38
Instrumentation	39
Validity/Reliability/Trustworthiness.....	39
Summary	40
Chapter 4 Results	41
General Attributes	41
Readability	45
Word Analysis	46
Sentence Analysis	52
Discourse.....	59
Contributions of Visual Analysis.....	66
Conclusion	68
Chapter 5 Discussion and Implications.....	74
Findings and Discussion	75
Research Question One:.....	75

Research Question Two:	78
Research Question Three	81
Research Question Four:	82
Considerations for Practitioners, Parents, and Prize Committees.....	84
Future Directions	85
Final Thoughts	86
References.....	88
Appendix A.....	99
Appendix B.....	102
Appendix C.....	108
Appendix D.....	112
Appendix E.....	114
Appendix F.....	116
Appendix G.....	123
Appendix H.....	125
Appendix I.....	132
Appendix J.....	134
Appendix K.....	136
Appendix L.....	138
Appendix M.....	140

Appendix N.....	151
Appendix O.....	153
Appendix P.....	155
Appendix Q.....	157
Appendix R.....	159
Appendix S.....	161
Appendix T	163
Appendix U.....	165
Appendix V.....	168

Chapter 1

Introduction

Through My Window (Mitchell, 1954) was the first book I ever read. I loved it beyond measure and I had anyone near me read it to me over and over until I could “read” it myself. Perhaps I had just memorized it, but I could point to each word and turn the pages appropriately. What is important is that I still remember the pride I felt in being able to read it myself. That epiphany, when you read a book all by yourself and realize that worlds have opened up to you, is powerful.

So, it was surprising for me to learn that no one has researched the Theodor Geisel Award books for beginning readers. It is these very books that children will remember all their lives. The Theodor Geisel Award was initiated in 2004 to honor the memory of Theodor Seuss Geisel, who wrote books for children under the pseudonyms Dr. Seuss, Theo LeSeig, and Rosetta Stone. This award is given by the American Library Association (ALA) to contemporary authors with distinguished work “that encourages and supports the beginning reader” (ALA, 2015). These books are specifically designed for young readers in pre-K through grade 2.

Early readers have been published for many years and are distinguished from other children’s books in that they are designed specifically for independent (child) reading. Early commercial publishing endeavors that specifically targeted early readers include the Little Golden Book series, first published by Simon and Schuster in 1942, and the Beginner Book series, first published by Houghton Mifflin in 1958. (Marcus, 2007; Morgan & Morgan, 1995). These series set themselves apart from other trade publications for children in that their production and design were founded in the most

current research-based theories in literacy. The production of Little Golden Books was overseen and endorsed by Mary Reed, head of the Kindergarten and Elementary Education Department at Columbia Teachers' College. A progressive educator of young children, Reed's research advocated for child-centered stories about everyday experiences. Sixteen years later, The Beginner Book series had its own promoter, Dr. Seuss's publisher, Phyllis Cerf, who tirelessly researched the best early reading practices of the day, which included limited and controlled vocabulary, fewer words per page, larger typeface, and illustrations directly parallel to the text (Morgan & Morgan 1995).

Although the Theodor Geisel Award books are not produced by a single publisher with an underlying philosophy, the criteria for the award stipulate several factors well known to be important to early reading efforts: "Illustrations which function as keys or clues to text"; repetition of words; simple and straightforward sentences; and a motivating "plot, sensibility, and rhythm that can carry a child along from start to finish" (ALA, 2015). The goal to acknowledge the importance of beginner books in the lives of nascent readers is crucial today because trade books are a vital supplement to basals and workbooks in use in the nation's PreK-2 classrooms. Beginner books are a pivotal bridge between decoding and fluency necessary for comprehension of text and commonly used both as supplements and as part of the core curriculum.

Need for the Study

Book awards reflect the importance the greater educational community puts on literacy. The Newbery and Caldecott award winners have been examined with some frequency, but unlike other national award books for children, the Theodor Geisel Award winners have yet to be examined and analyzed. The exploration and analysis of these

texts will provide insights into the values and substance of early literacy education with respect to the elements of learning to read, such as foundational skills, vocabulary development, and comprehension, as well as larger cultural values such as diversity and fairness we wish to inculcate in young children. Ultimately, this study is intended to describe a collection of texts selected by professionals to be “distinguished” in this genre and to evaluate in what ways they live up to or fall short of this measure.

Statement of the Problem

There are a number of considerations that led to this analysis. Lists of award-winning books are often used as a collection development tool by teachers; however, the selection of these books is done by a panel of librarians rather than educators. While other awards deal in aesthetics, such as Caldecott, or narrative (storyline/theme), such as Newbery, Theodore Geisel Award books have a purpose beyond enjoyment—to teach. So, it is important that educators be part of the conversation. Moreover, there is an absence in the research of early reading regarding semantic and syntactic analyses (Mesmer, Cunningham, & Hiebert, 2012). New computational models make it possible to examine texts in ways heretofore unavailable. Finally, research tells us that many teachers of early reading are unaware of the many different attributes inherent in beginning texts and the means to address them with young children in ways that will foster success (Juel & Minden-Cupp, 2000).

Considerations that form the basis of this study include analysis of foundational skill supports, text complexity including vocabulary, and an examination of reader/task considerations. This study will examine in depth the attributes of the Theodor Geisel Award and honor books in order to highlight the appropriateness of both texts and images

in facilitating emergent and early reading. Further, it will address how well the collection supports the Theodor Geisel Award criteria and the new and more stringent standards of the Common Core Curriculum. Thought will be given to the relationship of interactional theory in a selection process that does not include the reading audience.

While experts may disagree on the theoretical framework for teaching early reading and the effectiveness and importance of differing methodologies, the reality in today's preschool and elementary classrooms is that the formal introduction of reading skills is occurring earlier, fostered in part by the political climate surrounding student performance and accountability since the adoption of No Child Left Behind. Over time, this directive has resulted in the "pushing down" of academic skills to younger and younger grades/audiences.

New standards are evident in the documentation of the Common Core State Standards (CCSS) that deal with text complexity, so a comparison of these measures to trade books for emergent readers is appropriate. Stimulated in part by the CCSS which address primarily upper elementary and secondary texts, Mesmer, Cunningham, and Hiebert (2012), call for an examination of the "underlying assumptions about early-grade texts" (p. 236). Among their suggested considerations are issues related to (1) word complexity ie. structural complexity, word familiarity, and semantic features; (2) syntactic complexity ie. sentence length and syntactic awareness; and (3) discourse structure, ie. text length, cohesion, and genre. It is, in part, this challenge that prompted the following research. While the CCSS standards do not address illustrations, they are, in fact, an important aspect of early literacy and play a crucial part in supporting early readers. This study treats visual analysis as critical to early reading, especially given the

criteria of the Theodor Geisel Award, which refer specifically to artwork in early reading texts.

Beyond simply describing the collection of Theodore Geisel Award and honor winners, the study hopes to shed light on larger questions: What do teachers and students want these books to do? Should they prepare students for more difficult reading? Are they meant simply to entertain? How can educators access the potential of these texts to promote reading skills as well as content? How do we want beginning-to-read books to convey the values of society?

Research Questions

This study will address the following research questions.

Research Question One: To what extent do the Theodor Geisel Award and honor books address foundational reading skills?

Research Question Two: To what extent do the Theodor Geisel Award and honor books measure up against both quantitative and qualitative standards for text complexity in the CCSS?

Research Question Three: To what degree do the Theodor Geisel Award and honor books reflect the core literacy values related to the variety of human experience including behavior, and gender?

Research Question Four: How can computational linguistic analysis be used to inform knowledge of texts for application in the classroom?

Terms and Definitions

ATOS Level. This readability measure uses quantitative methods based on sentence length, word length, and word level.

Independent Clause. A group of words including a subject and predicate and able to stand alone.

Complex Nominal. This refers to the complexity of the group of words used to express a single entity. The nominal can include adjectival, prepositional, participial, and infinitive phrases in addition to the noun.

Computational Linguistics. Computer based methods of statistical modeling of natural human language.

Controlled Vocabulary. This restricted vocabulary is chosen from specific word lists (Spache, 1953).

Dependent Clause. A group of words which cannot stand alone as a sentence.

Emergent Literacy/Early Literacy. This term refers to a child's knowledge of reading and writing before acquiring formal reading and writing skills.

Grammatical Words. Words which have grammatical rather than content value, sometimes referred to as structural words.

Lemma. A lemma is a base word which stands for all of a word's forms. For example: *run* is the lemma for *runs*, *running*, *ran*.

Lemmatization. In education this refers to the process of finding the base word as well as the set of all forms that have the same meaning.

Lexical Diversity. This measure refers to the number of occurrences of a word within a text such as Type/Token ratio.

Lexical Words. These are words which convey meaning or content, typically nouns, verbs, adverbs, and adjectives.

Lexile Level. The Lexile Framework is a readability scale which uses quantitative methods, based on sentence length and word frequency.

Orthography. This is recognition of graphemes, written patterns of syllabication, morphemes, and derivational suffixes.

Phoneme-Grapheme Correspondence. This is the correspondence of letters or letter combinations to speech sounds.

Phonological Awareness. This is recognition of rhyme, alliteration, assonance, and syllabication.

POS tagging. This is the process of labeling the part of speech which each word in a sentence occupies.

Prosody. From text analytics, this is a way of measuring language flow based on syntactic units and sentence complexity.

Qualitative Text Complexity. From the CCSS, this refers to qualitative measures including levels of meaning, structure, language conventionality and clarity, and knowledge demands.

Quantitative Text Complexity. From the CCSS, this refers to quantitative measures such as readability scores.

Readability. This term refers to the difficulty of reading a passage. This measure is based on a number of different formulas.

Readers' Advisory – A service which involves suggesting appropriate titles for readers based on their interests and abilities.

Reading Foundational Skills. From the CCSS, these include phoneme/grapheme correspondence, phonological awareness, and orthography.

Sentence. For the purposes of this study a sentence is delineated by orthographic characteristics: i.e., text that begins with a capital letter and ends with final punctuation is considered a sentence. For example, “Wow!”

Stoplist. In computational linguistics this refers to a list of words filtered out of a text. These are generally high frequency or grammatical words.

Syntactic Complexity. This refers to measures of word patterns such as noun and verb clauses and dependencies within these patterns.

Token. The instance of a sequence of characters occurring together, generally words. In the case of this research words and punctuation will be tokenized separately.

Tokenization. A process of counting and iterating every token in a text. For the purposes of this study, tokenization consists of counting every discreet word in a text even if it occurs multiple times.

T-unit. This is the shortest grammatical unit which can stand alone as a full sentence. Sentences with co-ordinate clauses constitute multiple t-units.

Type. The net total of tokens. For example in the sentence “To be or not to be.” there are 4 types.

Vocabulary Tiers. From the CCSS, Tier I words are defined as coming from everyday speech, Tier II words are referred to as general academic words, and Tier III words are defined as domain specific (CCSS, 2012).

Organizational Overview

This study consists of five sections, an Introduction, Review of Literature, Methodology, Results, and Analysis along with references and appendices. The introduction posits the need for a study of the Theodore Geisel Award books and the

questions and purposes inherent in this need. The literature review that follows addresses several pertinent areas of research related to reading, in particular the historical foundations of early readers, a brief overview of research related to foundational reading skills, a description of the use of text analytics in content analysis, and a brief discussion of the CCSS for measuring text complexity. A short portion of the literature review addresses illustration and its importance in the reading process, and finally, cultural factors as they relate to both text and image are considered.

The methodology section briefly addresses the description and design of the tools used to assess the Theodor Geisel Award collection. Each book is analyzed in a comprehensive process that looks at general attributes such as title/author, number of pages, words per page, number of syllables, and vocabulary and reading levels. Computational linguistics is used to enhance the analysis of qualitative factors in texts such as sentence structure and figurative language. Literary characteristics such as character, plot, use of dialogue, and theme are noted and discussed as appropriate. Some consideration is given to artwork in form and substance, the socio-cultural values it portrays, and its relationship and reflection of words in the text. Lastly, a narrative of personal impressions and pedagogical judgements which ultimately informs the discussion, was written for each text.

The data is related in tables, graphs, and narratives, with discussion on the three aforementioned elements: quantitative measures, qualitative measures, and reader-text relationship. Analysis and conclusions examine the data relative to the four original research questions that ask about the extent to which the Theodor Geisel Award and

honor books conform to a variety of standards and how technology can be used by professionals to assist in their students' attainment of early reading skills.

Chapter 2

Review of Literature

Looking in depth at the intrinsic attributes of works, rather than at the reader, is relatively new in reading research, but the CCSS and its ideas related to text complexity have put new focus on the texts themselves. For the most part, the CCSS addresses upper elementary and secondary language arts standards, but texts for learning to read also have special attributes that make them appropriate for study. It is vital to understand how readers acquire foundational reading skills as well as how the texts they read might support or hinder reading fluency. What follows is an introduction to the literature that informs this examination of beginning-to-read books.

The sections are titled 1) reading foundational skills, 2) readability, 3) linguistic measures, 4) cultural considerations, 5) corpus linguistics and 6) reader and task considerations. The first section (reading foundational skills) relates to foundational reading skills that foster early reading success. This includes elements necessary for decoding: alphabetic knowledge, sound discrimination, and an ability to partner these two skills. In addition to addressing decoding skills, this section explores research related to the influence of prosody on reading comprehension. The section following reading acquisition moves to issues related directly to text analysis. This portion of the literature review includes three divisions. The first (readability) deals with the history and philosophy of readability measures. The second section (linguistic measures) goes beyond the quantitative aspects of readability to include the history of more qualitative measures and how the areas of linguistics and disciplinary literacy inform our knowledge of texts. A third section (cultural considerations in image and text) is devoted to a review

of literature that addresses the relationship of illustrations to texts and the concomitant cultural messages that are present in these images. A brief discussion of corpus linguistics follows with ideas regarding the application of computational linguistics to pedagogical practice. Lastly, this review will touch on the reader–text relationship with particular emphasis on the teacher’s role in reader’s advisory.

Reading Foundational Skills

In order to fully appreciate how texts can nourish new readers, it is important to understand the essential abilities necessary to learn to read. These primary skills include knowledge of how words are created, how sentences are composed, and how discourse is organized. There has been a large amount of research done at the word level. Today we have both wide and deep analysis of decoding as it relates to reading acquisition. From this we know that prereading children must have alphabetic knowledge or letter recognition, the ability to discriminate sounds, and the ability to match letters and sounds to construct or deconstruct words. Phonological awareness progresses from sensitivity to letter sound correspondence and syllabic beats to an ability to manipulate them (Morais, 1991). There is a general acceptance in the literature of two approaches to phonemic awareness: synthetic and analytic. Synthetic phonics is constructive and requires children to blend sounds together, i.e., breaking the word *cat* into discreet sounds, K-A-T, and blending them together (National Literacy Trust, 2015). Analytical phonics deconstructs words and requires students to find sound patterns. An example children might encounter: What do these have in common? Dog, desk, dim.

Much of the early literature related to decoding focuses on theories of developmental stages or phases. Frith (1985) described three phases: logographic,

alphabetic, and orthographic. The first refers to a child's ability to recognize objects and symbols—what we might refer to as environmental print such as associating the big yellow *M* with McDonalds. The alphabetic phase begins with the development of letter/sound correspondence. The final stage is orthographic knowledge in which readers are able to automatically recognize whole words without needing to sound them out (Frith, 1985). Jeanne Chall (1983), around the same period, proposed a developmental stage model of reading. Early readers begin in stage 0, which includes print awareness skills or “pretend reading.” Stages 1 and 2 mirror Frith's alphabetic and orthographic phases, which lead to the end of the “learning-to-read” phase and moves readers toward automaticity and fluency.

There is some research to suggest that oral language vocabulary supports the acquisition of phonemic awareness. This theory, the lexical restructuring model, suggests that children learn words as wholes and the more closely they are related in sound, the more children are forced to discriminate individual sounds (Walley, Metsala, & Garlock, 2003). In other words, when children learn words like *house* and *ball*, the sparse relationship between the sounds of these two words makes them easy to remember as wholes, but words with more common sounds like *these* and *tease* force the learner to discriminate on the basis of single phonemes or sounds.

While this refers to oral language discrimination, other theorists believe that early readers also recognize letter combinations that become intermediary units of reading. This is particularly true of “transparent” or completely phonetic languages in which most or all sounds are represented by a single orthography, as is true of Spanish or Italian (Seymour, Aro, & Erskine, 2003). These mediating supports are evident in English as

well. According to Ziegler and Goswami (2005), children who grow up with languages that have irregular phoneme/grapheme correspondence use intermediary supports as well, such as morphemes and rhyme.

Beyond the decoding of words, fluency in reading is accomplished through automaticity, which includes whole word recognition and decoding speed and prosody, which includes volume, duration, pitch, and pauses in reading. While we know that prosody is important to fluent reading and ultimately comprehension, research in this area is only in its infancy. To date, there have been two ways in which this factor has been measured: by testing oral reading prosody and by examining the syntax of texts. On its most simple level, prosody in oral reading can measure a child's ability to make a distinction between declarative and interrogative sentences by changing the inflections at the end of a sentence. A child's understanding of syntax might be exemplified by his or her ability to complete a sentence with the proper part of speech. For example, "Humpty dumpty sat on a ____." The expectation would be that a child could fill in the blank with a noun that makes sense.

Many theorists have understood that the ability to parse sentence syntax is vital to reading fluently (Young, Bowers, & MacKinnon, 1996; Grassmer, McNamara, Louwse, & Cai, 2004). In a 1991 study, Schreiber suggested the connection between prosody and syntax by demonstrating a lack of syntactic understanding for elements such as lengthening of the final syllable, pitch changes, and pauses that occur in phrases and sentences, which was significant in children who have difficulty comprehending (Schreiber, 1991). Esser and Polomski (1988) went a step further to suggest that "prosody

interacts with both syntax at the phrase level and semantics at the passage level” (cited in Benjamin & Schwanenflugel, 2010, p. 390).

Though there is little related to the pedagogy of prosody and syntax for the early reader, research indicates this as an area of weakness among elementary students. Both National Assessment of Education Progress (NAEP) and the National Institute of Child Health and Human Development (NICHD) have conducted studies using syntactic measures to assess fluency. The latter research found that only 10 percent of fourth-grade students in their study had enough understanding of prosody in the reading process to read with “meaning, phrasing, appropriate representation of syntax, and interpretive expression (NICHD, 2000).

Solutions to this problem have been hinted at by Benjamin and Schwanenflugel, who found in their 2010 study of prosody that on three of four variables, children scored better for prosody with difficult text. This skill seems to be reciprocal; according to Young, Bowers, and McKinnon (1996), student ability to parse sentences improves as reading matter becomes more complex. Another study that could perhaps have a future impact on teaching literacy skills was conducted by Holliman et al. (2014), which tested stress, intonation, and timing as discreet skills at three different linguistic levels: word-level, phrase-level, and sentence-level. The study found that these elements all impact reading ability and appear to have distinct relationships to a variety of reading measures. Further, these elements of prosody had a greater impact on reading than the size of the linguistic units (Klauda & Guthrie, 2008). However, none of these studies have led to measures that can be used to identify text difficulty and potential pedagogical frames for choosing texts. Until now, educators have relied on readability measures to select texts.

Readability

According to the CCSS, “the terms *quantitative dimensions* and *quantitative factors* refer to those aspects of text complexity, such as word length or frequency, sentence length, and text cohesion, that are difficult if not impossible for a human reader to evaluate efficiently, especially in long texts, and are thus today typically measured by computer software” (CCSS, 2012). Quantitative measures have a long and respected history in reading research. Readability studies and the development of readability formulas inform most of the quantitative constructs for complex text evaluation supported by the CCSS. While the standards use Lexile as the measure of choice, using the wider variety of available readability scales gives a truer picture of the difficulty level of books in the Theodor Geisel collection. In general, these formulas were derived with the intent of analyzing texts so that they could be graded and placed appropriately in the hands of students at a level at which the child could comfortably read and understand the material. For the most part these formulas are based on a combination of grammatical difficulty usually measured by sentence length and vocabulary. Vocabulary analysis varies among the studies in that sometimes it is measured by word length, for example, number of letters or number of syllables in a word, or by frequency, either measured directly or compared against a predetermined list.

The earliest attempts to grade texts were manifested in basals in which lexical control was imperative and relied heavily on lists such as those by Thorndike, Dolch, and Gates. Prior to these criteria, researchers like Thorndike and his successors relied heavily on word frequency counts as a measure of controlling for difficulty in texts (Dale, 1956). Ultimately, this strategy made for dull and contrived texts that later researchers would

reject and which ultimately lead to the whole-language movement significantly complicating the process of determining text complexity.

In 1948, Rudolph Flesch published one of the first widely accepted readability measures, the Reading Ease Readability Formula. It consisted of two measures: average sentence length and average syllables per word. These were derived from a 100-word sample of text and resulted in a measure from 0 to 100 (Flesch, 1948). He would later revise his formula together with J. Peter Kincaid. The Flesch-Kincaid readability formula was based on the earlier calculation but translated the results into an approximate grade level (Kincaid, 1975). Fifteen years after Flesch's first scale was published, Jeanne Chall coauthored a new readability formula with Edward Dale in an attempt to more accurately judge reading level (Dale & Chall, 1948). This formula sought to improve on the average syllables per word (ASW) portion of Flesch's readability scale by excluding a list of 3,000 easy words, so that the measure was calculated on the proportion of words not appearing on the list. Moreover, the computation of readability was then calculated on the aggregate of 100 word samples taken every 10 pages in a book or selection (Dale, 1956). These were the gold standard until the late 1960s. In 1968, Edward Fry would publish his readability graph, especially popular for its ease of use. It consisted of taking averages of the number of words and the number of sentences in three 100-word selections and plotting them on a graph to determine grade level (Fry, 1968).

Later readability measures would attempt to create more complex and therefore more accurate formulas. Among these are Lexile, DRP, and ATOS. Lexile measures differ from earlier measures in their attempt to scale both reading ability and text complexity. To define text complexity, Stenner and Smith, the creators of the Lexile

indices, based their construct on three characteristics: word frequency in large samples of written works, the likelihood of encountering those words (dispersion), and whether the word is concrete or abstract (Stenner, Smith & Smith, 1988). DRP is a non-normed readability scale that also aligns readers and texts. It uses four linguistic variables for its scale letter count, words on the Dale-Long list, sentence length, and cloze scores. DRP is then indexed to report independent, instructional, and frustrational levels. ATOS is the most contemporary readability scale and uses data from DRP tests to weigh common readability variables: words per sentence, average grade level of words, and characters per word. Further, unlike other readability scales, ATOS accounts for derivatives. It also is easy to use because it is scaled by grade level (Krippendorf and Bock, 2009).

Yet, despite all the advances in readability formulas, many prominent educators have spoken out against the idea of identifying complex texts, citing shortcomings in any and all of the quantitative measures used to evaluate them. But educators know that formulas are only one aspect of readability considerations.

Linguistic Measures

Throughout the history of readability studies, there has been a quiet knowledge that quantitative dimensions of readability are limited. These dimensions are considered so important, in fact, that they come first in the CCSS standards and are explained in great length. Their history, however, is more complicated and derives from several areas of educational research.

Klare (1963), in his foundational book on readability cited four limitations of the readability formulas. Firstly, they measure “only one aspect of writing – style”; readability measures do not account for content or complexity of ideas presented.

Second, these measures only account for difficulty in style as indicated by sentence length. Thirdly, they are at best, rough estimates of level of difficulty. And, finally, they do not in any way suggest the value of a text (Klare, 1963). This is supported by the CCSS, which state that “qualitative dimensions and qualitative factors refer to those aspects of text complexity best measured or only measurable by an attentive human reader” (CCSS, 2012).

Edward Fry, long a proponent of readability formulas, articulated the difference between quantitative and qualitative measures in his article “Readability versus Leveling” (2002). According to Fry, leveling is more subjective and “takes a number of ‘text support’ factors into consideration.” These include: content, illustrations, length, curriculum, language structure, judgment and format (Fry, 2002). Leveling is not an entirely new concept. Marie Clay (1991), in her work with early readers in primary grades, developed a system for leveling texts because readability formulas are not particularly sensitive at the lowest levels (Fry, 2002). Fountas and Pinnell (2001) also offered a qualitative method of leveling books. In addition to text difficulty based on vocabulary and sentence complexity, a number of additional factors are taken into account when leveling a book: book and print features, content, text structure, language and literary features, and themes and ideas (Fountas & Pinnell, 2001). The CCSS lists four aspects of qualitative analysis: levels of meaning (literary texts) or purpose (informational texts), structure, language conventionality and clarity, and knowledge demands (CCSS, 2012).

In beginning readers, language considerations are particularly significant. There are phonemic aspects that must be considered in books designed to “teach” children to

read. Juel and Minden-Cupp's article about skills and strategies associated with beginning reading provides a strong theoretical foundation in the literature related to both teaching and learning phonics (2000). There are two main foci of the article: one relates to what linguistic units assist early readers, and the second is instructional strategies that teachers use to foster beginning reading. The authors outline decoding in several ways: sounding and blending phonemes, sounding and blending an onset and rime, using word walls for support, combining initial letter-sound with words that might potentially make sense, visual strategies based on what a word "looks like," and specific knowledge of phonics rules (Juel & Minden-Cupp, 2000). While the authors suggested that all of these strategies are dependent upon the type of practice provided and the types of activities offered, such as spelling and reading, teacher modeling, and the degree of individualized instruction, it is imperative to consider what inherent qualities in the reading materials students encounter that might support or undermine this instruction.

The Common Core Standards provide a framework for looking at higher grade level texts which includes levels of meaning, structure, language conventionality and clarity, and knowledge demands. However, these categories alone are not fully adequate to the task of a qualitative evaluation of beginning-to-read books. Though there has been no linear historical path to determining qualitative dimensions of text complexity, two areas of educational research seem to inform the debate: English as a second language and content-area reading. Hadaway, Vardell, and Young (2004) stated that teachers of English as a second language (ESL) need a working knowledge of the basic elements of linguistics such as syntax, semantics, pragmatics, discourse structure, and rhetorical registers. A deeper look into these aspects of language leads to the underlying

contributions of psycholinguistics to the qualitative analysis of texts features. Patricia Carrell (1987) analyzed these ideas in her article “A View of Written Text as Communicative Interaction: Implications for Reading in a Second Language.” She referred to Beaugrande and Dressler’s (1981) seven standards of textuality: cohesion, coherence, intentionality, acceptability, informativity and situationality. These six standards offer an additional framework for addressing qualitative factors.

In this context, *cohesion* refers to sequencing of surface text and grammatical structures. This is echoed in the second aspect of the qualitative dimensions of the CCSS: Structure. “Texts of low complexity tend to have simple, well-marked, and conventional structures, whereas texts of high complexity tend to have complex, implicit, and (particularly in literary texts) unconventional structures.” Linguistic *coherence* relates to implicit or inferred knowledge not explicitly stated in the text (Carrell, 1987). Again, looking at the CCSS guidelines we find reference to language conventionality and clarity, which suggest that literal, clear language is less complex than “ironic, ambiguous, purposefully misleading, archaic or otherwise unfamiliar language,” which constitute more difficult text features (CCSS, 2012).

Intentionality and *acceptability* are aspects of text that deal with whether or not the text achieves its goals and if the communication of those goals is useful and relevant. These aspects are considered in the first level of qualitative measures, levels of meaning or purpose. Understandably, literary texts such as satire, “in which the author’s literal message is intentionally at odds with the underlying message,” are considered complex (CCSS, 2012). Conversely, informational texts with clearly stated purposes are considered less complex. While it may appear that these relate to advanced

comprehension, in early reading books these levels of incongruence are often conveyed by images. The fifth element of linguistic textuality that informs the CCSS is *informativity*. A balance of *informativity* is critical to the idea of complex texts—finding equilibrium between the demands of the text and the breakdown of communication. The interplay between text and image is, in fact, an important aspect of early literacy and plays a crucial part in supporting early readers. Children “use” images to read a story (Albers, 2008). Visual supports also help children to read character emotions, understand plot/action, and even build schema about cultural norms.

The last linguistic standard, *situationality*, refers to context (Carrell, 1987). Jerome Bruner, another linguist, extended this idea to content by offering that disciplines communicate through various structures unique to their perspectives on the world and their methods of constructing knowledge (Bruner, 2009). This is no less so in beginning reading books, which include fiction and nonfiction texts and formats such as graphic novels and diaries. Content provides the foundation for literacy and the use of critical and academic language in texts, just as it does for second language acquisition. It is no wonder that Gibbons asserted that language instruction should not be taught “in isolation from the rest of the curriculum” (Gibbons, 1993).

As mentioned above, content-area studies have also been concerned with their own set of qualitative measures. Smith (1964) identified several distinct patterns in content-area reading, such as statement of fact, explanation of process, classification, descriptive problem solving, abbreviations, and equations, which she attributes to the texts themselves. Herber (1978) referred to these as Readability-Reading Ability/Subjective Data and discussed the text in terms of literal level statements,

interpretive level statements, and applied level statements. He wrote that the whole idea of content-area reading is meeting the challenge of students as they “encounter and deal with new knowledge and new ideas” (Herber, 1978). Though at first glance this would seem not to apply to early-reading books, these challenges are particularly important as the wide range of books for emergent readers are the backbone of developing schema as it relates to different “types” of reading.

Cultural Considerations in Image and Text

As early as the 1930s, researchers provided evidence that underlines the relationship between language, story, and identity (Mead, 1934). According to Taylor (2000), as children learn to read they become acculturated to greater societal messages. This speaks to the analysis of ethnicity, gender, and age within the Theodor Geisel collection and how these messages influence young readers. Since Mary Reed’s participation in the production of Little Golden Books, it has been generally accepted that literature for children should be child centered stories about everyday experiences. To foster a love of reading, young children have to see themselves in the text.

Early readers and children’s books present an additional visual component that is as equally meaningful as the written word. In “Theorizing Visual Representation,” Albers (2008) suggests a strong connection between art and literacy. Through illustration of text, children “experience and build beliefs about the world” (Albers, 2008, p. 165). Art not only offers clues to the text, but often allows children to reflect on or even challenge their perspectives (Albers, 2008). Our stories are a primary vehicle for transmitting cultural values, and children’s books are an important component of this transference (Arbuthnot,

1964; Kortenhaus & Demarest, 1993). Often, illustrations provide a particular interpretation of the greater culture that may not be evident in the text.

One aspect of this concerns race and ethnicity. However little research evidence exists to suggest that motivation is greater when materials reflect a student's race or ethnicity (Williams, 2008). Still this does not negate the idea that literature can in fact support a child's identity with "authentic perspectives, flavor, and illustrations" (Collier, 2000, 236). More importantly, especially with respect to the Theodore Geisel Collection is the ability of children's books to portray characters within the greater societal context (Collier, 2008). Exposure to cultural and racial diversity in texts prevents what Banks (2009) termed "ethnic encapsulation." As Kruse (2001) eloquently put it, children need to "see the world reflection in all its diversity as well as to confirm the universality of human experience," (27).

Our human experience is reflected in representation of gender as well. Taylor (2000) suggested that along with supporting children's reading skills, children's books transmit a number of cultural norms including ideas about what it means to be male or female (Taylor, 2000). When we understand that gender identity, a child's self-awareness of gender and its meaning, begins to be internalized as early as age 4, examining children's books for this characteristic becomes a compelling issue

Peterson and Lach (1990) put it very poetically, "Picture books offer young children a macrocosmic resource through which they can discover worlds beyond their own life space" (p. 189). In terms of the perpetuation of gender roles that can be defined as stereotypic male or female behaviors, which include assumptions about occupations, attitudes, appearances, abilities and interests, children's books contribute significantly.

As children are exposed to literacy either by reading themselves or being read to, they become acquainted with normative roles and responsibilities of the sexes (Taylor, 2000). Over the last several decades, researchers have examined children's books for a number of gender markers and issues.

Several studies have taken a quantitative approach by simply examining the number of times males or females appear in the titles or as major characters in a text. According to Hamilton, Anderson, Broaddus, and Young (2006), at the beginning of the 21st century, children's books continued to under represent females, who encompassed less than half of the title and main characters among books produced in the late 1990s and early 2000s. Additionally, this study calculated the appearance of nurturing traits, domestic or outdoor settings, and occupational status. Their findings, when compared with earlier research revealed no reduction in sexism (Hamilton et al., 2006).

Though studies beginning in the early 1970s looked at the invisibility of women and girls, they also closely examined gender traits. Weitzman, Eiffer, Hokada, and Ross (1972) found that girls were portrayed as passive followers and were presented primarily in indoor scenes while boys were depicted as active leaders and appeared engaged in outdoor activities. A later study by Williams, Vernon, Williams, and Malecha (1987) indicated that children's books during the 1980s depicted females as less independent than their male counterparts and for the most part, without career goals. Still more recent studies on Caldecott books found that women and girls are still encountered in traditional gender roles, doing domestic chores, while males tended to be linked with outside occupations (Crabb & Bielawski, 1994). Overall stereotyping and gender inequity has shown some improvement in terms of the ratios of female to male characters over the last

50 years, but the definition of gender roles seems to have remained static (Hamilton et al., 2005).

Another aspect of gender messages evident in children's literature relates to parenting. In 2005, Anderson and Hamilton looked specifically at the presence or absence of fathers in children's literature. According to their study, annual sales of children's books approach \$1 billion per year. With so many books in circulation, they expressed concern that the "representations of parental roles may influence the socialization" of both children who listen and parents who read to them each evening (Anderson & Hamilton, 2005). When present in children's books, fathers more frequently showed anger and were presented as disciplinarians. Even books tagged as being "nonsexist" still fail to show males engaged in traditionally female roles with respect to household participation (Diekman & Murnen, 2004).

Others research studies demonstrate that children's literature has an influence on the behavior of children and their attitudes toward gender roles (Trepanier-Street & Romatowski, 1999). Narahara's (1998) research on nonsexist books reinforced these findings; children's behaviors, attitudes, and self-concepts regarding gender roles appeared to be more positive after exposure to non-stereotypic characters.

Overall, gender bias in books for children is an important discussion because we know that the effects of stereotyping influence children's attitudes toward their futures both in terms of career aspirations and their roles as parents (Hamilton et al., 2006). This relates to one final aspect of the human experience evident in children's books – the representation of age. Early research in this field reported a prevalence of stereotyping with respect to the elderly. (Ansello, 1978, Taylor, 2000). However, Mavrogenes (1982)

suggested that there is a trend toward more realistic portrayals of the elderly and increasingly positive representations. In general, we know that the stories children read give children a schema for how they see themselves and others in the world and provide implicit messages about what it means to be male or female, young or old, and how these roles are sanctioned by society (Peterson & Lach 1990; St Peter, 1979; Rachlin & Vogt, 1974).

Corpus Linguistics

A brief explanation of how the application of computational methods to written language is used to understand texts is presented here. The simplest definition of corpus linguistics is the use of statistical tools for the analysis of digitized texts. These analyses are based on the calculation of frequencies of occurrences of individual linguistic units and co-occurrences of multiple linguistic units. These units, called tokens, can include phonemes, words, morphemes and punctuation. Over time, as more and more texts have been analyzed, computers have learned to demonstrate these frequencies in a variety of ways that communicate lexical, syntactic, and semantic relationships. Stefan Th.Gries (2009), explains that “formal differences reflect, or correspond to, functional differences.” In other words, frequencies of linguistic units indicate patterns which in turn can be interpreted by researchers and used as a means to talk about how we use language.

Marks, Doctorow, and Wittrock (1974) were early researchers that posited a direct correlation between word frequency and comprehension on which much of the pedagogical application of corpus analysis is based. More contemporary researchers have expanded this basic concept and applied it to text analytic software in order to

measure a variety of text attributes, specifically lexical sophistication, which refers to the level of difficulty of a word, and lexical diversity, which refers to the number of words in a text. (Stevens, Lu, Baker, Ray, Eckert, & Gamson, 2015).

Biber and Conrad (2001), in their research on English language learning used corpus analysis to explore syntactic features. Their study focused heavily on language register and found that the 12 most frequent verbs occurred in 45 percent of conversational language. Additionally, they isolated the most common verb tenses in both conversation and literature. Conrad (2000) in a separate study, illustrated how grammatical structures commonly occur with a fairly limited number of words and forms. She went on to discuss the pedagogical implications of corpus analysis adding that “the teaching of grammar will become more integrated with the teaching of vocabulary.”(549).

Syntactic complexity is based on word frequencies as well but using measures of multiple word phrases and clauses. Covington (2006), conducted studies based on child language acquisition which ranked syntactic constructions based on language development. One can easily see the value of such research to early readers.

Reader and Task Considerations

Qualitative dimensions, like quantitative analysis, are critical components of how we look at texts but exist in tandem with reader and task considerations. A child’s encounter with early readers, like those in the Geisel collection, demands some analysis of the reader/text relationship because knowledge of how children experience texts both in terms of reading ability and prior knowledge can help teachers facilitate reading growth. Transactional theory is an important aspect of any reading experience, however,

in the case of the beginning readers, books are not always chosen by their intended reading audience. To extend this idea, some scholars even suggest that authors and illustrators, though not present physically, form a teacher-student relationship (Smith, 1992; Spencer, 1988). Juel and Minden-Cupp's discussion of a self-teaching hypothesis also supports the idea that authors of children's books act as teachers of reading (2000). E. A. Betts (1946) dealt with the question of matching readers to texts. His formula, as it were, was based on the number of errors made while reading orally. In a nutshell, Betts assigned levels of ability based on a 5% error measure. Reading materials in which students encountered a 5% error rate were considered instructional level, more than 5% were considered frustration level, and less than 5% were considered independent reading level (Betts, 1946, p. 439). But the entire idea of leveled texts has, until now, implied that what students read should be at an independent level. It sets a low bar for inaccuracy and understanding. Shanahan (2013) took issue with this. According to him, "instructional level theory . . . ignores the guidance, support, and scaffolding provided by the teacher." Nor does it help students learn the most or achieve optimum levels of understanding. Shanahan (2013) further asserted that with assistance students can "thrive when working with harder texts" and cited W. R. Powell's work, which identified an 85% reading accuracy rate as a measure for optimum learning to occur (1968).

To internalize this point, one needs only to harken back to the earlier discussion of ESL, in which every text can be said to be a complex text. Krashen's work on language acquisition, particularly his idea that comprehension precedes production and that production is sequential, supported the impetus to put students in texts above their ability to read independently (Krashen & Terrell, 1983). Just as second language learners

understand more than they can produce, learners in challenging texts can glean something from texts they do not fully understand, keeping in mind that with appropriate scaffolding students will continue to grow into the text.

However, early-reader books advisory is often done by teachers, many of whom are ill-prepared to scaffold early readers. Moats (1994) offered three areas in which early readers struggle based on aspects of difficulty found in the literature: phonological awareness, specifically the ability to discriminate sounds; structural awareness; and understanding morphemes and how they differ in spelling and pronunciation. However, she also suggested that although there is a solid research foundation for instructional practice, teachers either don't have the skills themselves or lack an understanding of how those skills facilitate reading. Moreover, the basic teacher certification standards do not enable teachers to "meet the diverse needs of students who are at risk for reading/writing failure" (Moats, 1994, p. 82).

Moats reported that even literate teachers have an insufficient understanding of spoken and written language to recognize student errors and even less capacity to teach explicitly to a variety of reading problems. She offered an examination of several challenges to teacher knowledge that became evident through her study. "Being able to interpret and respond to student errors, being able to pick the best examples for teaching decoding and spelling, being able to organize and sequence information for instruction, being able to use knowledge of morphology to explain spelling, and being able to integrate the components of language arts instruction," are all critical areas in which lack of knowledge inhibits teacher efficacy (Moats, 1994, p. 98). Teachers of reading should

understand and be able to talk about language in all its facets in order to measure student ability, facilitate student progress, and recommend reading materials.

Conclusion

Historically, the idea that texts have intrinsic qualities, in contrast to the readers themselves, has been informed by quantitative measures such as readability studies and qualitative measures that have their origins in literary, content-area, and language research. In addition to these significant variables that affect learning to read, we must also take the reader-task relationship under consideration. With respect to this last variable, as educators, we must be conscious of what strategies and supports a young reader will be able to avail himself of, given the texts and images in beginning readers. All of these aspects of texts give students the kinds of knowledge necessary for higher level reading such as comprehension and written expression.

Chapter 3

Methodology

Purpose and Importance of Study

First readers are powerful mechanisms in the lives of children. They build literacy skills and frame a child's vision of the world. The following research is an in-depth study of the Theodor Geisel Award collection, focusing on multiple aspects of the books: image, text, and the reader-text relationship. This is the first time this collection has been documented and analyzed. This study provides a valuable framework for publishers, parents, authors, and educators that sheds light not only on what our children are reading, but on what factors are deemed important in early literacy and to what degree these factors are represented in beginning-to-read books. It addresses the following research questions:

Research Question One: To what extent do the Theodor Geisel Award and honor books address foundational reading skills?

Research Question Two: To what extent do the Theodor Geisel Award and honor books measure up against both quantitative and qualitative standards for text complexity in the CCSS?

Research Question Three: To what degree do the Theodore Geisel Award and honor books reflect core literacy values related to the variety of human experience including behavior and gender?

Research Question Four: How can computational linguistic analysis be used to inform knowledge of texts for application in the classroom?

The study undertaken here examined the entire collection of the Theodore Geisel Book Award and Honor winners. What follows is a list of the gold medal winners, and a full list including honor books is available in Appendix C.

2016 Medal Winner

Don't Throw It to Mo!, written by David A. Adler, illustrated by Sam Ricks (Penguin Young Readers Group)

2015 Medal Winner

You Are (Not) Small, written by Anna Kang, illustrated by Christopher Weyant (Two Lions, New York)

2014 Medal Winner

The Watermelon Seed, written and illustrated by Greg Pizzoli (Disney Hyperion Books, an imprint of Disney Book Group)

2013 Medal Winner

Up, Tall and High!, written and illustrated by Ethan Long (G. P. Putnam's Sons, a division of Penguin Young Readers Group)

2012 Medal winner

Tales for Very Picky Eaters, written and illustrated by Josh Schneider (Clarion Books, an imprint of Houghton Mifflin Harcourt Publishing Company)

2011 Medal winner

Bink and Gollie, written by Kate DiCamillo and Alison McGhee, illustrated by Tony Fucile (Candlewick Press)

2010 Medal winner

Benny and Penny in the Big No-No! by Geoffrey Hayes (Toon Books, a division of RAW Junior, LLC)

2009 Medal winner

Are You Ready to Play Outside? by Mo Willems (Hyperion Books for Children, an imprint of Disney Book Group)

2008 Medal winner

There Is a Bird on Your Head by Mo Willems (Hyperion)

2007 Medal winner

Zelda and Ivy: The Runaways by Laura McGee Kvasnosky (Candlewick)

2006 Medal winner

Henry and Mudge and the Great Grandpas written by Cynthia Rylant and illustrated by Suçie Stevenson (Simon & Schuster Books for Young Readers)

Each text was examined through an analysis at the word, sentence, and discourse level and compared against a variety of standards for the development of early reading skills. An array of data mining and statistical software was employed, and narrative inquiry rounded out the research. It is my hope to both define and critique beginning-to-read books and to deepen the discussion of ways in which these kinds of texts can support fundamental reading skills and facilitate the reading proficiency of blossoming readers.

Study Design

According to Krippendorff and Bock (2009), content analysis is “a research technique for making replicable and valid inferences from texts . . . to the contexts of their use.” This idea of inference in context is particularly salient when analyzing the books on the Theodore Geisel Award list. Analysis in this case is used to describe in depth the content of written text, illustrations, and the overall character of the Theodor Geisel Award books.

A mixed methods approach best suits this pursuit and several methodologies, both quantitative and qualitative were employed. This approach allowed a context for the sample as well as an evaluation of the books themselves. This investigation is influenced by a number of quantitative approaches beginning with readability models. Other methodologies draw from text analytics and employ powerful computer analysis to make visible the more technical aspects of text in a large body of work. Specifically, Natural Language Toolkit (NLTK) (Bird, Klein, and Loper, 2009) as well as Lexical Complexity Analyzer (LCA) and Syntactic Complexity Analyzer (L2SCA) (Lu, 2010, Lu, 2012) were used to evaluate the data.

However, while computer modeling can magnify and allow for analysis of detail in the reading process, much of what we understand and appreciate comes from the interaction between

the reader and the text. This is particularly true in the early stages of reading and encountering books. To facilitate a more holistic understanding, Connelly and Clandinin's narrative approach rounded out the analysis (1990).

According to Creswell (1994), phenomenology focuses on “describing what all participants have in common as they experience a phenomenon.” This general approach was adopted but with the substitution of the Theodor Geisel Award texts for “participants” in an attempt to discover their commonalities with respect to the phenomenon of what constitutes a beginning-to-read book. I used narrative inquiry as a means of interpreting the data of these texts. As Connelly and Clandinin wrote in “Stories of Experience and Narrative Inquiry,” teachers' narratives are “metaphors for teaching-learning relationships” (1990). I hope to demonstrate a cyclical analysis through narrative by exploring books that are designed to teach, relating what I learn from them as a reader and educator, and then teaching what I have learned from both the process and the product.

This inquiry culminated in the analysis of the Theodor Geisel Award collection in terms of the four research questions, specifically how these texts promote or impede foundational reading skills, how these early reading books fit into the standards of text complexity written into the Common Core State Standards, how they convey the human experience and finally, how corpus linguistics can contribute to the research and pedagogy of beginning-to-read books.

Sample and Data Sources

Forty-six books were examined. This is the complete collection of gold medal recipients and honor books given the Theodore Geisel Award from its inception in 2004 to the present. The full list can be found in Appendix C. The collection contains both fiction and nonfiction texts designed for early readers in grades Pre-K–2. The collection includes both hardcover and

softbound copies. As mentioned in the study design, an in-depth narrative inquiry into each of the books in the Theodore Geisel Award canon, as well as the researcher's early reading and classroom teaching experience was examined. These "field texts" constitute a context and point of reference for the data analysis.

Structure of Analysis

Examination of the Theodor Geisel Award collection was organized progressively from analysis at the word level to sentence level to text or discourse level. Basic information was collected regarding title, author, illustrator, publisher, series title if applicable, copyright, subjects, and summaries based on Library of Congress data. Surface structures that refer to elements of organization such as chapters, titles, pages, and typography was noted. Analysis at the word and sentence level was applied to issues related to reading foundational skills. The Common Core State Standards were addressed through analysis at the word level, in terms of vocabulary, and analysis at the sentence level in terms of readability measures and qualitative factors such as levels of meaning. Cultural considerations were covered at the discourse level. Results are presented with the reader/text relationship in mind. Commentary on the applicability of computational models to early reading texts occurs throughout the analysis.

Computer analysis at the word, sentence, and discourse level was conducted with the use of NTLK, LCA, and L2SCA, all software programs for text analytics. Several levels of linguistic description were explored; from smallest to largest they are: orthography, morphology, syntax, semantics, pragmatics, and discourse. Natural language processing has an accompanying linguistic annotation for each of these linguistic units. These include, tokenization, part-of-speech tagging, syntactic parsing, semantic analysis, named entity recognition, coreference resolution, sentiment analysis, and discourse analysis (Wilcock, 2009). For the purposes of this

study, computational analysis did not include named entity and co-reference results, and sentiment and discourse analysis was done by hand. The results of each was recorded and discussed individually in chapters 4 and 5.

Analysis of Text at the Word Level

Lexile levels and ATOS levels were assigned to each book as they represent two different readability methodologies, the first based on word frequency and sentence length, the second based on word length, sentence length, and total number of words. Lexical analyses were conducted including gross number of words, net number of words, word occurrences, and sight words from the Natural Language Tool Kit (NLTK) Stoplist (Appendix D). Through examination of phonemic/graphemic elements, orthographic features of individual words, not included on the sight word list, were analyzed and counts of morphemes were conducted. Vocabulary was considered using the CCSS framework based on three levels of words: Tier I referring to everyday words, Tier II words, which are “general academic” words found across all disciplines, and Tier III words, which are “domain specific” or related to a particular content area (Beck, McKeown, & Kucan, 2008).

Analysis of Text at the Sentence Level

One aspect of reading fundamental skills is prosody. Though this generally refers to oral reading, the reader not the texts, advances in computer science have offered a way to examine this with respect to the texts themselves. More specifically, computational linguistics uses syntactic parsing as a way of measuring prosody, so in the absence of other measures, this theoretical framework was used to similarly predict and point out issues of prosody inherent in the text itself. For an example of how this works, consider the word *contract* in the following sentences. The diaphragm will contract when you exhale; and the attorney will prepare the

contract for the employer. The word *contract* occupies two distinct functions: In the first sentence contract acts as a verb, in the second it functions as a noun. Linguistic annotation tells the computer how to “read” the sentence with the proper inflection.

In addition to looking at sentence structure as a proxy for prosody, the analysis included subsequent counts of phrases and clauses as well as notation of figurative language and literary devices. Because rhythm and rhyme are related to prosody, this section included evidence of syllabication (in terms of cadence), rhyme, alliteration and assonance.

Analysis of Text at the Discourse Level

Literature reflects our attitudes about diversity, equity, and other aspects of the human condition as well as what themes and topics we value as a society. These attitudes were measured in terms of the presentation of text and images that reflect ethnicity, gender, age, and character attributes. While analysis of these elements were computer assisted, most were derived through data generated from checklists and subsequent coding. This section also makes note of important literary elements such as point of view, and genre.

While this study does not focus on illustration, it included counts of images in relationship to the text and general attributes of style. Coding was used to analyze images and these data were examined against text features to measure and point out incongruences. Thought was given to mood, character affect, and most importantly, how the visual aspects enhance, change, or influence the reading of the text. Consideration was given to the degree to which the illustrations meet the criteria set down in the awards protocol.

Analysis of the Reader–Text Relationship

Transactional theory is an important aspect of any reading experience, however, in the case of the Theodore Geisel Award, books are not chosen by their intended reading audience

(Rosenblatt, 1978). Moreover, picture books are unique in that author and illustrator are often not one and the same. Both of these phenomena imply an interpretive stance and speak to reader–text considerations. So this aspect of analysis has been taken into specific consideration.

In contrast to the clinical relationship between the books and their selectors, a child’s encounter with early readers is emotional as well as cognitive. Some scholars suggest that authors and illustrators, though not present physically, form a teacher–student relationship (Smith, 1992; Spencer, 1998). In order to fully explore this exchange I borrowed the ideas of relationship and collaboration proposed by Connelly and Clandinin (1990) and through narrative inquiry, explore my own story of early readers as well as my impressions as a reader and an educator as I encounter these picture books today.

Instrumentation

There were a number of instruments designed and modified for this analysis. Early analysis consisted of checklists for a variety of quantitative and qualitative measures. Lexical and syntactic analysis consisted of various word counts, tallying of phonemic units, counts of phrases, recording of sentence structure, and syntactic parsing. These operations were done using a variety of natural language processing software. Additionally, each book was hand coded for genre, theme, character attributes, and figurative language. Statistical analysis of coded aspects were conducted using Excel. Finally, journal and story present the means through which issues of reader–text relationship were explored.

Validity/Reliability/Trustworthiness

Several measures were used to ensure validity. Analysis instruments from several studies were examined and edited for the purposes of this research. In order to increase objectivity, descriptions and subject headings supplied by the Library of Congress were used wherever

possible. Readability data was obtained from credible and vetted websites and databases. Finally, with respect to field texts, a thorough exploration of researcher bias was conducted (Merriam, 2002). Together these measures form a foundation for establishing trustworthiness of the study.

Summary

This section focused on the procedures that were used to prepare an in-depth analysis of the Theodore Geisel Award collection. This examination looked at multiple aspects of the books: image, text and the reader–text relationship in both quantitative and qualitative ways. The methodology is structured to include analysis at the word, sentence, and discourse levels and gives attention to the relationship between teacher, text, and learner. This study was conducted and data were analyzed during the spring 2016 semester.

Chapter 4

Results

This foray into the world of beginning-to-read books has pointed out the wealth of material and potential for analysis in even the simplest texts. Books are exceedingly complex and while this research attempts to provide a comprehensive look at a small purportedly representative collection of texts, as the work progressed more questions came to the fore.

These results are organized according to the methodology framework with exclusion of personal narrative which will inform the discussion in chapter 5. This chapter also includes sections on general attributes and visual analysis. The order in which these results will be reported is: general attributes, word analysis, sentence analysis, discourse analysis and contributions of visual analysis.

General Attributes

Ordinarily one might assume that factors such as author, illustrator, publisher might be simply a mundane litany of attributes, however there are a number of interesting considerations which became apparent when the data were aggregated and sorted.

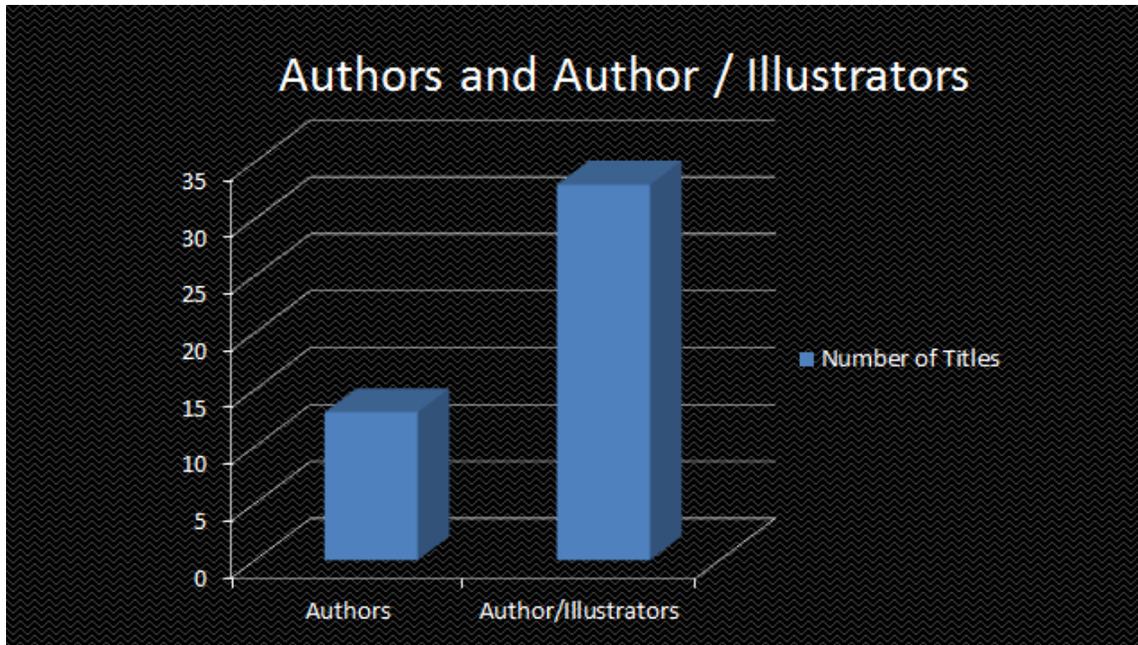


Figure 1 Author/Illustrator (corresponds to Appendix E)

Authors and Illustrators. The entire collection of 46 award and honor books is represented by 36 authors. Thirty-three texts were produced by author illustrators, who, like Theodor Geisel, created both the text and pictures. Several of the titles are either part of a series or have companion books, some of which have also received recognition in the collection.

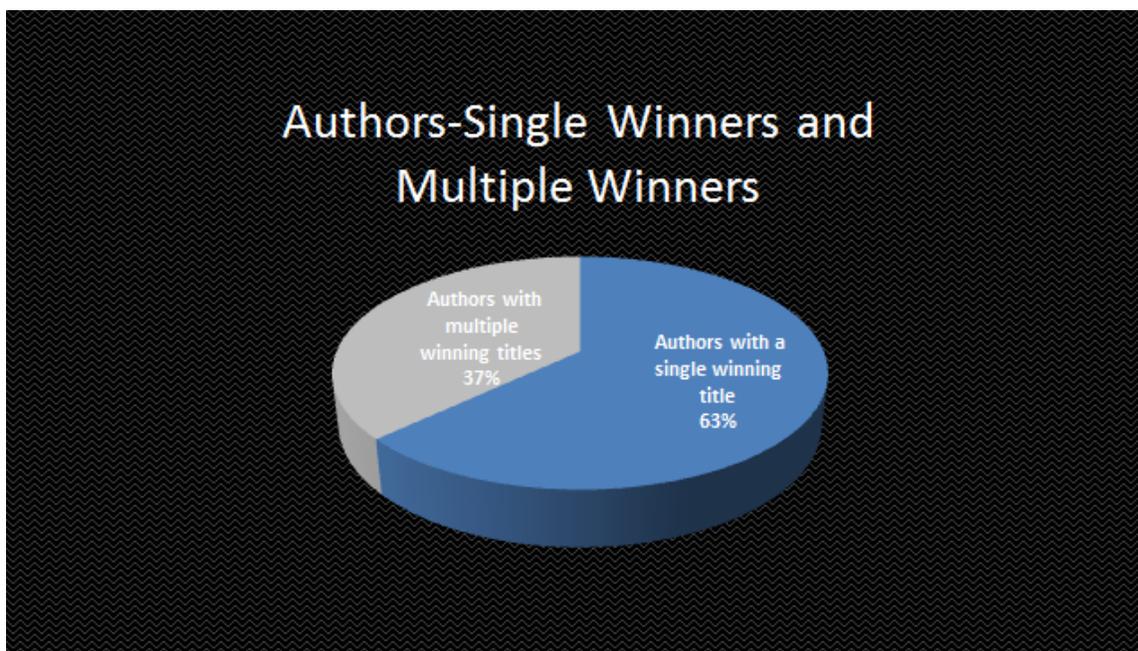


Figure 2 Authors -- Single Winners and Multiple Winners (corresponds to Appendix E)

Figure 2 notes that 6 authors represent 37 percent of the Theodor Geisel Winner and Honor books. Cynthia Rylant, Kate DiCamillo, Laura Seeger, and Kevin Henkes each won twice, either gold or silver medals. Tedd Arnold and Mo Willems won for books in the same series; *The Fly Guy* series won twice and the *Elephant and* series won 7 times.

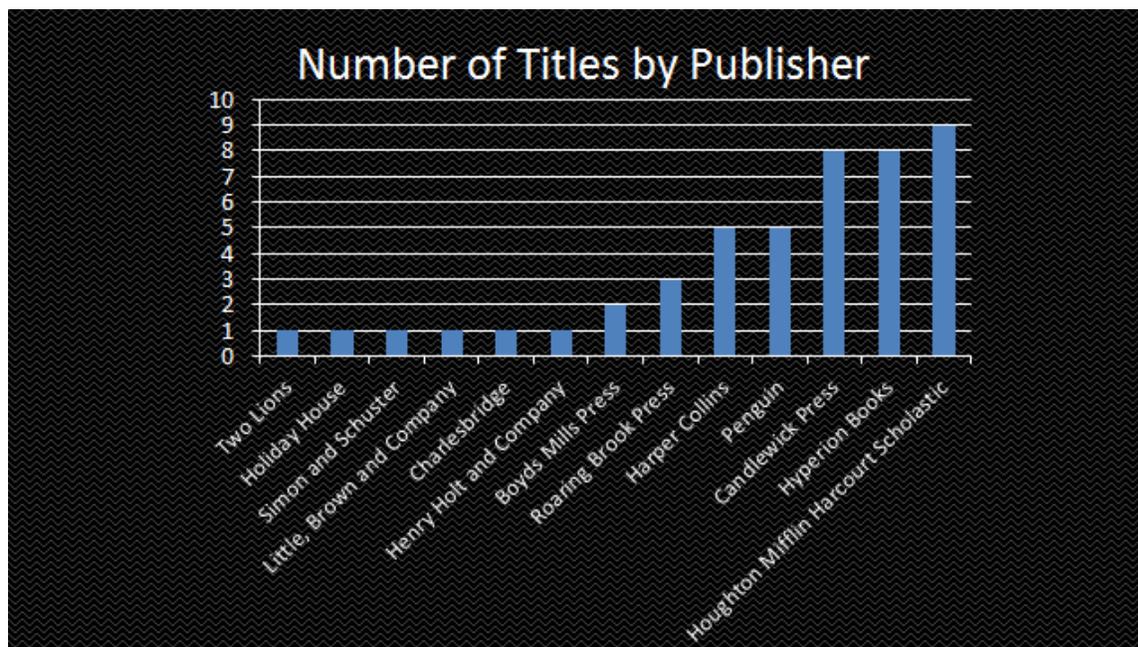


Figure 3 Publishers (corresponds to Appendix E)

The 46 texts in the Theodor Geisel collection were published by 14 different publishing houses, with well over half of the titles having been produced by only 5 publishers. One of the limitations of this work is the inability to see the entire pool of competitors in any year. An interesting question that remains is to what degree small presses are represented in the larger pool and how much the selection process is influenced by availability of catalogs and promotional materials to members of the committee.

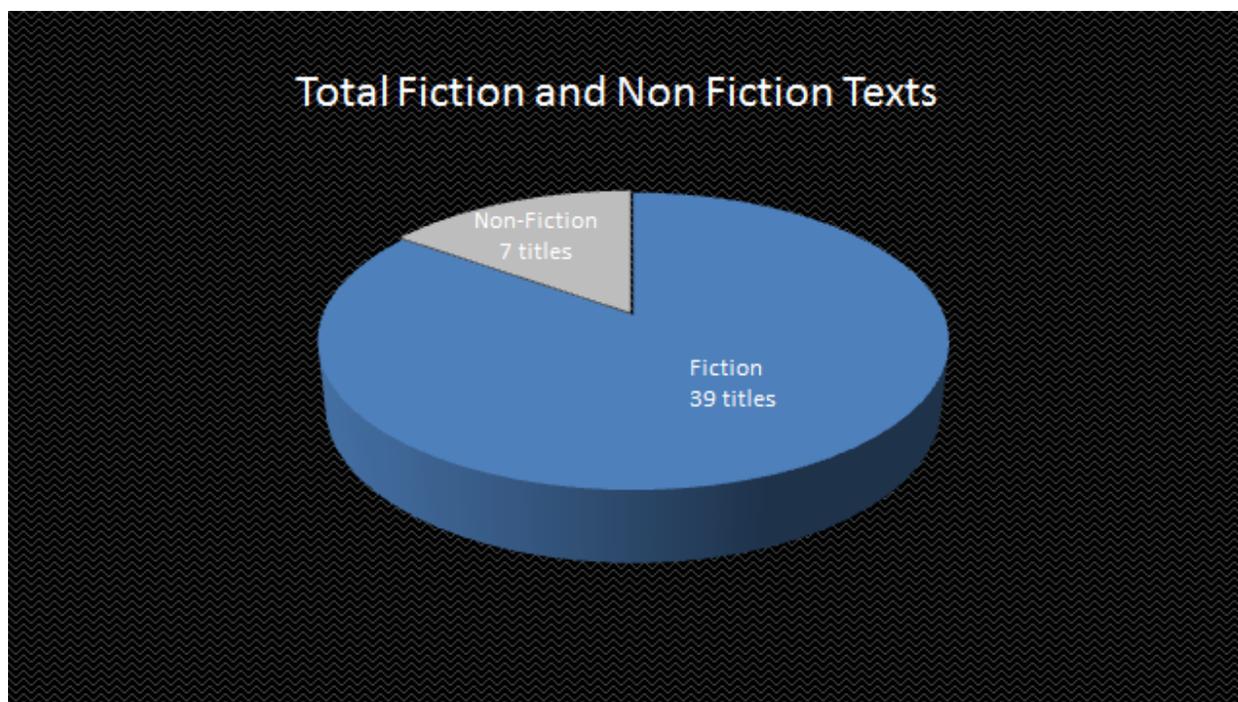


Figure 4 Genre (corresponds to Appendix F)

Genre, Subject, and Theme. Most of the books in this collection are books of fiction. There are 7 titles that are either informational or concept books. The collection addresses 89 discreet subject headings (Appendix P). It is important to note, however, that the subject cataloguing of children's books can include genre and format, in addition to theme. Moreover, it is customary to list the identity of the major characters in the subject field when, in fact, the piece is not *about* the character. For example, the subject headings for Mo Willem's *Elephant and* books all contain elephant and pig, even though they are characters in the story, not the subject of the story. Accounting for these kinds of subject headings, the number of true subject labels diminishes to just 52 which coincide with the themes of these texts (Appendix Z). The value of these headings greatly affects readers' advisory as both librarians and teachers use these to help integrate literature into the curriculum.

Readability

Many of the texts in the Theodor Geisel collection are designated by their publishers as early readers and leveled according to the publisher's standards. (See Appendix Q for each series' criteria.) A number of readability measures were carried out for this research, including Lexile and ATOS (See Appendix R for additional data.)

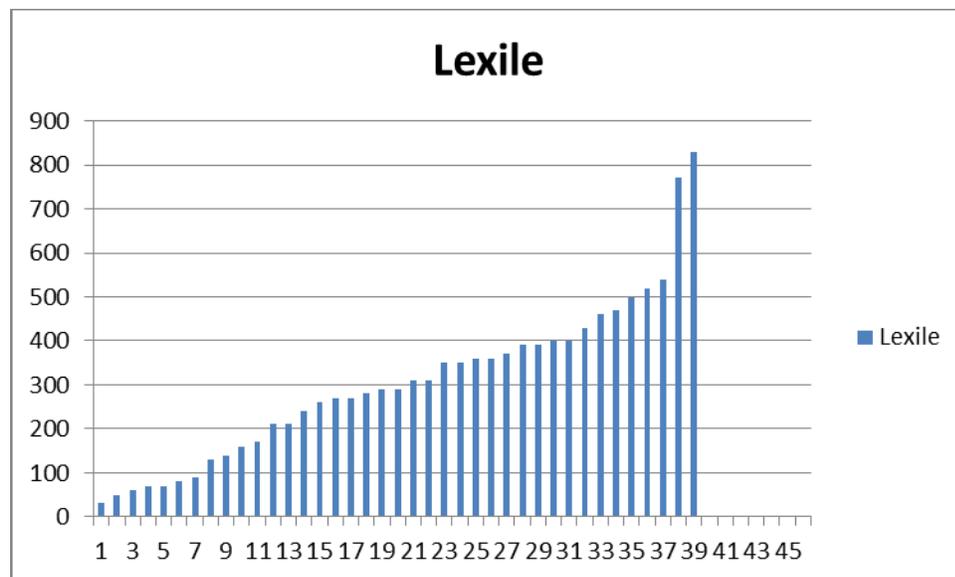


Figure 5a (Readability Appendix G)

Lexile measures are particularly salient as they form the basis of quantitative measures in the CCSS. From the Figure above, it is evident that all books in the collection are not included. This is a result of coding eccentricities in Lexile. The algorithm is not sensitive to books at the lowest levels so they are given a BR code for “Beginning Reader.” Another aspect of Lexile, not well known, is its category analysis (See Appendix H for the full list and definitions). Important to this collection are the designations AD, adult directed, GN, graphic novel, and NP, non-prose. Within this collection, Lexile highlights 6 books as BR, 8 as AD, 3 as GN, and 1 as NP. However, a hand analysis notes that there are 6 titles written entirely in verse; these are *Jazz Baby*, *Move Over Rover*, *Vulture View*, *Pete the Cat*, *See Me Run*, and *One Boy* in addition to *First the Egg*. There are two texts, both concept books, which consist of a single sentence; this

skews the results. After inserting periods at natural breaks, *One Boy* goes from a Lexile score of 1450 to 130 and *First the Egg* goes from 1210 to 470.

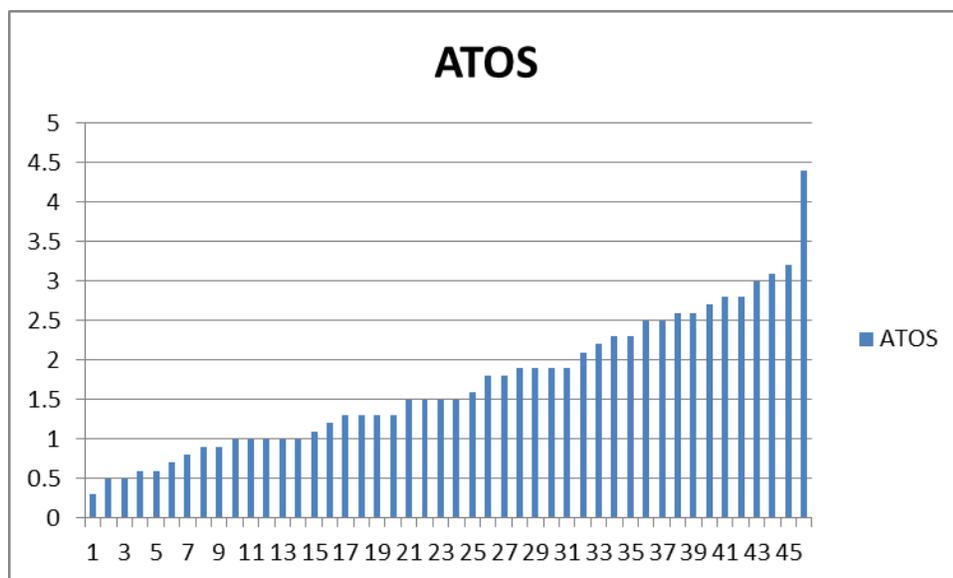


Figure 5b (Readability Appendix G)

ATOS analysis is presented here because of its wide popularity across the nation's schools. The ATOS Analyzer at Renaissance Learning was able to level every book and because ATOS is reported by year and month; for example, 2.3 coincides with second grade third month. The range for this collection is .5 or 5th month of kindergarten through 4.4, fourth month of fourth grade. The two texts with the highest reading levels are both non-fiction, *Wolfsnail*, and *First the Egg*.

Word Analysis

Fry, Chall, Spache, even Theodor Geisel himself believed in the importance of controlled vocabulary and repetition. This section highlights lexical variation throughout the collection. With a range of 54 to 2230 words per text, the average number of words in each of the Theodor Geisel Award and Honor winners is 517 with a median of 279, but these numbers do not adequately address the influence and complexity of words in the collection. Each book was

examined for a variety of characteristics of words including: total counts, type-token ratios, high frequency words, and parts of speech.

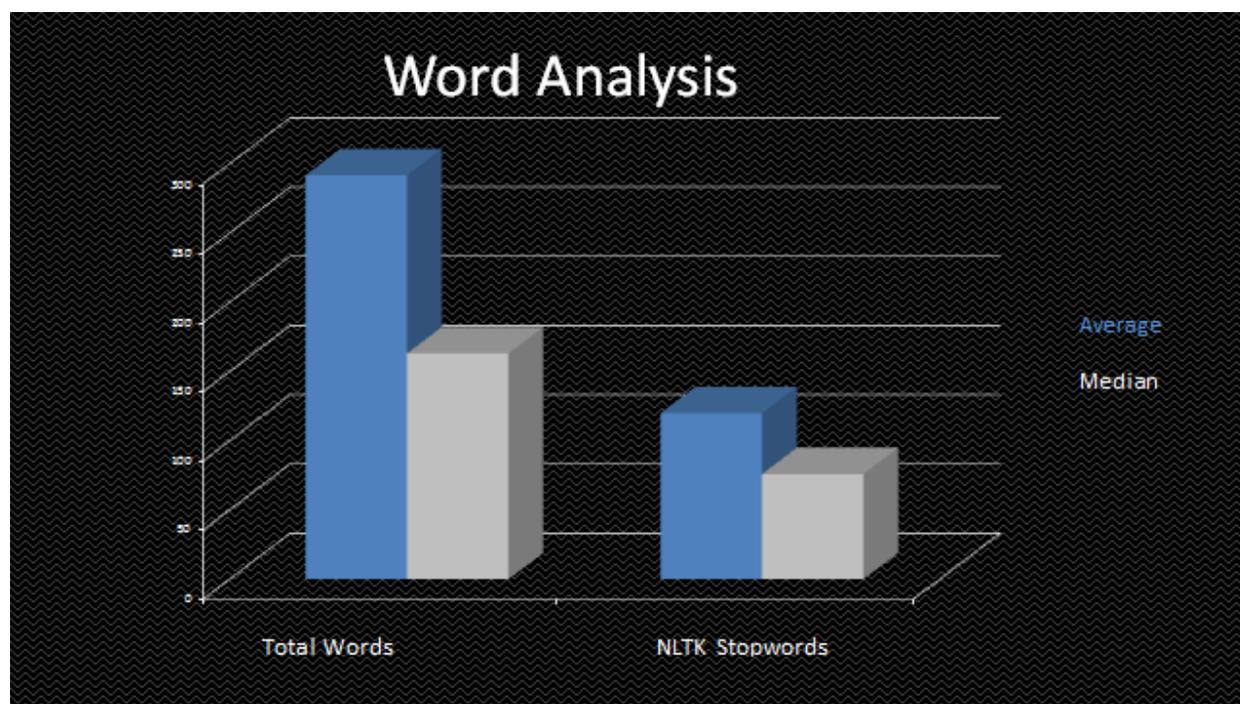


Figure 6 (Word Analysis Appendix I)

The entire corpus of the Theodor Geisel Award and Honor winners was measured against the Natural Language Tool Kit (NLTK) Stoplist of words that function mainly as grammar tools in a sentence (e.g. *am, any, and, because, by*. See Appendix S for full list). This list in many ways mirrors Fry's list of the 100 most frequently used words, with 70 words in common. The primary difference is that the NLTK Stoplist includes very few lexical words (nouns and verbs), whereas Fry includes some simple verbs and commonly occurring nouns.

Figure 6 illustrates how much of these texts consists of lexical words or content words (mainly nouns and verbs) versus structural words, or those that function as grammar tools, many of which are high frequency words. This has significance to pedagogy both in teaching vocabulary and spelling since high frequency words are generally taught as sight words, because

they often have unusual or difficult orthography (See Appendix I for more data on ratios of lexical words to high frequency words).

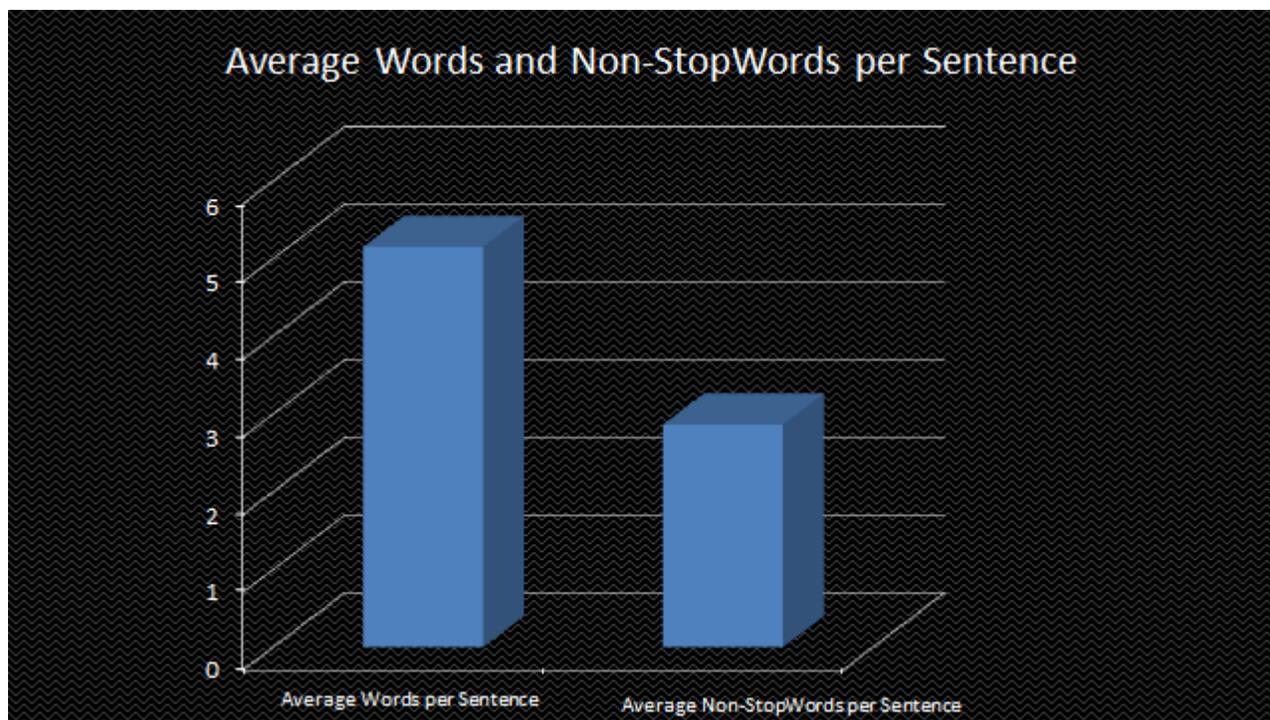


Figure 7 (Sentence Analysis Appendix J)

Another way of looking at words is number of words in a sentence as sentence length is a common measure in readability scales. The data show that the average sentence contains 5.2 words, and of these 2.9 are content words, i.e., not among the 125 high frequency words on the NLTK Stoplist. The medians are not significantly different at 5.1 words per sentence and 2.3 lexical words per sentence.

Parts of Speech. Tagging parts of speech is an integral step to determining complexity of text, but even simple counts can convey a great deal of information. Figure 8 shows averages and means for nouns, verbs, pronouns, adjectives, and adverbs.

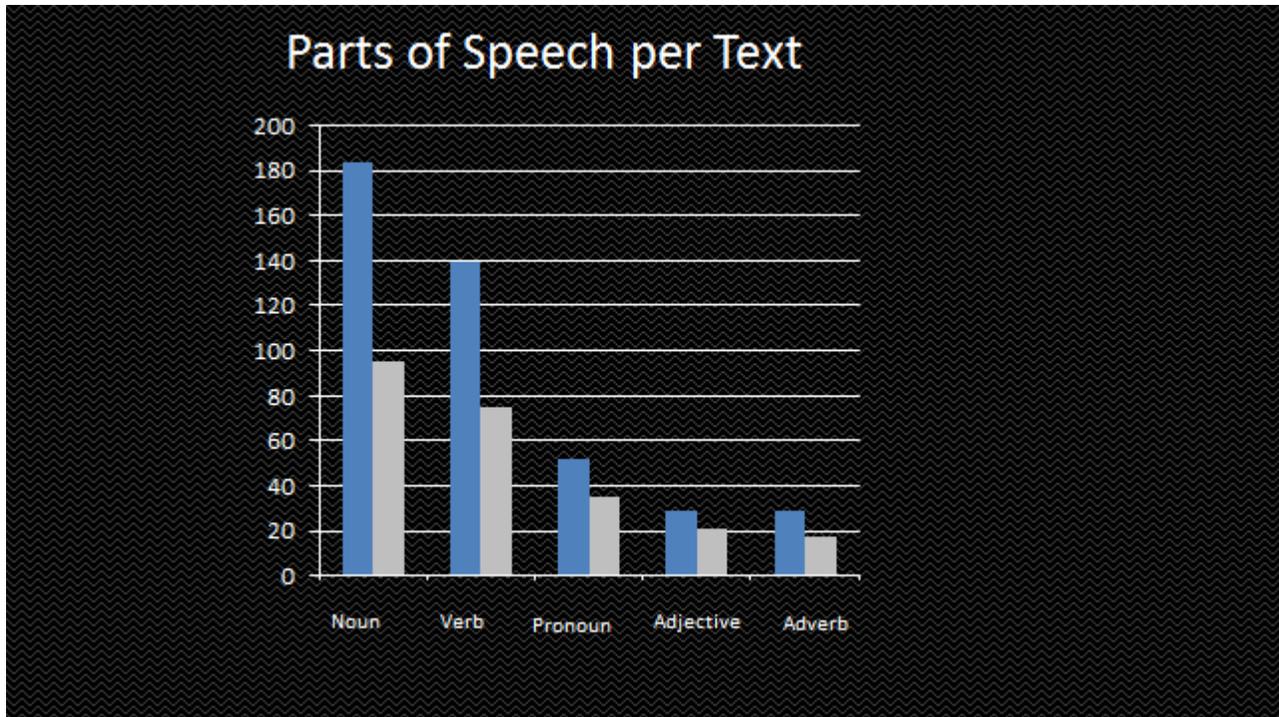


Figure 8(Parts of Speech Appendix K)

In the Theodor Geisel Collection, the average number of parts of speech are as follows: nouns—184, verbs—140, pronouns—52, adjectives – 29, and adverbs – 29. In the case of the Theodor Geisel collection, it makes sense that books for young readers would have a significant number of nouns and verbs to make up the core of a sentence and fewer modifiers which add to complexity. Since prepositions, conjunctions, determiners, and interjections are not content words, they are not shown in Figure 8.

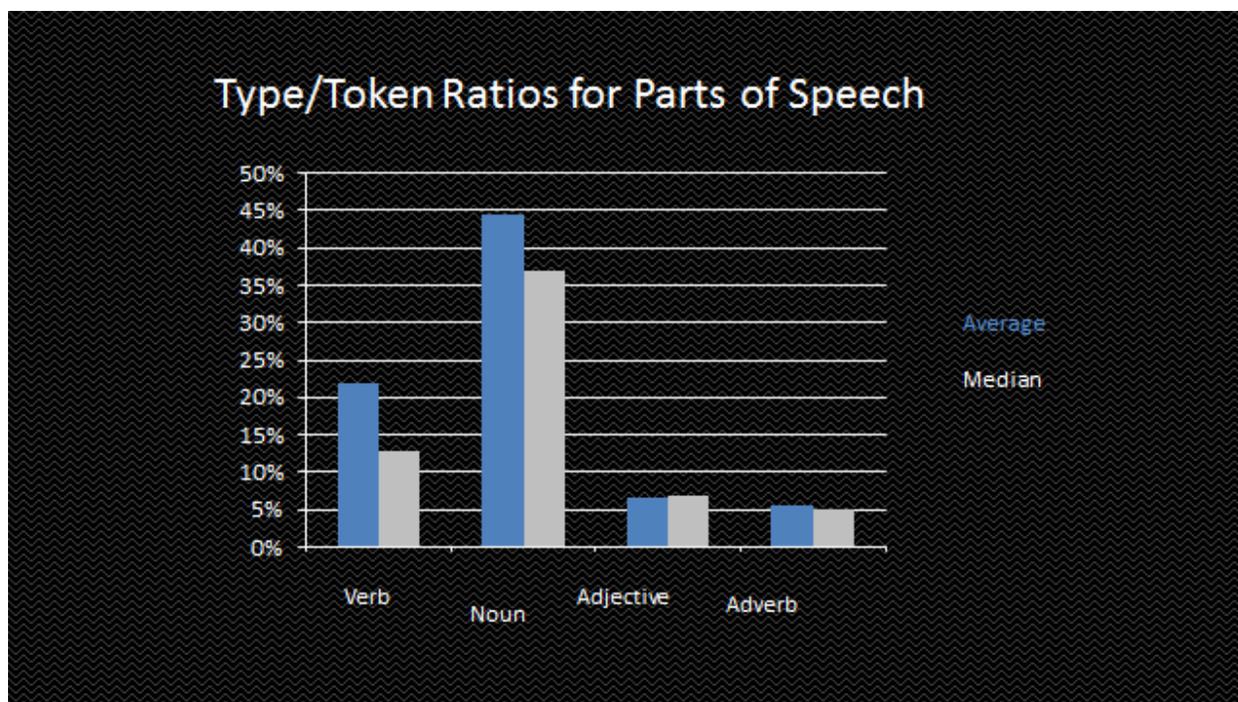


Figure 9 (Lexical Variation Appendix L)

Early reading researchers such as Thorndike (1944) and Dolch (1948) pointed out the significance of word frequency in determining text difficulty. Today, tokenization of texts through computational linguistics reveals the validity of this idea. This Figure shows the type-token ratios of the most significant parts of speech. For example, this measures the net number of verbs in a text against the total number of verbs in the text. This measure indicates the amount of repetition of a particular part of speech. For example, *Supertruck* has a verb type-token ratio of .06 which indicates that there is a high degree of variation of verbs within the text or put another way, there is a very low degree of repetition of verbs in the text.

A close look at the data suggests that though modifiers are used relatively infrequently, they tend to be repeated often, while nouns and verbs make fewer repeated appearances in a single text. Longer texts show a higher type-token ratio because the more words a work contains, the more likely it is to have multiple occurrences of the same word.

Mouse and Mole had the highest type-token ratio for verbs at 66 percent but this text is filled with parallel sentences such as “Mouse looked to her left. Mole looked to his right.” Additionally, the text repeats character actions throughout all the chapters, “Mole rubbed his snout. Mouse twirled her tail.” At the other end of the spectrum are many of the concept books which have a very small type-token ratio for verbs because there is little “action” that takes place.

Orthography. The NLTK Stopword list was used to filter out high frequency words in the Theodor Geisel Award and Honor books. Those words which do not appear on the Stoplist are included (Appendix M) and are examined for complexity both with respect to decoding and as a measure of the CCSS.

Sample of words with difficult orthography

Ghost	Haiku	Hover
Lasagna	Laugh	Mustache
Panicked	Paused	Skiwies
Spaghetti	Stomach	ventured

Figure 10 (Complete Lexical Word List Appendix M)

One aspect of complicated orthography that children encounter in these texts is related to the presence of onomatopoeia as it often does not follow spelling patterns. Onomatopoeia is a common attribute of Graphic Novels which use many words such as *waaah*, *whapp*, and *grrr*.

Sentence Analysis

The examination of prosody and its relationship to fluency is an emerging field in early reading research. As Schreiber (1991) suggests, without an understanding of syntax, prosody suffers. So in order to address this, sentences were analyzed for several measures including counts, punctuation, syntax, and prosody. (For sentence length see Word Analysis).

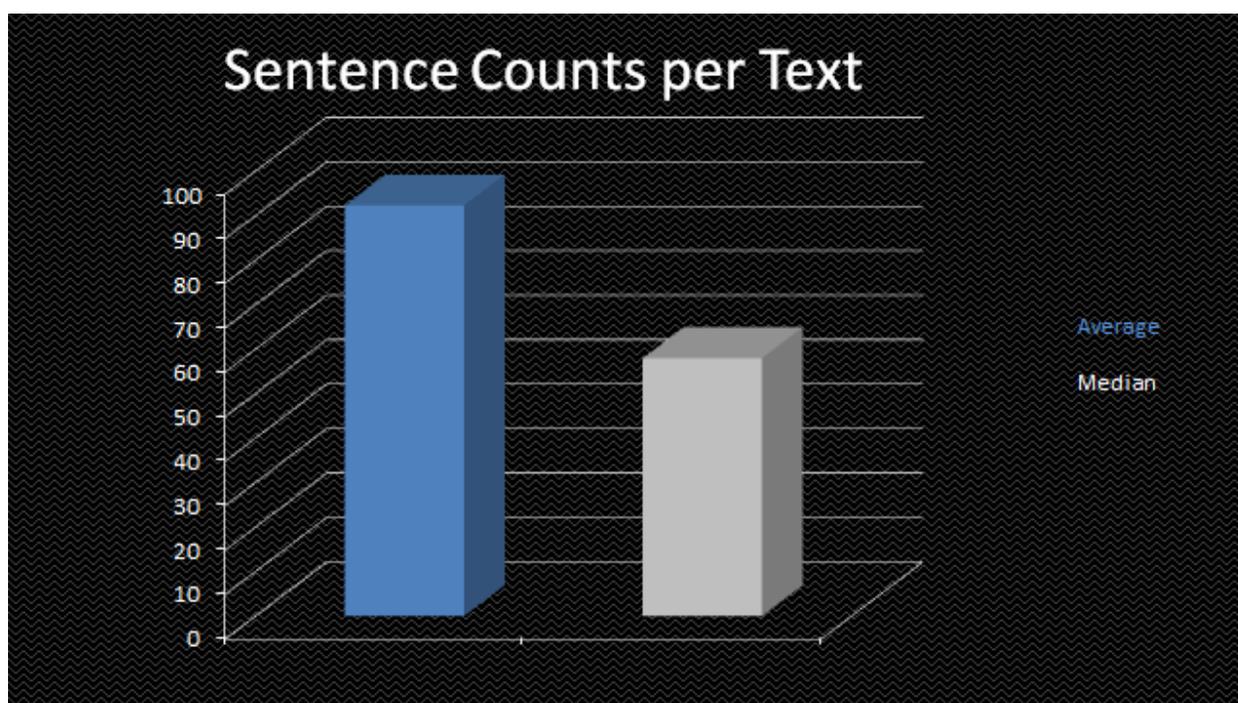


Figure 11 (Sentence and Punctuation Analysis Appendix N)

Counts. The average number of sentences per text among the aggregated collection is 92.3 with a median of 58. However, the range is quite large, spreading from just 12 sentences in the shortest text, *Not a Box*, to 350 in the longest text, *Mercy Watson Goes for a Drive*. The average sentence length is 5.2 words

Punctuation. This section looks at the punctuation in whole texts. Included in this section are end punctuation and use of quotations.

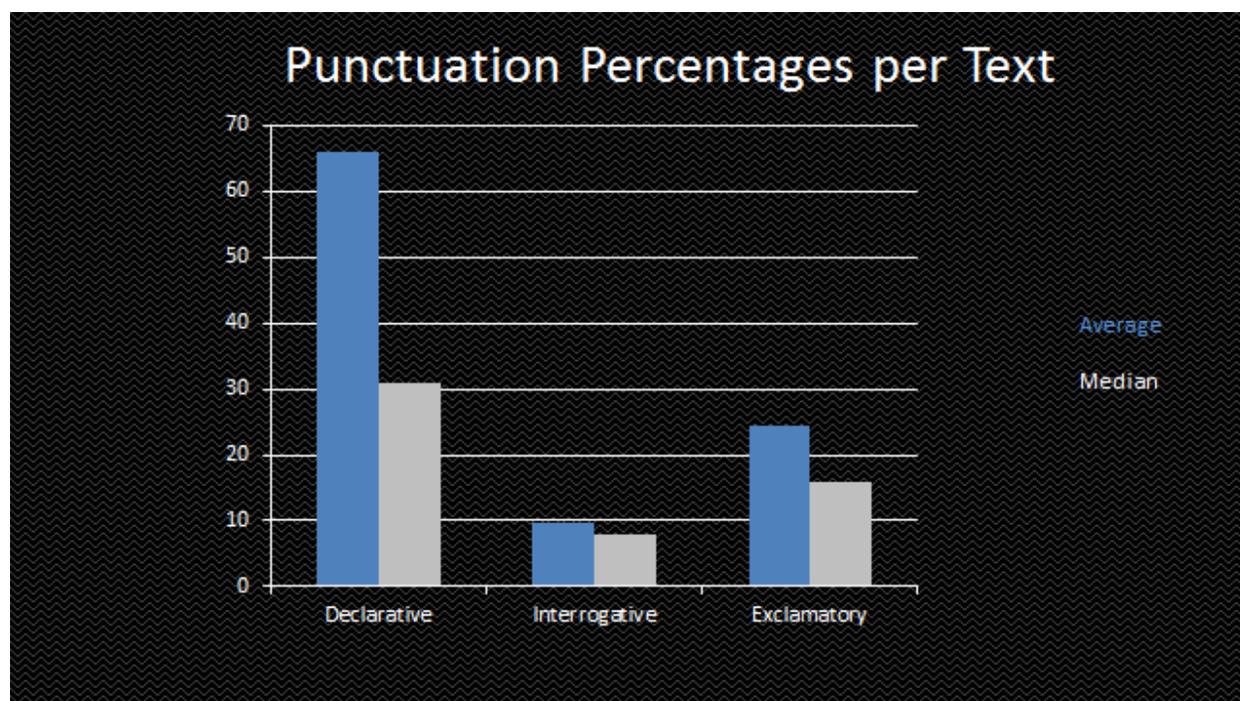


Figure 12 (Sentences and Punctuation Appendix N)

The Figure above contains the means and medians of three major end punctuation marks: period, question mark, and exclamation point. No differentiation was made between imperative sentences and declarative sentences whether represented by a period or an exclamation point. In aggregate, declarative and imperative sentences averaged 66 percent per text, interrogative, 10 percent per text, and exclamatory, 25 percent per text, with ranges of 0-396, 0-42, and 0-121 respectively.

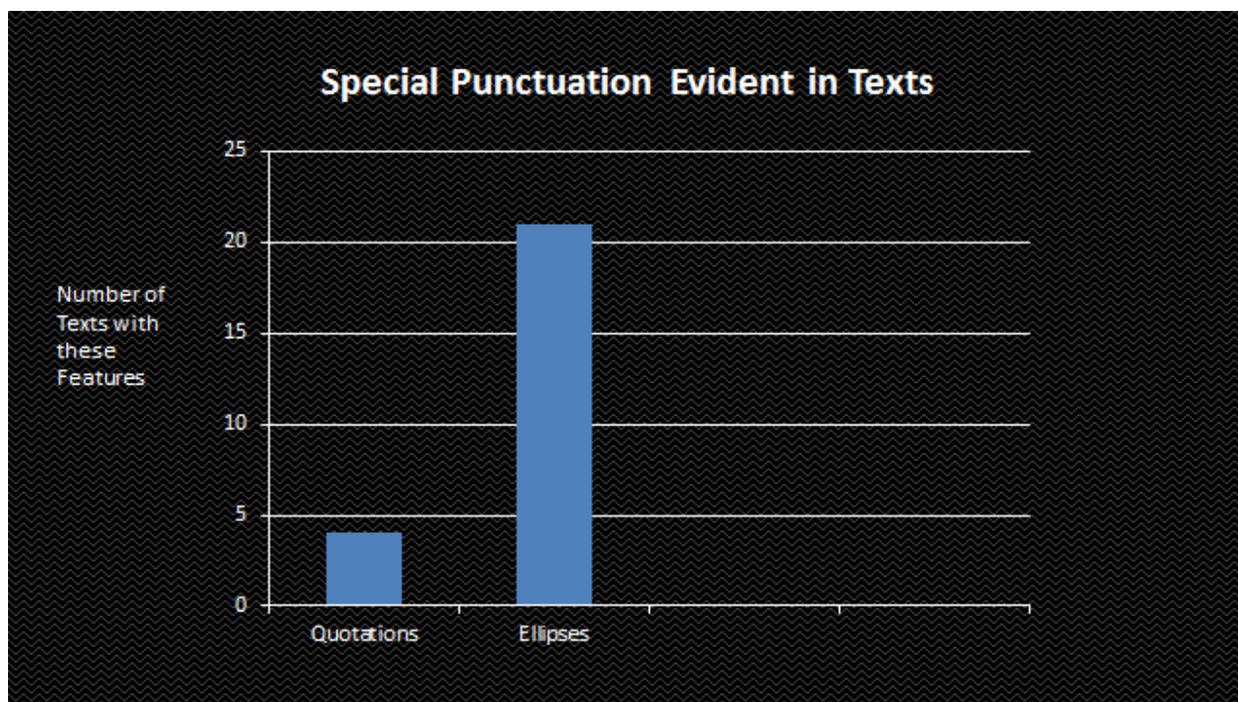


Figure 13 (Punctuation Appendix O)

In texts which used dialogue, the amount of spoken text is quite high, keeping in mind that each quotation can include more than one sentence. However, assuming each quotation is only one sentence, quotations make up better than 40 percent of those texts which use dialogue. Graphic novels are excluded from this count as speech and thought bubbles rather than quotation marks are used to indicate dialogue. Twenty-one texts use ellipses, the counts of which are greatest in graphic novels.

Syntax. Syntactic analysis was done on both sentences and t-units; this distinction is very important. In computational linguistics a sentence is represented by its frame, i.e., it begins with a capital letter and ends with a final punctuating mark. However, a t-unit, as defined earlier, is a grammatically complete thought with a subject and a predicate. The differences between these two ways of looking at syntax are significant in the Theodor Geisel Collection. Both were examined for verb phrases, dependent clauses, coordinate phrases and complex nominals.

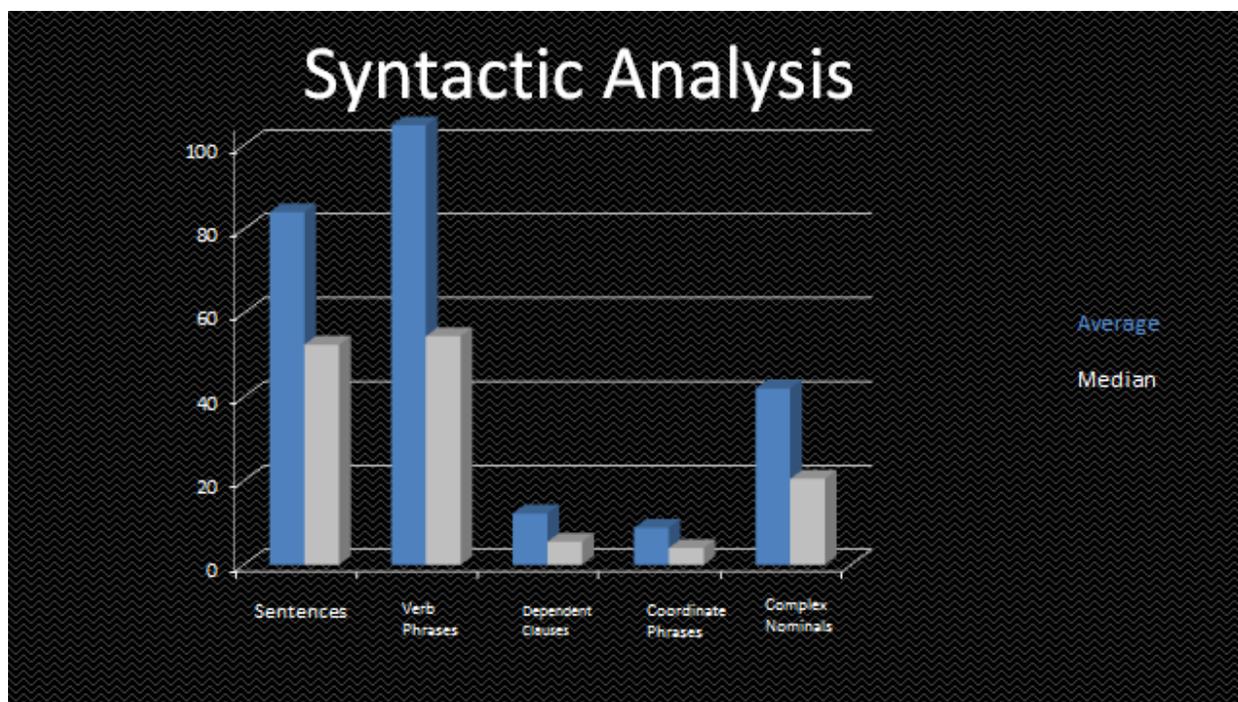


Figure 14 (Syntactic Analysis Appendix P)

As is evident from the Figure above, medians are a much more appropriate measure for syntax. The median number of sentences per book is 53. It is worth noting the large discrepancy in the averages. Keeping in mind the definition of a sentence, this suggests that there might be sentences with multiple verb phrases, e.g., *Today is a day to sit in the shade or do nothing*, or alternatively, verb phrases, without subjects e.g., *Halt!* There are many selections in which sentences begin with *but* or *and*, followed by a verb phrase.

Dependent clauses do not make up a significant percentage of the sentences in the entire group of books, but there seems to be a higher correlation between the number of dependent clauses and longer texts. A final factor considered in this analysis is the number of complex nominals, e.g., *the troll living in the basement*, in each of these texts, which ranges from 0 to 168 with an average of 42 percent and a median of 21 percent of texts which contain complex nominals.

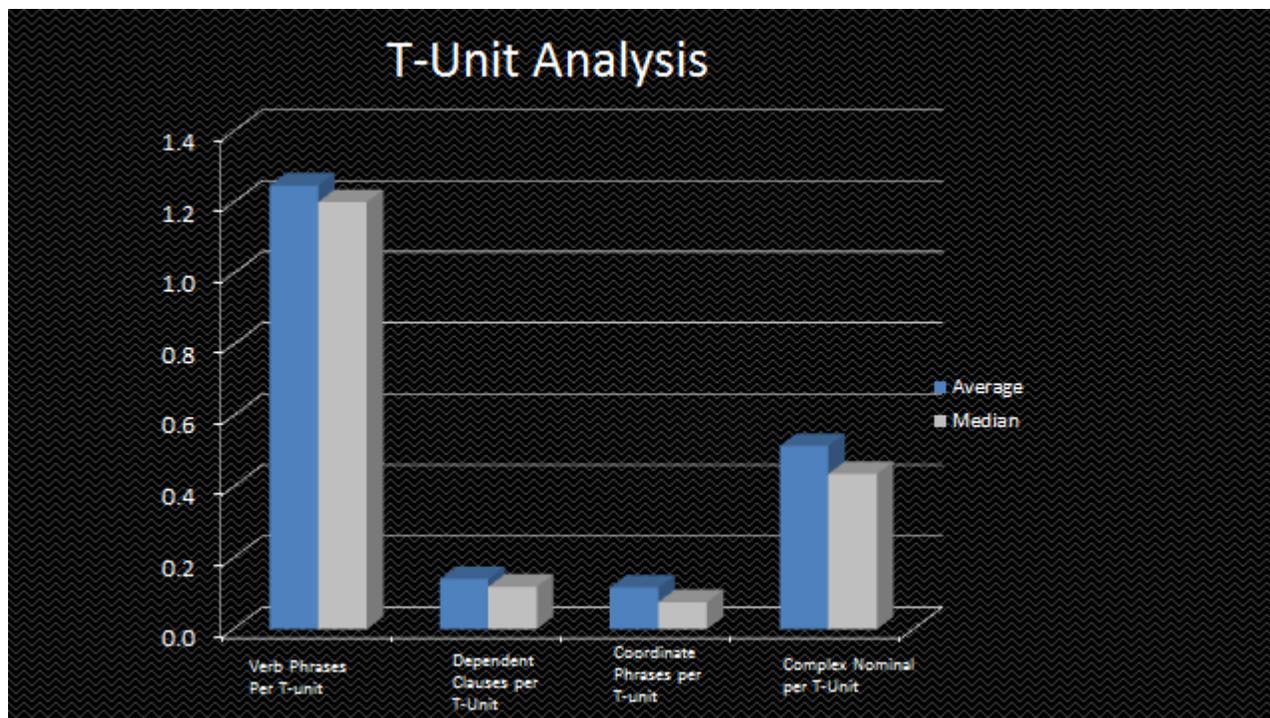


Figure 15 (T-unit Analysis Appendix Q)

The average number of t-units per sentence is .91. Cases in which the t-unit count is less than one are explained by lower level and heavily graphic texts such as Mo Willem's, *A Big Guy Took My Ball*. The graphic novels in the collection have the lowest ratios of t-units per sentence. *Benny and Penny and the Big No-No* contains .82 t-units per sentence, *Stinky* contains .8 t-units per sentence and *Little Mouse Gets Ready* has a t-unit ratio of .64 t-units to a single sentence. A good example of how this is possible is illustrated in the phrases like "Dumb girl," and "Birds and butterflies," which are written as sentences but are not complete t-units.

Relative Complexity of T-Units

Title	T-units	Verb Phrases Per T-unit	Dependent Clauses per T-unit	Coordinate Phrases per T-unit	Complex Nominals per T-unit
A pig a fox and a box	74	1.081	0.054	0.014	0.284
Chicken said Cluck!	72	1.028	0.028	0.000	0.069
Bink and Gollie	95	2.042	0.453	0.116	1.453
Tales for Very Picky Eaters	91	2.143	0.582	0.176	1.593

Figure 16 (T-unit Analysis Appendix Q)

The usefulness of t-unit analysis can be seen in the contrast between complex books such as *Bink and Gollie* and *Tales for Very Picky Eaters* and books with simpler syntax such as *A Pig, A Fox, and a Box*, and *Chicken said "Cluck!"* The former have 2 to 1 ratios of verb phrases per t-unit, with nearly half of all t-units containing dependent clauses.

Prosody. As discussed earlier, prosody is a significant factor in building fluency; however, there is no real measure for analyzing prosody within a text. There are, however, a number of markers that can give clues to the constraints and supports available to young readers. These include sound devices, typography, and punctuation.

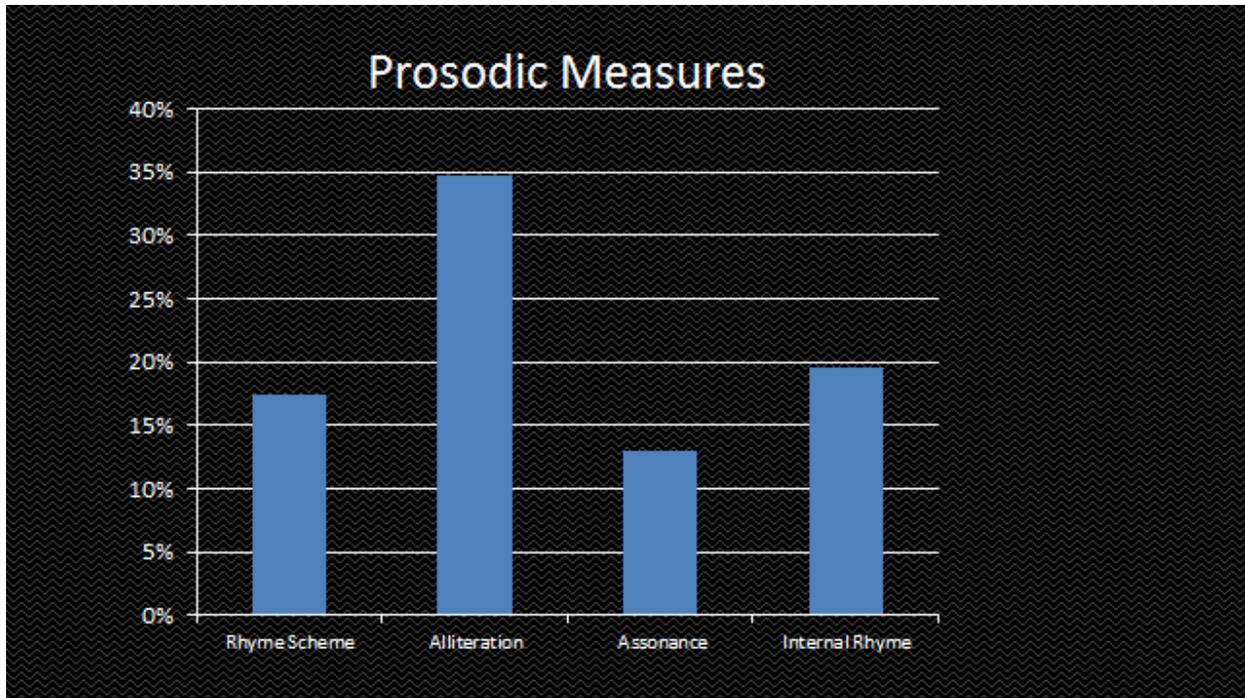


Figure 17 (Prosodic Measures Appendix R)

Sound Devices. There are a number of sound devices evident in the texts of the Theodor Geisel Collection. As seen in Figure 17, alliteration is present to a greater or lesser degree in 35 percent of the texts. Rhyme is present in 22 percent of the texts and consists primarily of couplets with the exception of *Move Over Rover*, which uses an AB rhyme scheme and *Jazz Baby* which uses an ABCB pattern. While many texts do not use formal verse, internal rhyme and assonance also are common sound devices evident in 19 percent and 13 percent of texts respectively.

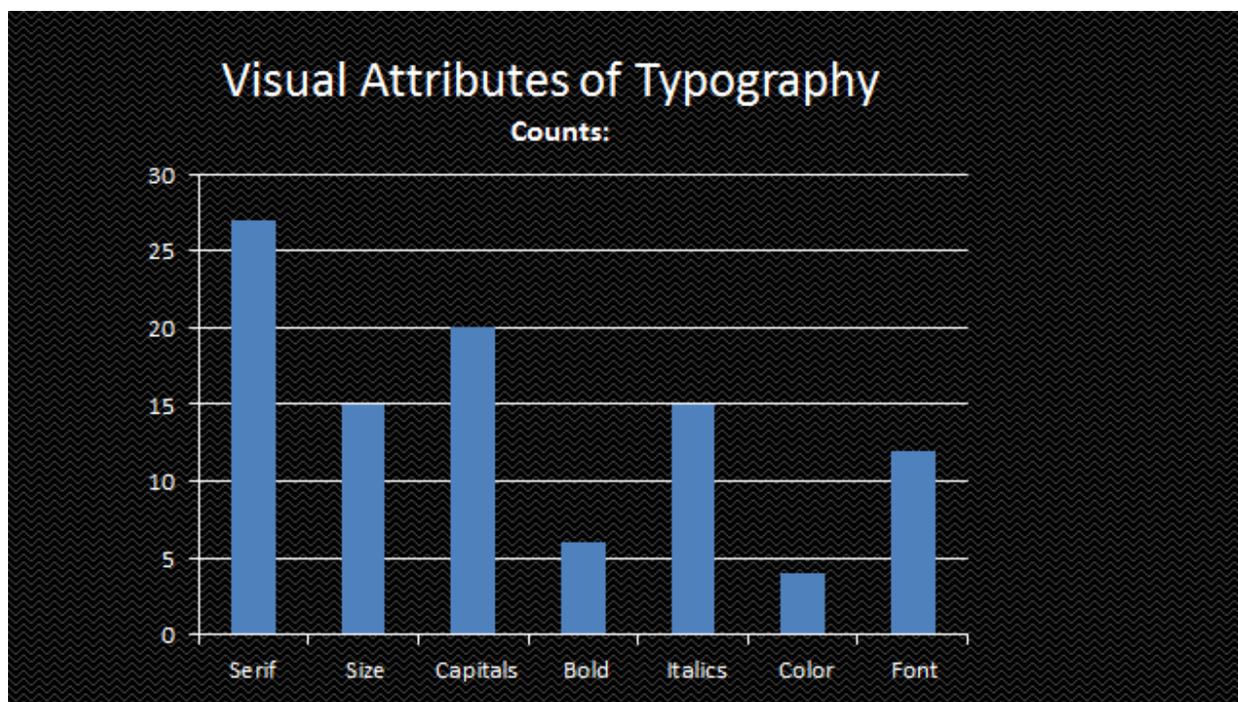


Figure 18 (Visual Aspects of Type Appendix S)

Typography. The English language is fairly limited in its ability to promote prosody. Beyond question marks and exclamation points there is little indication of how a sentence should be read and almost no way to indicate the operative word in a sentence. Enhanced typeface is extremely important in these early texts. Thirty-four of the 46 books, or 74 percent, use some kind of altered type to convey meaning.

Altering the type size of individual words, presenting words in all capital letters, and inserting italics in a sentence, are the most prevalent uses of typography to convey prosody. With respect to font, 12 of the texts intersperse either different typestyles or hand lettering. Colored fonts are used in 4 books; however, in *I Want My Hat Back*, it is used to convey a change of character in a spoken dialogue, as this text does not use quotation marks.

Discourse

The CCSS make qualitative measures of text complexity a priority in analyzing texts. As Carrell (1987) emphasized, written language has the propensity to be indirect and carry aesthetic

meaning, in addition to semantic meaning. The following results point out the relevance of this complexity to the assessment of the Theodor Geisel Collection.

Literary Devices. Seven literary devices were counted in the examination of the Theodor Geisel books: simile, metaphor, personification, onomatopoeia, hyperbole, idiom, and pun. Each of these was present in 2 or more texts in the collection, with personification and onomatopoeia being present in more than 60 percent of the collection.

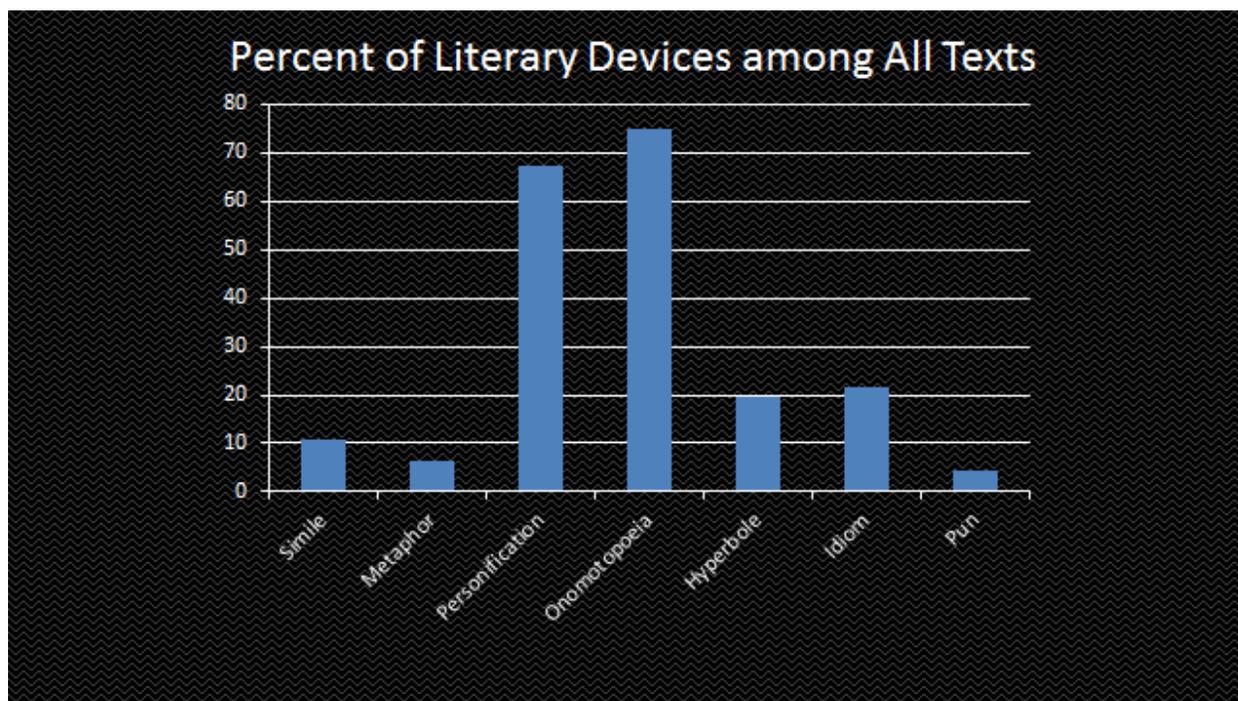


Figure 19 (Literary Devices Appendix T)

Simile and Metaphor. There is very little use of simile and metaphor among the Theodor Geisel collection. This seems appropriate as comparison, whether direct or indirect, requires a higher level of comprehension. It is worth noting, however, that in the instance of similes, there is often crossover with hyperbole. *Amanda Pig and the Really Hot Day* provides many examples: “I’m as hot as a fried egg, I’m as hot as toast. I’m as hot as oatmeal.”

Personification. Personification is widely used in texts for young children. This is mostly obvious as 58.6 percent of the texts use animal characters in lieu of human characters. (See

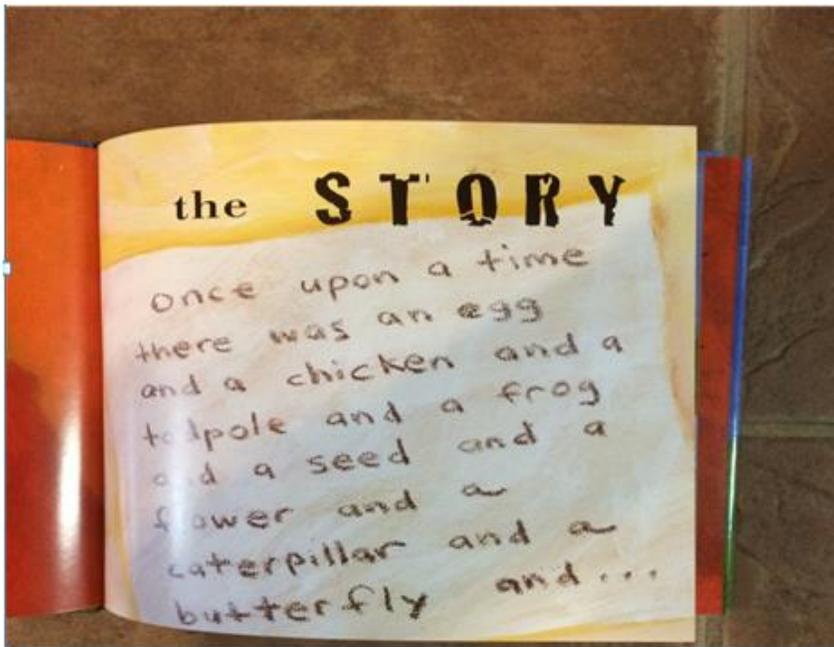
discussion of ethnicity) However, there are instances in which personification is more subtle as in the following reference to “oatmeal” in *Tales for Very Picky Eaters*: “Growing oatmeal has quite an appetite. It will start eating the other foods.”

Idiom. Use of idiom was somewhat surprising as 28 percent of the texts used this device. Examples include, “grey matter,” “from the boots up,” “eureka,” “rats,” “plan b,” and “bug.” While these enrich the reading, they can present difficulties in comprehension for children who are still in a concrete phase. Similarly, though puns are less common, they require a sophistication beyond the capabilities of many early readers. Puns may not impede comprehension, but if a child has not achieved some fluency, then the magic of the language is lost because of the effort spent first on decoding, second on comprehension, and lastly on word play. Examples include the contrast of *pets* and *pests*, in *Hi Fly Guy*, and *forward-looking/look backwards* in *Mercy Watson Goes for a Drive*.

Onomatopoeia. Onomatopoeia is the most common literary device among the winners and honor recipients of the Theodor Geisel Award. It is evident in 65 percent of the texts in this collection. Onomatopoeia is an engaging mechanism although it also carries with it a problem for poor decoders. While the incorrect decoding of onomatopoeia will have little effect on comprehension, it is possible that for less skilled readers the orthography of these words can be problematic.

Register. Among these texts there are 7 non-fiction or concept books. Because many of these texts have animals as central characters, it is sometimes difficult to distinguish between fiction and non-fiction. For this reason, the designation for this element was dependent upon whether or not the books had all 5 characteristics of plot: exposition, rising action, climax, falling action, and resolution.

In addition to the peculiarities of identifying genre, the Theodor Geisel collection comes with a number of other eccentricities with respect to register. Elements commonly found in non-fiction are a part of many of the texts. For example, eleven titles have a formal table of contents while none of the informational or concept books include this feature. Even more surprising is the use of literary language in non-fiction selections. For example, *First the Egg*, an introduction to developmental biology, contains the following literary “Once upon a time” opening.



In a final note related to register, the increase in popularity of the graphic novel and its presence in the collection warrants acknowledgement here. To date, the Geisel collection includes 3 graphic novels with other books in the collection adopting graphic elements such as speech balloons.

Human Experience. Because text conveys cultural and social standards, the Theodor Geisel Award and Honor Collection was examined for representations of the human experience which included gender, ethnicity, and age.

Gender. This section includes character counts as well as references to how gender is presented within the collection.

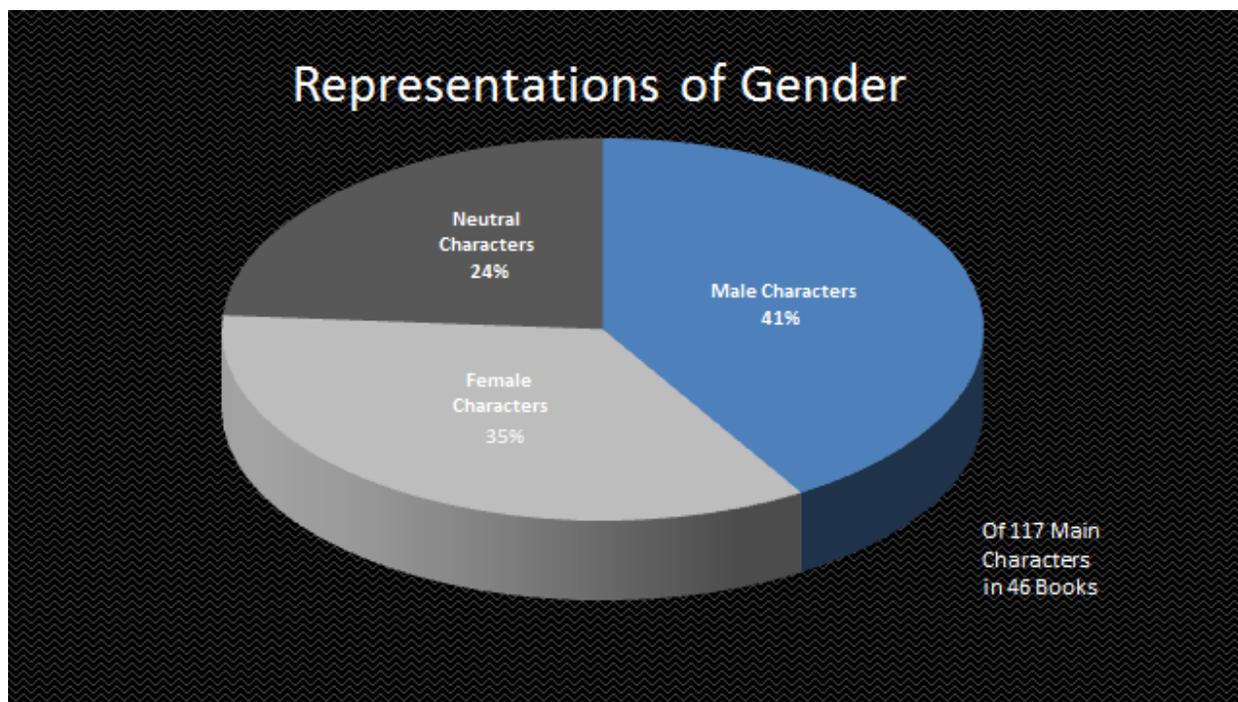


Figure 20 (Gender Appendix U)

Among the 110 major characters of the Collection, 45 were male, 38 female, and 26 were neutral gender (i.e., animals or objects with no identifying gender markers in text or image). In the exploration of the fiction titles, pronouns as well as gendered names provided evidence of gender. The seven non-fiction titles are gender neutral with the exception of *One Boy*. Three of the fiction selections had no reference to gender in the text, but demonstrated gender in images, most of which were stereotypic: for example, girls wore hair bows or dresses.

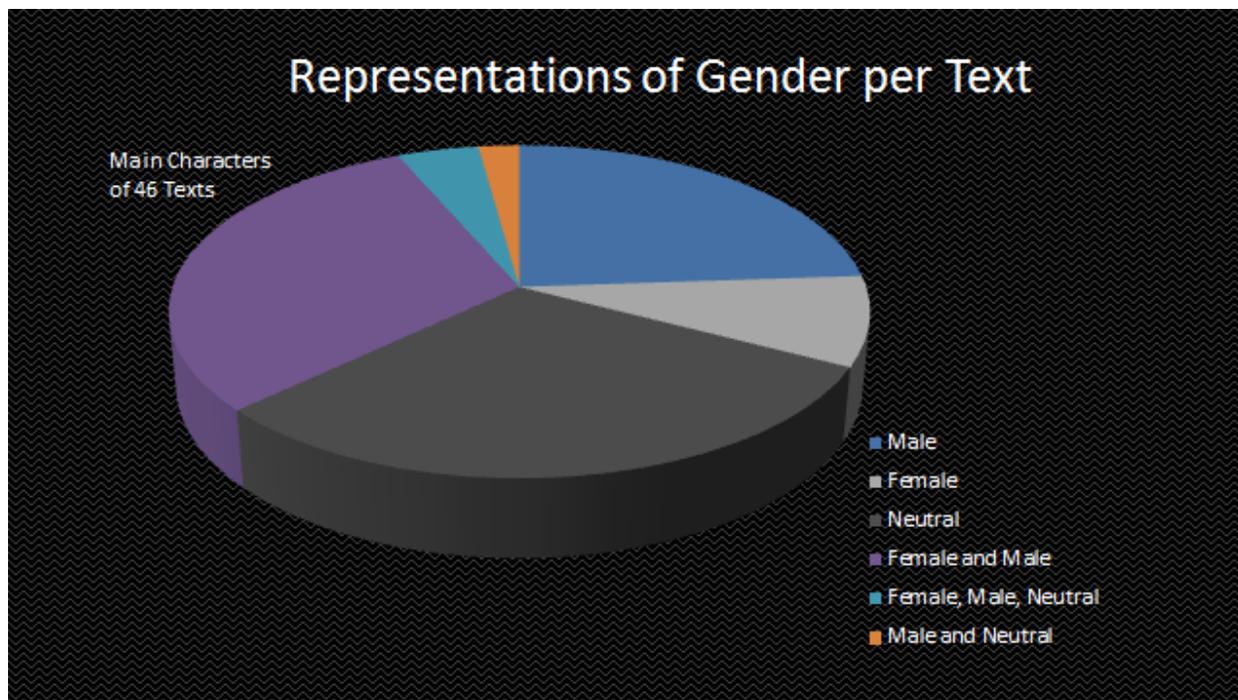


Figure 21 (Gender Appendix U)

Because many of the Theodor Geisel texts had more than one major character the 46 texts were divided into six categories, books with: male protagonists, female protagonists, neutral protagonists, female and male together, female, male and neutral together, and male and neutral together. The breakdown can be seen in Figure 21.

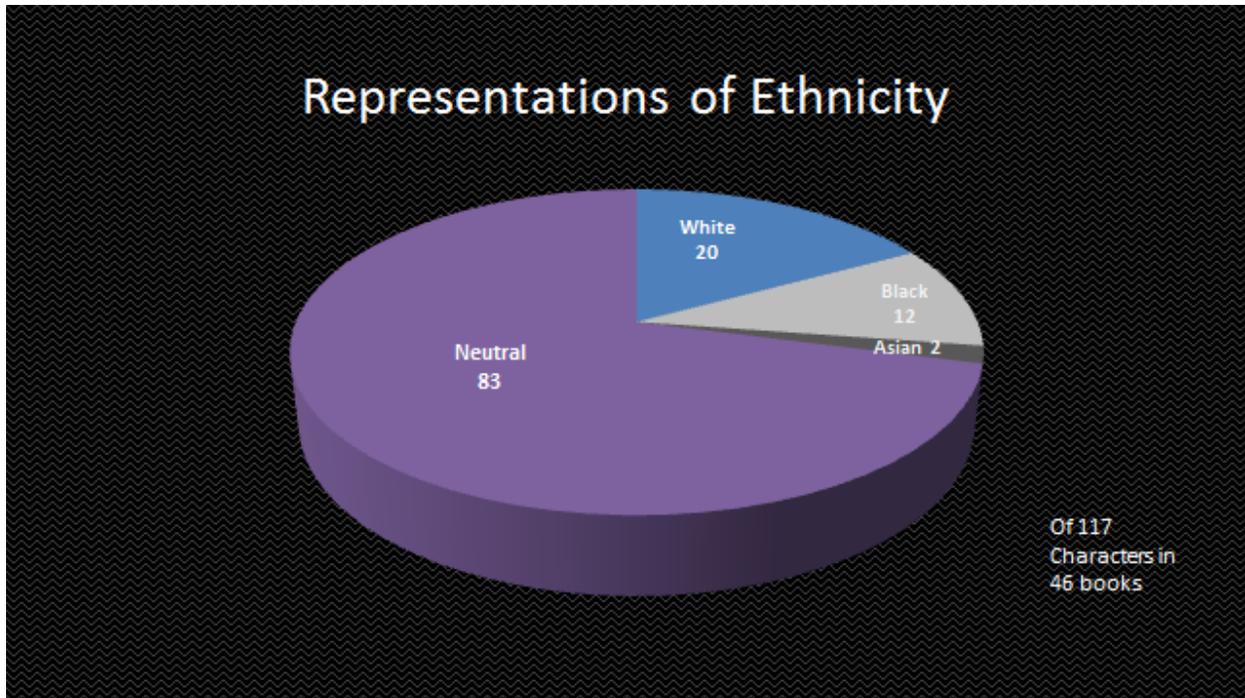


Figure 22 (Ethnicity and Age Appendix U)

Ethnicity. Ethnicity in the Geisel Collection is evident primarily in images except in the case of *Ling & Ting: Not Exactly the Same* in which there are cultural references in the text. In those books that had non white main characters, all but *Jazz Baby* showed mixed ethnicity in the subordinate characters. By far the greatest number of characters are represented by animals with no indication of ethnicity.

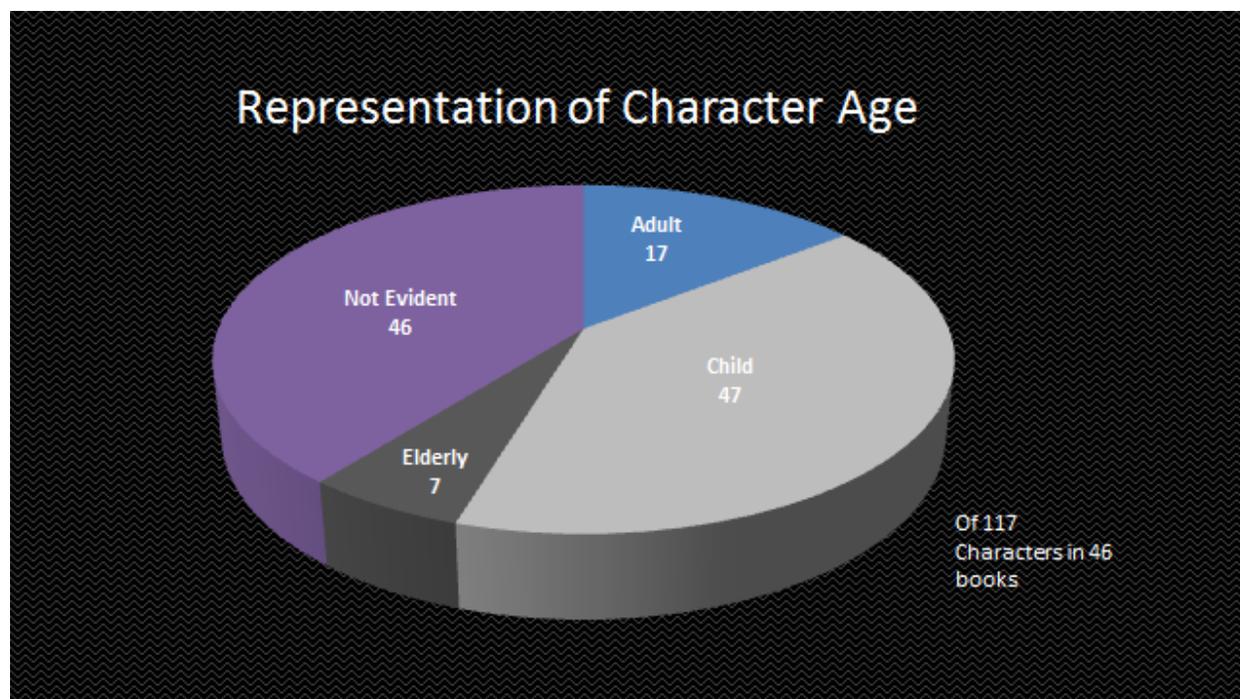


Figure 23 (Ethnicity and Age Appendix U)

Most of the primary characters are children or child animals, but supporting characters often include adults; in most texts it was parents, with the exception of *Don't Throw It To Mo!* in which the football coach, Steve, is a main character. The elderly are also present in these texts, both as supporting and main characters. Mr. Putter in *Mr. Putter and Tabby* is elderly as is Mrs. Goodwin in *Penny and her Marble* and Eugenia in *Mercy Watson Goes for a Ride*. Again, the evidence for this is solely conveyed in images of grey hair and walking canes.

Contributions of Visual Analysis.

Visual analysis was conducted on the Theodore Geisel collection with respect to image and the way it connects to a variety of issues not easily pigeonholed into word, sentence, or discourse analysis. These aspects are nonetheless important to this analysis and include how the art conveys text, and the significance of typeface to legibility.

Images and Text. One noticeable aspect of the visuals in these children's books is their ability to indicate emotions and values that are not evident in the text. This is particularly true in

works by author/illustrators. A good example of this is *There is a Bird on My Head*. In the story, Elephant has a nest of birds on his head, which is upsetting. suggests that Elephant simply ask the birds to leave. In the following illustration, Elephant thanks Piggie for his sage advice, but the image of Piggie depicts something beyond the text.



Typography. Typeface was analyzed here as well as on the previous page, this time for legibility including size and style. An effort was made to determine font type of all 46 books; however, the search yielded tpestyle for only 19. Of these only two styles repeated: Optima and Adobe Caslon, Optima being the preferred type for Boyds Mill Press, but there seemed to be little consistency across the board. Of the two series books, Mo Willem's *Elephant and* stories

have the same type though it is not identified in the text. Conversely, Tedd Arnold's *Fly Guy* books use completely different fonts from one another.

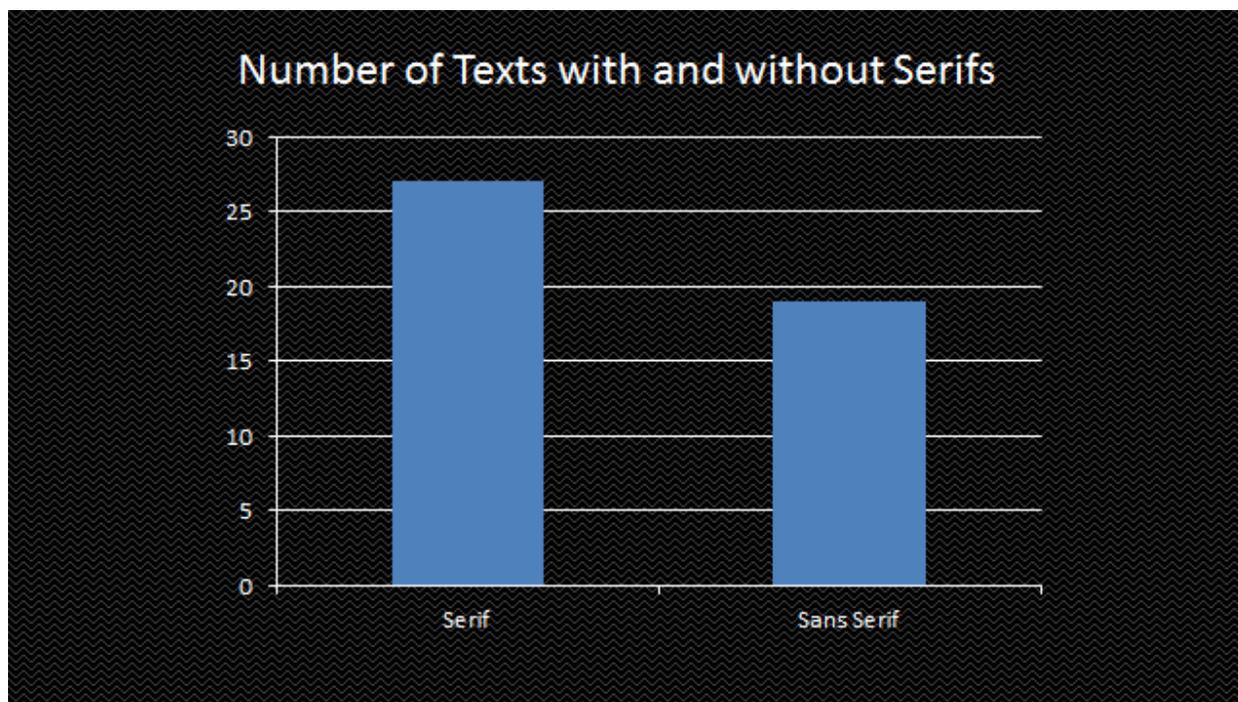


Figure 24 (Typography Appendix V)

While the typefaces could not all be identified, an analysis of serifs was conducted. Within the Geisel collection, 27 texts or 63 percent, use serif types and 19 texts or 36 percent, use sans serif. Point size is a problematic dimension of type because it is not a consistent measure across font styles so the dimensions of typeface were not considered in the scope of this research.

Conclusion

The preceding pages show an overwhelming amount of data and findings; though clearly these same data could be utilized to ask and answer myriad other questions, certain information most directly addresses the original research questions. The following concluding paragraphs will underscore what is specifically relevant to this study and why.

Analysis of General Information

Figures 1, 2, and 3, which are associated with general information (Author/Illustrator, Single and Multiple Winners, and Publisher), might seem perfunctory to a content analysis; however, in the context of a prestigious award such as the Theodor Geisel Medal, these elements take on greater meaning. In the history of the Newbery Award, 1922-2016, only 5 authors have received the award more than once; similarly between its inception in 1938 and today, the Caldecott committees awarded only a handful of wins to the same illustrators and over a much longer period of time, so the fact that in a 10 year span 6 authors account for nearly 40 percent of the winners for the Theodor Geisel Award might suggest an inadequate pool of texts in the competition and even more importantly, limited variety in materials for beginning readers.

Also worthy of consideration are genre, subject, and theme (Figure 4, Appendix F) in so far as all of these attributes are a common access point for teachers in the selection of texts. While there are a number of peer reviewed and published works that include lists of themes in literature, most do not sufficiently address early reading texts -- either because the themes were designed for higher level texts and were therefore too complex, or they were too limited because they are designed for student use. Lists related to subject and theme were compiled from Library of Congress Subject Headings for Children's and Young Adult Cataloging which are standardized and have the added benefit of allowing for access by teachers and librarians to build their collections, support classroom curricula, and match children to texts.

Readability Analyses

Readability measures have long been the standard in grading texts which assists teachers in placing appropriately leveled books in the hands of readers. Figures 5a and 5b show the variety of Lexile levels and ATOS levels among the texts in the Theodor Geisel collection. It is important to note that the Lexile algorithm is not sensitive to books at the lowest levels so they

are given a BR code for “Beginning Reader.” This inadequacy presents a problem when considering the importance of complexity in the CCSS which uses Lexile measures as their standard. ATOS was included in part because it does level texts at the lowest levels and also because it is used in more than 60,000 schools nationwide. Ultimately, the value of Lexile and ATOS measures in this research relates to both the complexity levels of the Common Core State Standards and as a contrast to other methods of computational linguistics.

Word Analyses

In the course of this research, word analysis proved to be a wellspring of valuable information, particularly with respect to Reading Foundational Skills. We know that the size of a child’s vocabulary is a critical aspect of early reading, and it is interesting to note the sheer number of words contained in this small collection. In just these 46 texts there are nearly 2000 discreet content words comprised of better than 1100 root words. This knowledge informs the tiered vocabulary analysis of the CCSS.

The brief examples in Figure 9 speak to the decoding difficulties students might encounter because of difficult or complex orthography. In addition to providing visibility of spelling patterns, Appendix T from which this table is derived, sheds light on the number and forms of cognates in the collection. Beyond these factors, word analysis informs reading fundamental skills with respect to sight words. Figures 5 and 6 demonstrate the ratios of high frequency words on both the text and the sentence level.

The figures dealing with parts of speech, 7 and 8, are again important to reading foundational skills. Repetition is a common and valuable tool for achieving fluency in young readers. The type-token ratios for parts of speech illustrated in Figure 8 not only demonstrate

which kinds of words have low repetition ratios, but reinforce the potential benefits of computational analysis of words.

Sentence Analyses

Applying aspects of prosody is vital to learning to read. Sentence analysis reveals the limitations of punctuation to promote prosody but also the presence and power of other devices such as type to enhance the teaching of this elusive but necessary skill. Figure 13 outlines the punctuation counts in each text, but this element alone does little to support early readers. Figure 18 demonstrates a variety of alternative prosodic supports to foundational reading skills within the Theodor Geisel texts. Manipulation of typography plays an important role in facilitating a child's ability to read with expression and proper emphasis. Sound devices shown in Figure 17 not only inform how these books can be used to enhance basic skills, but also have bearing on the qualitative requirements of the CCSS.

Understanding syntax is another aptitude that constitutes a fundamental reading skill. This area of analysis is highlighted in Figure 14, a syntactic analysis of sentences and Figure 15, a syntactic analysis of T-units. Because the CCSS relies on the quantitative measures of Lexile software which uses sentence length as a measure, sentence syntax is included in this study. Figure 15 offers a clearer picture of the potential complicating elements such as dependent clauses, coordinating phrases and complex nominals in T-units, which as discussed earlier are much more informative of syntactic complexity as it relates to foundational reading skills.

Discourse Analyses

A wide variety of literary devices is evident in the Theodor Geisel Collection. This factor is critical to the discussion of qualitative measures in the CCSS, which assumes that literary

language increases complexity. Figure 19 points out seven different literary devices as well as their prevalence in the collection.

Figures and tables were inadequate to addressing register within the collection, but again, the CCSS puts a great deal of emphasis on the purposes for reading, so in an effort to account for this factor, images were included. These images highlight a very interesting pattern that emerged, specifically, a kind of discourse confusion. Within these beginning-to-read texts, elements that we think of as belonging to informational texts are apparent in fictional books, and literary elements are apparent in the nonfiction texts in the collection. Graphic novels are addressed in the narrative because they have some impact with respect to reading foundational skills.

Characteristics of human experience are subsumed under discourse because written communication conveys social and cultural norms. Figures 21-24 examine these aspects of the Geisel collection which include gender, ethnicity, and age. While these features are present, the study makes clear the limitations of the expression of human experience in the collection.

Visual Analyses

The final section of results, visual analysis, does not conform to the original framework of the methodology, yet it is, perhaps, one of the most significant elements in early reading texts. This section speaks to all of the research questions: reading foundational skills, Common Core standards, and the representation of the human experience. Peculiar to these beginning-to-read texts is how images, even in isolation, can be read in the same way as words, sometimes parallel to the text but sometimes in conflict with it. Picture 2 embodies this characteristic.

Summary

While each section of these results provides interesting and informative data, taken together they speak to the advantages and limitations of computational linguistics for use in providing appropriate materials and informing reading teachers on what the texts our children read have to offer. Chapter 5 will discuss these results in terms of the 4 research questions as well as offer applications to pedagogy and suggestions for further study.

Chapter 5

Discussion and Implications

This study began as a personal journey to discover why the very first texts we read seem to stay with us throughout our lives. Early readers generally aren't great stories. The plots are thin and formulaic, the characters are flat and often stereotypic; they aren't part of our lived experience, so why do they remain so vivid?

From chapbooks to penny dreadfuls, from Marvel Comics to Little Golden Books, from *Dick and Jane* to *Cat in the Hat*, children's literature has a long and varied history with much to tell us about the children who read it. These are not necessarily literary masterpieces or lasting treatises, but they cannot be ignored because they speak to the skills and abilities necessary to "breaking the code" as well as a child's interest in and love for reading.

Beginning readers are distinguished from the larger genre of children's literature in that they are designed specifically for independent (child) reading. Since their inception, commercial readers have set themselves apart from other publications for children in that their production and design have been founded in the most current research-based theories in literacy. Today most beginning-to-read books are produced with an understanding of several factors we know to be important to early reading efforts: limited and controlled vocabulary, fewer words per page, simple and straightforward sentences, larger typeface, illustrations directly parallel to the text, and motivating stories centered in a child's everyday experiences (ALA 2016, Morgan & Morgan 1995).

Research also tells us that in order to read these texts, children must have alphabetic knowledge or letter recognition, the ability to discriminate sounds and the ability to match letters and sounds to construct or deconstruct words. We know that this knowledge of the relationships

between the written letter and the sound facilitates fluency in reading, and that fluency aids in comprehension. But just knowing the letters and sounds isn't the whole story. Reading accuracy alone doesn't translate into understanding; punctuation, proper syntax, and line breaks are equally vital to comprehension. In addition to these skills based and practical considerations, context is also important to the understanding of any literature, particularly children's literature because it is based in sociological and philosophical underpinnings of the world we want to convey to our children.

This study examined all of these aspects of beginning-to-read books among the titles which received gold or silver medals of the Theodor Geisel Award sponsored by the American Library Association. The findings of this research are addressed question by question as multiple criteria i.e., lexical analysis, cultural content, syntactic analysis are together employed to inform each question in a different way.

Findings and Discussion

Research Question One: To what extent do the Theodor Geisel Award and honor books address foundational reading skills?

There is an inherent tension between the act of reading and the purpose of reading. Children learning to read are motivated not only by the content (the story) or information as is the adult, but they are motivated by the act itself without context much like other early skills. Toddlers often walk or run for the sheer joy of locomotion, not to go anywhere in particular. So while reading has great value in terms of what it can teach, for the early reader there is value in the skill as well.

Findings related to Vocabulary. According to the Theodor Geisel Award criteria, new words should be "added slowly enough to make learning them a positive experience, and words

should be repeated in an easily recognizable pattern to ensure knowledge retention” (ALA 2016). The results of this study demonstrate the degree to which the collection supports building vocabulary.

The word analysis showed that high frequency words make up 43 percent of each text. However when accounting for repetition, high frequency words make up less than 25 percent of texts. Of these unique content words, nouns clearly have the highest rate of repetition, followed by verbs, adjectives and adverbs respectively. In addition to repetition, verbs potentially have the benefit of having easily accessible cognates, however, descriptive words occur less frequently and have a very low ratio of repetition.

Implications. It becomes clear from these results where students are constrained by a lack of controlled vocabulary. Educators can and should use this knowledge in the selection of texts and the integration of vocabulary. We know that in order to add a word to a child’s vocabulary, it must be heard or seen dozens of times. The results suggest that many content words do not occur frequently enough; however, this might be mitigated by the fact that these books are designed to be read multiple times, so the number of repetitions may be sufficient.

Findings Related to Prosody. Prosody, which includes timing, intonation, and stress, directly affects meaning; it lends an emotional quality to words or phrases and helps us understand what we are reading. What is important to know about this foundational skill is that rhythmic patterning of speech is vital to success in learning to read. When children read they create a network in the brain which allows access to spoken language through visual stimuli. How words are put together and the sound patterns evident in texts is important. Results related to punctuation and sentence complexity both demonstrate the importance of syntax and prosody.

Punctuation is a limited way of conveying prosody. The question mark implies a raised inflection at the end of a sentence, while an exclamation point indicates emphasis, however, neither offer an access point for children to understand or communicate what words are operative in the sentence. Among the books in the collection, punctuation counts were not surprising, with the exception that exclamations account for over 20 percent of texts. An unexpected finding was the significance typography played to the communication of prosody. Many of the Theodor Geisel texts reflect an effort to increase a child's ability to read with appropriate inflection through the use of fonts. This was accomplished by increasing font size, printing words in all capitals, bolding words, using italics, and hand lettering.

Our current standards of readability measure syntactic complexity based on the length of sentences measured in words, and word difficulty based on number of letters or syllables in a word. However, the T-unit analysis elucidated sentence structure and syntactic complexity within the collection in a much more comprehensive way. While simplicity of sentences is a hallmark of the collection, the results related to sentence complexity were striking when the texts were compared against one another. For example Cynthia Rylant's *Henry and Mudge and the Great Grandpas* and Josh Schneider's *Tales for Very Picky Eaters*, have close averages in sentence length of 7.8 and 8.1 words per sentence, respectively. However the T-unit analysis is much more revealing of difficulty. *Tales for Very Picky Eaters* has 29 percent more verb phrases, 58 percent more dependent clauses, and 8 percent more complex nominals than Rylant's book. It is interesting to note that *Henry and Mudge and the Great Grandpas* has 52 percent more coordinating phrases. This is important to note because in many cases, coordinating phrases can increase the length of a sentence without significantly increasing difficulty. This difference is illuminated in simple 8 word sentences from both texts: "Henry and Mudge and everyone there

felt happy.” versus “He works all night long perfecting his recipe.” Complexity of sentences is inadequately addressed by current readability measures.

Implications: Sensitivity to prosody is highly correlated with vocabulary, reading accuracy and comprehension. Changes in typography can make texts much more accessible to young children in terms of learning this foundational skill. Information related to syntactic analysis could potentially be used to enhance our ability to determine the difficulty or complexity of beginning to read books; in particular, it has value in illuminating discrete factors such as coordinating phrases and dependencies. These measures of syntax can be used in tandem with readability scores to communicate more precisely the true difficulty or complexity of a text and its accessibility to students.

Research Question Two: To what extent do the Theodor Geisel Award and honor books measure up against both quantitative and qualitative standards for text complexity in the CCSS?

The CCSS for English Language Arts do not set standards for complexity in the early grades. However, as Mesmer, Cunningham, and Hiebert (2012) have noted, this void presents a “window of opportunity” to explore similar standards for the early grades and construct a theoretical model for beginning-to-read books.

Findings Related to Quantitative Measures. Though the CCSS does not set Lexile measures until grade 3, the Theodor Geisel Collection ranges from BR, beginning reader, to 540L, just within the first quartile of levels in the grade 2-3 band of 450-790. This includes several books designated as being most appropriate for read aloud.

Findings Related to Qualitative Measures. The CCSS outlines qualitative text considerations under the following headings: Levels of Meaning, Structure, Language Conventionality and Clarity, and Knowledge Demands.

Levels of Meaning. Though not prevalent, the collection does include some complex literary devices such as simile, metaphor, idiom, and pun, which have meanings that must be inferred or are at odds with the text. This is particularly true of the artwork. Often sarcasm and irony are evident in the visuals only, even though the words are not there to support it.

Structure. Texts within the Geisel Collection have a wide variety of structures. As noted in the results section, several of the literary texts include a table of contents, an attribute not generally found in adult fiction even when chapters are present. More perplexing is the insertion of different registers within a single text such as fiction, which includes lists and songs. Another consideration related to structure is the inclusion of graphic novels. They introduce not only left to right orientation but an up-and-down orientation as well. In addition, they contain a great deal of environmental print.

Language Conventionality and Clarity. For the most part, this collection adheres to clear, conventional and contemporary language with the exceptions noted above in levels of meaning. Still, throughout these texts, rhyme is prevalent both in fiction and non-fiction, and while the language and meaning are clear, the format is atypical for narrative.

Knowledge Demands. Within the collection, knowledge demands are appropriately limited to characters of a similar age to the reader and engaged in activities and dilemmas common to their age group and experience. The informational and concept books are primers and assume little prior knowledge.

Implications. With respect to quantitative measures, even the highest level texts within the Geisel Collection do not meet the minimum of the CCSS; however, as outlined in the implications related to prosody, many of the texts that do not meet Lexile standards are fairly complex in terms of syntax, and this should be no less valued. For example, *Bink and Gollie* has a Lexile level of 310, well below the minimum for grades 2 -3 in the CCSS; however, it has a high degree of complexity with an averages of 45 percent of T-units containing dependent clauses, 1.4 complex nominals per T-unit and 2 verb phrases per T-unit; these numbers are among the highest in the collection.

One of the most significant issues related to text complexity and the CCSS deals with discourse. Teachers of reading are taught to identify for their students typical structures in a variety of registers and among different kinds of discourse. From this examination, it is particularly important for educators to be aware of the potential for “discourse confusion” among these texts, as both fiction and non-fiction have elements normally associated with the opposite genre. For example, *Vulture View* is an informational text written entirely in verse. Teachers should also be aware of the potential graphic novels have to be confusing for new readers. Graphic novels have both vertical and horizontal orientations and a considerable amount of environmental print, which is often either distracting or at a level that is higher than the vocabulary or the language present in the text of the story.

There is quite a bit of literary play, which adds to the qualitative complexity suggested by the CCSS. Additionally, the rate at which new vocabulary is introduced and the sparsity of repetition in some categories can also be a function of complexity. However, it is worth keeping in mind that for most new readers, all vocabulary can be said to add to complexity. Finally, as regards knowledge demands, these earliest reading texts are designed to create schema, and in

many cases teach content from a zero base. A good example is *First the Egg*. On the surface it is designed to teach sequence as the entire book follows a pattern of “First...then.” Yet this book is catalogued as “developmental biology,” which it is. For example, the first two sequences are: “First the EGG, then the CHICKEN and First the TADPOLE then the FROG.” On an even deeper level, this text creates a kind of extended literary juxtaposition by ending the book with a child’s painting of “First the *CHICKEN*, then the *EGG*.”

Research Question Three: To what degree do the Theodor Geisel Award and honor books reflect the core literacy values related to the variety of human experience including gender, ethnicity and age?

The notion of respecting a child’s experience has been stimulated by the understanding that in order to foster engagement with reading, young children have to see themselves in the text. In *Reading the Word and the World*, Freire (1987) refers to words as “laden with the meaning of people’s existential experience” (p. 35). Literature reflects our attitudes about aspects of the human condition as well as what themes and topics we value as a society. For preschool and elementary students, these attitudes can be measured in terms of the presentation of both text and images which reflect ethnicity, gender, age, character role, appearance, and any number of other sociocultural messages.

Findings. It is clear from the results that these books lack a certain amount of diversity. There are no differently-abled characters, and there is a noticeable lack of diversity in terms of race, ethnicity, and culture. Male and female characters are relatively evenly divided with a fair number of characters with indeterminate gender (personified animals and objects). It is interesting to note that in cases where the author is also the illustrator, stereotypic characteristics are more prevalent, and gender is often indicated only visually and not in the text. Only one book

has a mixture of black and white major characters, though others do have subordinate or background characters with some diversity. All of the books depicted the life experiences of young children and promoted a positive view of the world.

Implications. As educators, we know that advocating for social justice begins with understanding. According to Creighton (1997) “Critical pedagogy brings to the classroom an awareness of the structure of cultural systems and positions of power therein.” (p. 439) Through her research, she suggests ways of fostering critical literacy in teacher practice. Teachers must be conscious of how texts mirror the children in the classroom, and how illustrations reflect stereotypes. When developing critical literacy in young children, it is important to look for and present literature which reflects common beliefs, values and experiences so that all children, from a multiplicity of backgrounds, can share in the text.

Research Question Four: How can computational linguistic analysis be used to inform knowledge of texts for application in the classroom?

Until recently, research on measuring prosody was conducted solely on readers, not texts. However, in the process of researching ways of measuring syntax for beginning readers, an interesting parallel with text to speech computer programming became apparent. Computers learn to read in much the same way as children. They are trained to read aloud in two ways: first they are trained to recognize words and sentence structure, and then they are taught to *hear* the sounds and rhythm of language with a large pre-recorded speech database with proper pitch, duration, and energy. Similarly, readers need texts which support prosody because without the proper signifiers and without rhythmic structures they would have a difficult time comprehending.

Findings and Implications. The results of this research would not have been possible without the utilization of computer assisted corpus linguistics. Corpus linguistics and computational methods have been used in ELL contexts for many years. Keith Forse (2011) discusses the value of corpus linguistics to developing vocabulary lists; Simpson-Vlach and Ellis (2002) have done important research related to patterns and sequences of multiple words and phrases; Susan Conrad studied register and demonstrated a relationship between grammar and vocabulary. Still other researchers have used computational linguistics to determine syntactic complexity based on indices of comprehension (Rosenberg and Abbeduto, 1987; Covinton et al, 2006) In addition to studying texts, corpus linguistic tools have been demonstrated to have great value in terms of analyzing student output. What is unique about this study, however, is not the use of computational methods, but rather their application to a corpus of early readers.

The findings throughout this project were informative in ways that traditional readability measures cannot match. As illustrated in the findings for Question 2, computational linguistics, specifically syntactic analyses, are far more informative than current quantitative measures such as Lexile and ATOS. Moreover, with this technology, educators have an opportunity to select texts to support specific skills. For example, looking at the results for the T-unit analysis (Appendix Q), a teacher wanting to focus on coordinating phrases could choose *Wolfsnail*, *Supertruck*, and *Mr Putter and Tabby Turn the Page*, which all average better than 2 coordinating phrases per T-unit. This knowledge could be helpful in selecting companion texts for a number of different grammar skills.

The ability to see into a text has great implications for the pedagogy of early reading with respect to vocabulary. Teachers can now preview and pinpoint words precisely linked to the texts they use in their individual curricula. Computational linguistic tools can be programmed to

search for words by syllables, parts of speech, even orthography; it can be used to determine the rate at which individual words are repeated in a text.

The power of these tools is ever improving in both the ability to extricate elements of text and the accuracy of these measures. In a 2003 article, O’Keeffe and Farr even suggest ways for teachers to integrate student use of this technology. Most encouraging is the relative simplicity of some functions of computational linguistics and the availability of free online programs, which require little training and could be easily integrated into reading instruction for teachers.

Considerations for Practitioners, Parents, and Prize Committees

Teachers. Teachers can benefit from several of the analyses conducted in this study. With respect to lexical measures, getting word frequency counts to uncover vocabulary is not only helpful, but quite easy. The only caveat to this is that currently there are no open source programs that will accept image files for analysis, so converting books to plain text formats is necessary. In terms of syntactic measures, while teachers could profit greatly in their ability to provide students with texts targeted to their individual needs, from a practical standpoint, until syntactic complexity measures are more easily accessible and more common, finding detailed information on specific texts will require initiative and time.

Teachers often depend on leveled texts to support their curriculum; this study shows that these measures are not always indicative of text difficulty. These analyses reveal much about books that is not apparent to the naked eye, but even without access to computational methods, it reminds us of the need to be aware of the infinite levels of complexity in seemingly simple texts.

Parents. For the most part, parents can feel confident that the overall collection is of good quality and presents an appropriate range of materials for early readers. That said, the most valuable tool in a parent’s arsenal is repetition, especially with regard to building and supporting

vocabulary acquisition. While in aggregate, these texts are all capable of being read by new readers, none are designed for a single reading. So when a child asks to “read it one more time,” read it two!

Selection Committees. With respect to the nominating and vetting process for prize winning books, I hope that committees will consider not only story and comprehension, but some of the underlying skills necessary to read good books and whether or not their selections support the reading process. There are many computational linguistic programs available and the knowledge gained can be valuable in terms of seeing inside texts where the human eye and emotion cannot.

Future Directions

The future directions for this research are truly unlimited. There are two areas which I would like to pursue, one related to computer analysis of texts and the other related to the study of prosody in early reading. Certainly, using computational linguistics to measure the content of reading programs in use across the nation is a natural next step. I would pose the question: To what extent do the reading selections in popular reading programs address foundational reading skills? Additionally, there are algorithms for more sophisticated syntactic analysis as well as other tools for data mining and sentiment analysis that could be applied to higher level materials and disciplinary texts. Using data mining trend analysis, an example question might be asked, How has the teaching of grammar changed based on a content analysis of grammar texts published by over the last 50 years?

Another aspect of this research that has led me to additional questions relates to prosody and syntactic learning. Having been impressed by the ability of graphic text to convey prosody, I would like to test the potential for this format to actually affect reading proficiency. For example,

based on an intervention using selected texts with a high degree of typographic variation: Does oral prosody improve in populations which receive the treatment? If so, is there any evidence of a correlational improvement in comprehension?

Final Thoughts

I am from the Little Golden Books generation in which picture and story are inextricably tied together. I remember vividly the story lines and the images of my first books. I can see my childhood bedroom, the ferns and roses outside my window and a magical landscape beyond – magical, not because it was so beautiful, but because the world outside was where I played pretend. Perhaps this is why the first book I choose to read was *Through my Window*. My bedroom was a thousand worlds to me because it is where I brought to life the worlds inside the pages of my books.

In the process of this adventure, I was curious to know if others felt such a deep attachment to their “first books.” I asked around – a 12 year old child, an 88yr old retiree, a physician...I was amazed at the clarity with which they all remembered the story lines and reimagined the illustrations; I was surprised at the affection with which they spoke -- eyes lit, wistful smiles. In an effort to rekindle that feeling, before I started on this journey, I hunted down my first book. As I opened the cover I became my 4 year old self, a little golden haired child just like the one in the book, wearing a blue pinafore and peering out the window. The words came rushing back to me and in an instant I was transported over decades.

By the end of the journey through this collection, I have come to the real truth of this research: I am utterly and completely in awe of the capacity of children to conquer the complexity of learning to read, and I have a new appreciation for the pedagogy of teaching children to be literate and all that the word connotes. I hope that this study will not only inform,

but will remind teachers of their personal moments of awe inspired by their first books. I hope it will bring an understanding that each of *these* books might be “the one” for some child, and what a privilege it is to be part of that.

References

- Albers, P. (2008). Theorizing visual representation in children's literature. *Journal of Literacy Research*, 40(8), 163-200.
- American Library Association. (2016). Theodore Seuss Geisel award terms and criteria. Retrieved from <http://www.ala.org/alsc/awardsgrants/bookmedia/geiselaward/geiselawardtermscriteria/geiselawardtermscriteria>
- Anderson, D. A., Hamilton, M. (2005). Gender role stereotyping of parents in children's picture books: The invisible father. *Sex Roles*, 52(3/4), 145-151.
- Ansello, E. F. (1978). Ageism—The subtle stereotype. *Childhood Education*, 54(3), 118-122.
- Arbuthnot, M. H. (1964). *Children and books*. Chicago, IL: Scott, Foresman.
- Banks, J. A., & Banks, C. A. M. (2009). *Multicultural education: Issues and perspectives*. John Wiley & Sons.
- Beaugrande, R., Dressler, W. (1981). *Introduction to text linguistics*. New York, NY: Longman.
- Beck, I. L., McKeown, M. G., & Kucan, L. (2008). *Creating robust vocabulary: Frequently asked questions and extended examples*. New York, NY: Guilford Press.
- Benjamin, R., & Schwanenflugel, P. (2010). Text complexity and oral reading prosody in young readers. *Reading Research Quarterly*, 45(4), 388-404.
- Betts, E. A. (1946). *Foundations of reading instruction*. New York, NY: American Book Company.

- Biber, D., & Conrad, S. (2001). Quantitative corpus-based research: Much more than bean counting. *TESOL quarterly*, 35(2), 331-336.
- Bird, S., Klein, E., & Loper, E. (2009). *NLTK Book*.
- Bruner, J. S. (2009). *The process of education*. Harvard University Press.
- Carrell, P. (1987). A view of written text as communicative interaction: Implications for reading in a second language. J. Devine (Ed.), *Research in Reading in English as a Second Language* (pp. 21-40). Washington, DC: TESOL.
- Chall, J. S. (1983). *Stages of reading development*. New York, NY: McGraw-Hill.
- Clark, R., Guilmain, J., Saucier, P., & Tavaréz, J. (2003). Two steps forward, one step back: The presence of female characters and gender stereotyping in award-winning picture books between the 1930s and the 1960s. *Sex Roles*, 49(9), 439–449.
- Clay, M. (1991). *Becoming literate: The construction of inner control*. Portsmouth, NH: Heinemann.
- Collier, M. D. (2000). Through the looking glass: Harnessing the power of African American children's literature. *Journal of Negro Education*, 235-242.
- Common Core State Standards Initiative. (2012). 2012 Common Core State Standards for English language arts & literacy in history/social studies, science, and technical subjects, Appendix A. Washington, DC: National Governors Association Center for Best Practices, Council of Chief State School Officers.
- Connelly, M. F., & Clandinin, J. D. (1990). Stories of experience and narrative inquiry. *Educational Researcher*, 19(5), 2–14.

- Conrad, S. (2000). Will corpus linguistics revolutionize grammar teaching in the 21st century? *Tesol Quarterly*, 34(3), 548-560.
- Covington, M. A., He, C., Brown, C., Naci, L., & Brown, J. (2006). How complex is that sentence? A proposed revision of the Rosenberg and Abbeduto D-Level Scale. CASPR Research Report, 2001(1).
- Crabb, P. B., & Bielawski, D. (1994). The social representation of material culture and gender in children's books. *Sex Roles*, 30(1/2), 69–79.
- Creighton, D. C., (1997). Critical literacy in the elementary classroom. *Language Arts*. 74(6) 438-445.
- Creswell, J. W. (1994). *Research design: Qualitative & quantitative approaches*. Thousand Oaks, CA: Sage Publications.
- Dale, R. (1956). The problem of vocabulary in reading. *Educational Research Bulletin*, 35(5), 113–123.
- Dale, R., Chall, J. (1948) A formula for predicting readability: Instructions. *Educational Research Bulletin*, 27(2), 37-54.
- Diekman, A. B., & Murnen, S. K. (2004). Learning to be little women and little men: The inequitable gender equality of nonsexist children's literature. *Sex Roles*, 50, 373–385.
- Dolch, E. W. (1948). *Problems in reading*. Garrard Press.
- Ehrlich, Hannah. (2015) The diversity gap in children's publishing. *The Open Book*. [blog.leeandlow.com/2015/03/05/the-diversity-gap-in -children's-publishing-2015/](http://blog.leeandlow.com/2015/03/05/the-diversity-gap-in-children-s-publishing-2015/).

- Esser, J., & Polomski, A. (1988). *Comparing reading and speaking intonation*. Amsterdam, The Netherlands: Rodopi.
- Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology*, 32(3), 221–233.
- Fountas, I. C., & Pinnell, G. S. (2001). *Guiding readers and writers, grades 3-6: Teaching comprehension, genre, and content literacy*. Portsmouth, NH: Heinemann.
- Fox, M. (1993). Men who weep, boys who dance: The agenda between the lines in children's literature. *Language Arts*, 70, 84-88.
- Freire, P., Macedo, D. (1987). *Literacy: Reading the word and the world*. South Hadley, MA: Bergin & Garvey Publishers, Inc.
- Frith, U. (1985). Beneath the surface developmental dyslexia. In K. Patterson, J. Marshall, & M. Coltheart (Eds.), *Surface dyslexia, neurophysiological and cognitive studies of phonological reading* (pp. 301–330). London, England: Erlbaum.
- Fry, E. (1968). A readability formula that saves time. *Journal of Reading*, 11(7), 513-516, 575-578.
- Fry, E. (2002). Readability versus leveling. *The Reading Teacher*, 56(3), 286-291.
- Gibbons, P. (1993). *Learning to learn in a second language*. Portsmouth, NH: Heinemann.
- Grassmer, A. C., McNamara, D. S., Louwrese, M. M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods*, 36(2), 193–202.

- Gries, S. Th. (2009). What is corpus linguistics? *Language and Linguistics Compass*, 3, 1-17.
- Hadaway, N., Vardell, S., & Young, T. (2004). *What every teacher should know about English language learners*. Boston, MA: Pearson.
- Hamilton, M., Anderson, D. A., Broaddus, M., & Young, K. (2006). Sex stereotyping and underrepresentation of female characters in 200 popular children's picture books: A 21st century update. *Sex Roles*, 55(11), 757–765.
- Heintz, K. E. (1987). An examination of sex and occupational-role presentations of female characters in children's picture books. *Women's Studies in Communication*, 11, 67–78.
- Herber, H. (1978). *Teaching reading in content areas, 2nd ed.* Englewood Cliffs, NJ: Prentice-Hall.
- Holliman, A., Williams, G., Mundy, I., Wood, C., Hart, L., & Waldron, S. (2014). Beginning to disentangle the prosody-literacy relationship: A multi-component measure of prosodic sensitivity. *Reading and Writing*, 27, 255–266.
- Juel, C., & Minden-Cupp, C. (2000). Learning to read words: Linguistic units and instructional strategies. *Reading Research Quarterly*, 35(4), 458–492.
- Kincaid, J. P., Fishburne, R. P., Rogers, R. L, Chissom, B.S. (1975). Derivation of new readability formulas (automated readability index, fog count, and flesch reading ease formula) for Navy enlisted personnel. *Research Branch Report, 8-75*, Memphis: Chief of Naval Technical Training: Naval Air Station Memphis.
- Klare, G. (1963). *The measurement of readability*. Ames, IA: Iowa State University Press.

- Klauda, S. L., & Guthrie, J. T. (2008). Relationships of three components of reading fluency to reading comprehension. *Journal of Educational Psychology*, 100(2), 310–321.
- Kolbe, R., & LaVoie, J. (1981). Sex-role stereotyping in preschool children's picture books. *Social Psychology Quarterly*, 44, 369–74.
- Kortenhaus, C. M., & Demarest, J. (1993). Gender role stereotyping in children's literature: An update. *Sex Roles*, 28, 219–232.
- Krashen, S. D., & Terrell, T. D. (1983). *The natural approach*. San Francisco, CA: Alemany Press.
- Krippendorf, D., & Bock, M. (2009). *The content analysis reader*. Thousand Oaks, CA: Sage Publications.
- Kruse, M. (2001). Escaping ethnic encapsulation: The role of multicultural children's literature. *The Delta Kappa Gamma Bulletin*, 67(2), 26-32.
- Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics*, 15(4):474-496. [L2 Syntactic Complexity Analyzer]
- Lu, X. (2012). The relationship of lexical richness to the quality of ESL learners' oral narratives. *The Modern Language Journal*, 96(2), 190-208. [Lexical Complexity Analyzer]
- Marcus, L. (2007). *Golden legacy: How golden books won children's hearts, changed publishing forever, and became an American icon along the way*. New York, NY: Golden Books.

- Marks, C. B., Doctorow, M. J., & Wittrock, M. C. (1974). Word Frequency and Reading Comprehension. *The Journal of Educational Research*, 67(6), 259-262.
- Mavrogenes, N. A. (1982). Positive images of grandparents in children's picture books. *The Reading Teacher*, 35(8), 896-901.
- Mead, G. H. (1934). *Mind, self & society from the standpoint of a social behaviorist*. Chicago, Ill: University of Chicago Press.
- Merriam, S. B. (2002). *Introduction to qualitative research*. San Francisco, CA: Jossey-Bass.
- Mesmer, H. A., Cunningham, J. W., & Hiebert, E. H. (2012). Toward a theoretical model of text complexity for the early grades: Learning from the past, anticipating the future. *Reading Research Quarterly*, 47(3), 235–258.
- Mitchell, G. T. (1954). *Through my window*. Racine, WI: Cozy Corner Press.
- Moats, L. (1994). The missing foundation in teacher education: Knowledge of the structure of spoken and written language. *Annals of Dyslexia*, 44, 81–102.
- Moore, D. W., & Readence, J. E. (1983). Approaches to content-area reading instruction. *Journal of Reading*, 26(5), 397–402.
- Morais, J. (1991). Constraints on the development of phonemic awareness. In S. A. Brady & D. P. Shankweiler (Eds.), *Phonological processes in literacy: A tribute to Isabelle Y. Liberman*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Morgan, J., & Morgan, N. (1997). *Dr. Seuss & Mr. Geisel: A biography*. Boston, MA: Da Capo Press.

- Narahara, M. (1998). *Gender stereotypes in children's picture books*. (ERIC Document Reproduction Service No. ED419248.) East Lansing, MI: National Center for Research on Teacher Learning.
- National Literacy Trust. (2015). Phonics – methods of teaching. Retrieved from http://www.literacytrust.org.uk/resources/practical_resources_info/1035_phonics-methods_of_teaching
- National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups*. Washington, DC: U.S. Government Printing Office.
- Peterson, S. B., & Lach, M. A. (1990). Gender stereotypes in children's books: Their prevalence and influence on cognitive and affective development. *Gender and Education, 2*, 185–197.
- Powell, W. R. (1968). Reappraising the criteria for interpreting informal inventories. Washington, DC: ERIC 5194164.
- Rachlin, S. K., & Vogt, G. L. (1974). Sex roles as presented in children's coloring books. *Popular Culture, 8*(3), 149–156.
- Rosenblatt, L. M. (1978). *The reader, the text, the poem: The transactional theory of the literary work*. Carbondale, IL: Southern Illinois University Press.
- Schau, C. G., & Scott, K. P. (1984). Impact of gender characteristics of instructional materials: An integration of the research literature. *Journal of Educational Psychology, 76*, 183–193.

- Schreiber, P. A. (1991). Understanding prosody's role in reading acquisition. *Theory Into Practice, 30*(30), 158–164.
- Seymour, P. H., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology, 94*, 143–174.
- Shanahan, T. (2013). Rejecting instructional level theory. Retrieved from <http://www.shanahanonliteracy.com/>
- Smith, B. (1964). Patterns of writing in different school subjects. *Journal of Reading, 8*, 31–37, 97–102.
- Smith, F. (1992). Learning to read: The never ending debate. *Phi Delta Kappan, 7*(6), 432–441.
- Spache, G. (1953). A new readability formula for primary-grade reading materials. *The Elementary School Journal, 53*(7): 410–13. DOI 10.1086/458513. JSTOR 998915
- Spache, G. D. (1972). *Good Reading for Poor Readers*. 8th ed. Champaign, IL: Garrard Publishing.
- Spencer, M. (1988). *How texts teach what readers learn*. Victoria, BC: Abel Press.
- St. Peter, S. (1979). Jack went up the hill . . . but where was Jill? *Psychology of Women Quarterly, 4*, 256–260.
- Stenner, A. J., Smith, D. R., & Smith, M. (1988). *The lexile framework*. Durham, NC: MetaMetrics.
- Stevens, R. J., Liu, X., Baker, D. P., Ray, M., Eckert, S., & Gamson, D. (2015). Assessing the cognitive demands of a century of reading curricula: An analysis of reading text and comprehension tasks From 1910 to 2000. *American Educational Research Journal, 52*(3), 582-617.

- Stone, C. R. (1956). Measuring difficulty of primary reading material: A constructive criticism of Spache's measure. *The Elementary School Journal*, 57(1) 36-41.
- Taylor, F. (2000). Content analysis and gender stereotypes in children's books. *Teaching Sociology*, 31(3), 300–311.
- Thorndike, E. L., & Lorge, I. (1944). The teacher's word book of 30,000 words.
- Tognoli, J., Pullen, J., & Lieber, J. (1994). The privilege of place: Domestic and work locations of characters in children's books. *Children's Environments*, 11, 272–280.
- Trepanier-Street, M. L., & Romatowski, J. S. (1999). The influence of children's literature on gender role perceptions: A reexamination. *Early Childhood Education Journal*, 26, 155–159.
- Walley, A. C., Metsala, J. L., & Garlock, V. M. (2003). Spoken vocabulary growth: Its role in the development of phoneme awareness and early reading ability. *Reading and Writing: An Interdisciplinary Journal*, 16, 5–20.
- Weitzman, L., Eiffer, D., Hokada, E., & Ross, C. (1972). Sex-role socialization in picture books for preschool children. *American Journal of Sociology*, 77, 1125–1150.
- Wilcox, G. (2009). *Introduction to linguistic annotation and text analytics*. Toronto: Morgan & Claypool.
- Williams, A. J., Vernon, J. A., Williams, M.C. Malecha, K. (1987). Sex role socialization in picture books: An update. *Social Science Quarterly*, 68, 148–156.
- Williams, L. M. (2008). Book selections of economically disadvantaged black elementary students. *The Journal of Educational Research*, 102(1), 51-64

Young, A. R., Bowers, P. G., & MacKinnon, G. E. (1996). Effects of prosodic modeling and repeated reading on poor readers' fluency and comprehension. *Applied*

Psycholinguistics, 17(1), 59–84.

Ziegler, J., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory.

Psychological Bulletin, 131(1), 3–29.

Appendix A

Theodor Seuss Geisel Award Terms and Criteria

Theodor Seuss Geisel Award Terms and Criteria

Purpose: The Theodor Seuss Geisel Award recognizes the author(s) and illustrator(s) of a book for beginning readers who, through their literary and artistic achievements, demonstrate creativity and imagination to engage children in reading.

Criteria: The award shall be given annually to the author and illustrator of the most distinguished contribution to the body of American children's literature that encourages and supports the beginning reader published in English in the United States during the preceding year. There are no limitations as to the character of the book considered except that it will be original and function successfully as a book for beginning readers. Honor Books may be named. These shall be books that are also truly distinguished.

The Award is restricted to author(s) and illustrator(s) who are citizens or residents of the United States.

The committee in its deliberations is to consider only the books eligible for the award, as specified in the terms.

The "author(s) and illustrator(s)" may include co-authors and co-illustrators. The author(s) and illustrator(s) may be awarded the medal posthumously.

"Distinguished" is defined as: Marked by distinction: noted for significant achievement; Marked by excellence in quality; Marked by conspicuous excellence or eminence; Individually distinct; Providing a stimulating and successful reading experience for the beginning reader containing the kind of plot, sensibility, and rhythm that can carry a child along from start to finish.

"American children's literature" means that books originally published in other countries are not eligible.

"Contribution to the body of children's literature that encourages and supports the beginning reader" indicates the text of a book, which must be directed at readers from pre-K through Grade 2. The book must also contain illustrations, which function as keys or clues to the text. The committee shall consider all forms of writing—fiction, non-fiction, and poetry. Reprints and compilations are not eligible.

Committee members need to consider the following criteria: Subject matter must be intriguing enough to motivate the child to read; The book may or may not include short "chapters"; New words should be added slowly enough to make learning them a positive experience; Words should be repeated to ensure knowledge retention.; Sentences must be simple and straightforward; There must be a minimum of 24 pages. Books may not be longer than 96 pages; The illustrations must demonstrate the story being told; The book creates a successful reading experience, from start to finish; The plot advances from one page to the next and creates a "page-turning" dynamic.

Number and Frequency of Award: One Theodor Seuss Geisel Award for one book will be announced in January of each year during the American Library Association Midwinter Meeting to recognized author(s) and illustrator(s). The announcement of the first Theodor Seuss Geisel Award is January 2006, with the first award presented in 2006, and to continue every year thereafter. If a suitable candidate is not found for any given year, the award will not be presented that year.

Selection of Jury to Administer the Award: The President of the Association for Library Service to Children, a division of the American Library Association, will appoint the committee to administer the award. The first committee shall be appointed in fall 2004; and the first committee to sit in 2005. Membership shall be one chair and six members.

Deadline for Nomination of Candidates: The deadline for award entries is December 31. The committee will consider books for the award up to and including this date. The dates of the Midwinter Meeting may affect the date the nominations are due, but will be specified on a yearly basis.

Presentation of the Award: The award will be presented at the American Library Association Annual Conference.

Appendix B

Common Core State Standards for English Language Arts

The Standards' Approach to Text Complexity

To help redress the situation described above, the Standards define a three-part model for determining how easy or difficult a particular text is to read as well as grade-by-grade specifications for increasing text complexity in successive years of schooling (Reading standard 10). These are to be used together with grade-specific standards that require increasing sophistication in students' reading comprehension ability (Reading standards 1-9). The Standards thus approach the intertwined issues of what and how student read.

A Three-Part Model for Measuring Text Complexity

As signaled by the graphic at right, the Standards' model of text complexity consists of three equally important parts.

(1) Qualitative dimensions of text complexity. In the Standards, *qualitative dimensions* and *qualitative factors* refer to those aspects of text complexity best measured or only measurable by an attentive human reader, such as levels of meaning or purpose; structure; language conventionality and clarity; and knowledge demands.

(2) Quantitative dimensions of text complexity. The terms *quantitative dimensions* and *quantitative factors* refer to those aspects of text complexity, such as word length or frequency, sentence length, and text cohesion, that are difficult if not impossible for a human reader to evaluate efficiently, especially in long texts, and are thus today typically measured by computer software.

(3) Reader and task considerations. While the prior two elements of the model focus on the inherent complexity of text, variables specific to particular readers (such as motivation, knowledge, and experiences) and to particular tasks (such as purpose and the complexity of the task assigned and the questions posed) must also be considered when determining whether a text is appropriate for a given student. Such assessments are best made by teachers employing their professional judgment, experience, and knowledge of their students and the subject.

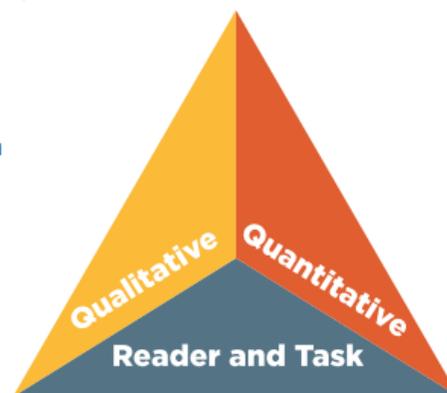


Figure 1: The Standards' Model of Text Complexity

The Standards presume that all three elements will come into play when text complexity and appropriateness are determined. The following pages begin with a brief overview of just some of the currently available tools, both qualitative and quantitative, for measuring text complexity, continue with some important considerations for using text complexity with students, and conclude with a series of examples showing how text complexity measures, balanced with reader and task considerations, might be used with a number of different texts.

Qualitative and Quantitative Measures of Text Complexity

The qualitative and quantitative measures of text complexity described below are representative of the best tools presently available. However, each should be considered only provisional; more precise, more accurate, and easier-to-use tools are urgently needed to help make text complexity a vital, everyday part of classroom instruction and curriculum planning.

Qualitative Measures of Text Complexity

Using qualitative measures of text complexity involves making an informed decision about the difficulty of a text in terms of one or more factors discernible to a human reader applying trained judgment to the task. In the Standards, qualitative measures, along with professional judgment in matching a text to reader and task, serve as a necessary complement and sometimes as a corrective to quantitative measures, which, as discussed below, cannot (at least at present) capture all of the elements that make a text easy or challenging to read and are not equally successful in rating the complexity of all categories of text.

Built on prior research, the four qualitative factors described below are offered here as a first step in the development of robust tools for the qualitative analysis of text complexity. These factors are presented as continua of difficulty rather than as a succession of discrete "stages" in text complexity. Additional development and validation would be needed to translate these or other dimensions into, for example, grade-level- or grade-band-specific rubrics. The qualitative factors run from easy (left-hand side) to difficult (right-hand side). Few, if any, authentic texts will be low or high on all of these measures, and some elements of the dimensions are better suited to literary or to informational texts.

(1) **Levels of Meaning (literary texts) or Purpose (informational texts).** Literary texts with a single level of meaning tend to be easier to read than literary texts with multiple levels of meaning (such as satires, in which the author's literal message is intentionally at odds with his or her underlying message). Similarly, informational texts with an explicitly stated purpose are generally easier to comprehend than informational texts with an implicit, hidden, or obscure purpose.

(2) **Structure.** Texts of low complexity tend to have simple, well-marked, and conventional structures, whereas texts of high complexity tend to have complex, implicit, and (particularly in literary texts) unconventional structures. Simple literary texts tend to relate events in chronological order, while complex literary texts make more frequent use of flashbacks, flash-forwards, and other manipulations of time and sequence. Simple informational texts are likely not to deviate from the conventions of common genres and subgenres, while complex informational texts are more likely to conform to the norms and conventions of a specific discipline. Graphics tend to be simple and either unnecessary or merely supplementary to the meaning of texts of low complexity, whereas texts of high complexity tend to have similarly complex graphics, graphics whose interpretation is essential to understanding the text, and graphics that provide an independent source of information within a text. (Note that many books for the youngest students rely heavily on graphics to convey meaning and are an exception to the above generalization.)

(3) **Language Conventiality and Clarity.** Texts that rely on literal, clear, contemporary, and conversational language tend to be easier to read than texts that rely on figurative, ironic, ambiguous, purposefully misleading, archaic or otherwise unfamiliar language or on general academic and domain-specific vocabulary.

(4) **Knowledge Demands.** Texts that make few assumptions about the extent of readers' life experiences and the depth of their cultural/literary and content/discipline knowledge are generally less complex than are texts that make many assumptions in one or more of those areas.

Levels of Meaning (literary texts) or Purpose (informational texts)

- Single level of meaning → Multiple levels of meaning
- Explicitly stated purpose → Implicit purpose, may be hidden or obscure

Structure

- Simple → Complex
- Explicit → Implicit
- Conventional → Unconventional (chiefly literary texts)
- Events related in chronological order → Events related out of chronological order (chiefly literary texts)
- Traits of a common genre or subgenre → Traits specific to a particular discipline (chiefly informational texts)
- Simple graphics → Sophisticated graphics
- Graphics unnecessary or merely supplementary to understanding the text → Graphics essential to understanding the text and may provide information not otherwise conveyed in the text

Language Conventuality and Clarity

- Literal → Figurative or ironic
- Clear → Ambiguous or purposefully misleading
- Contemporary, familiar → Archaic or otherwise unfamiliar
- Conversational → General academic and domain-specific

Knowledge Demands: Life Experiences (literary texts)

- Simple theme → Complex or sophisticated themes
- Single themes → Multiple themes
- Common, everyday experiences or clearly fantastical situations → Experiences distinctly different from one's own
- Single perspective → Multiple perspectives
- Perspective(s) like one's own → Perspective(s) unlike or in opposition to one's own

Knowledge Demands: Cultural/Literary Knowledge (chiefly literary texts)

- Everyday knowledge and familiarity with genre conventions required → Cultural and literary knowledge useful
- Low intertextuality (few if any references/allusions to other texts) → High intertextuality (many references/allusions to other texts)

Knowledge Demands: Content/Discipline Knowledge (chiefly informational texts)

- Everyday knowledge and familiarity with genre conventions required → Extensive, perhaps specialized discipline-specific content knowledge required

Quantitative Measures of Text Complexity

A number of quantitative tools exist to help educators assess aspects of text complexity that are better measured by algorithm than by a human reader. The discussion is not exhaustive, nor is it intended as an endorsement of one method or program over another. Indeed, because of the limits of each of the tools, new or improved ones are needed quickly if text complexity is to be used effectively in the classroom and curriculum.

Numerous formulas exist for measuring the readability of various types of texts. Such formulas, including the widely used Flesch-Kincaid Grade Level test, typically use word length and sentence length as proxies for semantic and syntactic complexity, respectively (roughly, the complexity of the meaning and sentence structure). The assumption behind these formulas is that longer words and longer sentences are more difficult to read than shorter ones; a text with many long words and/or sentences is thus rated by these formulas as harder to read than a text with many short words and/or sentences would be. Some formulas, such as the Dale-Chall Readability Formula, substitute word frequency for word length as a factor, the assumption here being that less familiar words are harder to comprehend than familiar words. The higher the proportion of less familiar words in a text, the theory goes, the harder that text is to read. While these readability formulas are easy to use and readily available—some are even built into various word processing applications—their chief weakness is that longer words, less familiar words, and longer sentences are not inherently hard to read. In fact, series of short, choppy sentences can pose problems for readers precisely because these sentences lack the cohesive devices, such as transition words and phrases, that help establish logical links among ideas and thereby reduce the inference load on readers.

Like Dale-Chall, the Lexile Framework for Reading, developed by MetaMetrics, Inc., uses word frequency and sentence length to produce a single measure, called a Lexile, of a text's complexity. The most important difference between the Lexile system and traditional readability formulas is that traditional formulas only assign a score to texts, whereas the Lexile Framework can place both readers and texts on the same scale. Certain reading assessments yield Lexile scores based on student performance on the instrument; some reading programs then use these scores to assign texts to students. Because it too relies on word familiarity and sentence length as proxies for semantic and syntactic complexity, the Lexile Framework, like traditional formulas, may underestimate the difficulty of texts that use simple, familiar language to convey sophisticated ideas, as is true of much high-quality fiction written for adults and appropriate for older students. For this reason and others, it is possible that factors other than word familiarity and sentence length contribute to text difficulty. In response to such concerns, MetaMetrics has indicated that it will release the qualitative ratings it assigns to some of the texts it rates and will actively seek to determine whether one or more additional factors can and should be added to its quantitative measure. Other readability formulas also exist, such as the ATOS formula associated with the Accelerated Reader program developed by Renaissance Learning. ATOS uses word difficulty (estimated grade level), word length, sentence length, and text length (measured in words) as its factors. Like the Lexile Framework, ATOS puts students and texts on the same scale.

A nonprofit service operated at the University of Memphis, Coh-Metrix attempts to account for factors in addition to those measured by readability formulas. The Coh-Metrix system focuses on the cohesiveness of a text—basically, how tightly the text holds together. A high-cohesion text does a good deal of the work for the reader by signaling relationships among words, sentences, and ideas using repetition, concrete language, and the like; a low-cohesion text, by contrast, requires the reader him- or herself to make many of the connections needed to comprehend the text. High-cohesion texts are not necessarily "better" than low-cohesion texts, but they are easier to read.

The standard Coh-Metrix report includes information on more than sixty indices related to text cohesion, so it can be daunting to the layperson or even to a professional educator unfamiliar with the indices. Coh-Metrix staff have worked to isolate the most revealing, informative factors from among the many they consider, but these "key factors" are not yet widely available to the public, nor have the results they yield been calibrated to the Standards' text complexity grade bands. The greatest value of these factors may well be the promise they offer of more advanced and usable tools yet to come.

Reader and Task Considerations

The use of qualitative and quantitative measures to assess text complexity is balanced in the Standards' model by the expectation that educators will employ professional judgment to match texts to particular students and tasks. Numerous considerations go into such matching. For example, harder texts may be appropriate for highly knowledgeable or skilled readers, and easier texts may be suitable as an expedient for building struggling readers' knowledge or reading skill up to the level required by the Standards. Highly motivated readers are often willing to put in the extra effort required to read harder texts that tell a story or contain information in which they are deeply interested. Complex tasks may require the kind of information contained only in similarly complex texts.

Numerous factors associated with the individual reader are relevant when determining whether a given text is appropriate for him or her. The RAND Reading Study Group identified many such factors in the 2002 report *Reading for Understanding*:

The reader brings to the act of reading his or her cognitive capabilities (attention, memory, critical analytic ability, inferencing, visualization); motivation (a purpose for reading, interest in the content, self-efficacy as a reader); knowledge (vocabulary and topic knowledge, linguistic and discourse knowledge, knowledge of

CORE STATE STANDARDS for ENGLISH LANGUAGE ARTS & LITERACY IN HISTORY/SOCIAL STUDIES, SCIENCE, AND TECHNICAL SUBJECTS

comprehension strategies); and experiences.

As part of describing the activity of reading, the RAND group also named important task-related variables, including the reader's purpose (which might shift over the course of reading), "the type of reading being done, such as skimming (getting the gist of the text) or studying (reading the text with the intent of retaining the information for a period of time)," and the intended outcome, which could include "an increase in knowledge, a solution to some real-world problem, and/or engagement with the text."⁴

Appendix C

Theodore Seuss Geisel Award Winners and Honor Books 2006–Present

Theodore Seuss Geisel Award Winners and Honor Books 2006-Present

2016 Medal Winners

Don't Throw It to Mo! written by David A. Adler, illustrated by Sam Ricks and published by Penguin Young Readers, an Imprint of Penguin Group (USA) LLC

Honor Books

A Pig, a Fox, and a Box, written and illustrated by Jonathan Fenske, and published by Penguin Young Readers, an Imprint of Penguin Group (USA) LLC.

Waiting, written and illustrated by Kevin Henkes and published by Greenwillow Books, an Imprint of HarperCollins Publisher.

Supertruck, written and illustrated by Stephen Savage, A Neal Porter Book published by Roaring Brook Press, a division of Holtzbrinck Publishing Holdings Limited Partnership

2015 Medal Winners

You Are (Not) Small, written by Anna Kang, illustrated by Christopher Weyant (Two Lions, New York)

Honor Books

Mr. Putter & Tabby Turn the Page, written by Cynthia Rylant, illustrated by Arthur Howard (Houghton Mifflin Harcourt Publishing)

Waiting Is Not Easy! written and illustrated by Mo Willems (Hyperion Books for Children, an imprint of Disney Book Group)

2014 Medal Winner

The Watermelon Seed, written and illustrated by Greg Pizzoli (Disney Hyperion Books, an imprint of Disney Book Group)

Honor Books

Ball, written and illustrated by Mary Sullivan (Houghton Mifflin Books for Children, an imprint of Houghton Mifflin Harcourt)

A Big Guy Took My Ball! written and illustrated by Mo Willems (Hyperion Books for Children, an imprint of Disney Book Group)

Penny and Her Marble, written and illustrated by Kevin Henkes (Greenwillow Books, an imprint of HarperCollins)

2013 Medal Winner

Up, Tall and High!, written and illustrated by Ethan Long (G. P. Putnam's Sons, a division of Penguin Young Readers Group)

Honor Books

Let's Go for a Drive!, written and illustrated by Mo Willems (Hyperion Books for Children, an imprint of Disney Book Group)

Pete the Cat and His Four Groovy Buttons by Eric Litwin, created and illustrated by James Dean (HarperCollins Children's Books, a division of HarperCollins Publishers)

Rabbit & Robot: The Sleepover, written and illustrated by Cece Bell (Candlewick Press)

2012 Medal winner

Tales for Very Picky Eaters, written and illustrated by Josh Schneider (Clarion Books, an imprint of Houghton Mifflin Harcourt Publishing Company)

Honor books

I Broke My Trunk, written and illustrated by Mo Willems (Hyperion Books for Children, an imprint of Disney Book Group)

I Want My Hat Back, written and illustrated by Jon Klassen (Candlewick Press)

See Me Run, written and illustrated by Paul Meisel (Holiday House)

2011 Medal winner

Bink and Gollie, written by Kate DiCamillo and Alison McGhee, illustrated by Tony Fucile (Candlewick Press)

Honor books

Ling & Ting: Not Exactly the Same!, written and illustrated by Grace Lin (Little, Brown and Company, a division of Hachette Book Group, Inc.)

We Are in a Book!, written and illustrated by Mo Willems (Hyperion Books for Children, an imprint of Disney Book Group)

2010 Medal winner

Benny and Penny in the Big No-No! by Geoffrey Hayes (Toon Books, a division of RAW Junior, LLC)

Honor books

I Spy Fly Guy! by Tedd Arnold, (Scholastic, Inc.)

Little Mouse Gets Ready by Jeff Smith (Toon Books, a division of RAW Junior, LLC)

Mouse and Mole: Fine Feathered Friends by Wong Herbert Yee (Houghton Mifflin Books for Children/Houghton Mifflin Harcourt)

Pearl and Wagner: One Funny Day, written by Kate McMullan, illustrated by R.W. Alley (Dial Books for Young Readers)

2009 Medal winner

Are You Ready to Play Outside? by Mo Willems (Hyperion Books for Children, an imprint of Disney Book Group)

Honor books

Chicken Said, 'Cluck!' by Judyann Ackerman Grant, illustrated by Sue Truesdell (HarperCollins Children's Books, a division of HarperCollins Publishers)

One Boy, written and illustrated by Laura Vaccaro Seeger (A Neal Porter Book published by Roaring Brook Press, a division of Holtzbrinck Publishing Holdings Limited Partnership)

Stinky, written and illustrated by Eleanor Davis (RAW Junior/TOON Books)

Wolfsnail: A Backyard Predator by Sarah C. Campbell, photographs by Sarah C. Campbell and Richard P. Campbell (Boyd's Mills Press)

2008 Medal winner

There Is a Bird on Your Head by Mo Willems (Hyperion)

Honor books

First the Egg by Laura Vaccaro Seeger (Roaring Brook/Neal Porter)

Hello, Bumblebee Bat, written by Darrin Lunde, illustrated by Patricia J. Wynne (Charlesbridge)

Jazz Baby, written by Lisa Wheeler, illustrated by R. Gregory Christie (Harcourt)

Vulture View, written by April Pulley Sayre, illustrated by Steve Jenkins (Holt)

2007**Medal winner**

Zelda and Ivy: The Runaways by Laura McGee Kvasnosky (Candlewick)

Honor books

Mercy Watson Goes for a Ride written by Kate DiCamillo and illustrated by Chris Van Dusen (Candlewick)

Move Over, Rover! by Karen Beaumont and illustrated by Jane Dyer (Harcourt)

Not a Box by Antoinette Portis (HarperCollins)

2006 Medal winner

Henry and Mudge and the Great Grandpas written by Cynthia Rylant and illustrated by Suçie Stevenson (Simon & Schuster Books for Young Readers)

Honor books

Hi! Fly Guy by Tedd Arnold (Cartwheel Books, an imprint of Scholastic Inc.)

A Splendid Friend, Indeed by Suzanne Bloom (Boyd's Mills Press)

Cowgirl Kate and Cocoa written by Erica Silverman and illustrated by Betsy Lewin (Harcourt, Inc.)

Amanda Pig and the Really Hot Day written by Jean Van Leeuwen and illustrated by Ann Schweninger (Dial Books for Young Readers, a division of Penguin Young Readers Group)

Appendix D

Natural Language Took Kit (NLTK) Stoplist

Appendix E**Title Author Illustrator Data**

Title, Author, Publisher

Title	Author	Illustrator	Publisher	Copyright
A Big Guy Took My Ball!	Willems, Mo		Hyperion	2013
A Pig, a Fox, and a Box	Fenske, Jonathan		Penguin	2015
A Splendid Friend, Indeed	Bloom, Suzanne		Boyd's Mills Press	2005
Amanda Pig and the Really Hot Day	Van Leeuwen, Jean	Schweninger, Ann	Dial/Penguin	2005
Are You Ready to Play Outside?	Willems, Mo		Hyperion	2008
Ball	Sullivan, Mary		Houghton Mifflin	2013
Benny and Penny in the Big No-No!	Hayes, Geoffrey		Toon Books	2009
Bink and Gollie	DiCamillo, Kate	Fucile, Tony,	Candlewick Press	2010
Chicken Said, "cluck!"	Grant, Judyann	Truesdell, Sue	Candlewick Press	2008
Cowgirl Kate and Cocoa	Silverman, Erica	Lewin, Betsy	HarperCollins	2005
Don't Throw it, Moi!	Adler, David A.	Ricks, Sam	Harcourt	2015
First the Egg	Seeger, Laura Vaccaro		Penguin	2007
Hello, Bumblebee Bat	Lunde, Darrin P.		Roaring Brook Press	2007
Henry and Mudge and the Great Grandpas	Rylant, Cynthia		Charlesbridge	2007
Hil Fly Guy	Arnold, Tedd		Simon & Schuster	2005
I Broke My Trunk	Willems, Mo		Scholastic	2011
I Spy Fly Guy!	Arnold, Tedd		Hyperion	2009
I Want My Hat Back	Klassen, J		Scholastic	2011
Jazz Baby	Weatherford, Carole Bost	Freeman-Hines, Laura	Candlewick Press	2011
Let's Go for a Drive!	Willems, Mo		Lee & Low Books	2002
Ling & Ting: Not Exactly the Same!	Lin, Grace		Hyperion	2012
Little Mouse Gets Ready	Smith, Jeff		Little, Brown	2010
Mercy Watson Goes for a Ride	DiCamillo, Kate		Toon Books	2009
Mouse and Mole: Fine Feathered Friends	Yee, Wong Herbert	Van Dusen, Chris	Candlewick Press	2006
Move Over, Rover!	Baumont, Karen	Dyer, Jane	Houghton Mifflin	2009
Mr. Putter and Tabby Turn the Page	Rylant, Cynthia	Howard, Arthur	Scholastic	2006
Not a Box	Portis, Antoinette		Houghton Mifflin	2014
One Boy	Seeger, Laura Vaccaro		HarperCollins	2006
Pearl and Wagner: One Funny Day	McMullan, Kate		Roaring Brook Press	2008
Penny and Her Marble	Henkes, Kevin		Dial/Penguin	2009
Pete the Cat and His Four Groovy Buttons	Litwin, Eric		Greenwillow/Harper	2013
Rabbit & Robot: The Sleepover	Bell, Cece		Harper	2012
See Me Run	Meisel, Paul	Dean, James	Candlewick	2011
Stinky	Davis, Eleanor		Holiday House	2011
Supertruck	Savage, Stepher		Toon Books	2008
Tales for Very Picky Eaters	Schneider, Josh		Roaring Brook Press	2015
The Watermelon Seed	Pizzoli, Greg		Clarion Books	2011
There is a Bird on Your Head	Willems, Mo		Disney Hyperion	2013
Up, Tail and High!	Long, Ethan		Hyperion	2007
Vulture View	Sayre, April Pulley		G.P. Putnam's Sons	2012
Waiting	Henkes, Kevin	Jenkins, Steve	Henry Holt	2007
Waiting is Not Easy!	Willems, Mo		Greenwillow/Harper Colli	2015
We Are in a Book!	Willems, Mo		Hyperion/Disney	2014
Wolfsnail: A Backyard Predator	Campbell, Sarah C.		Hyperion	2010
You Are (Not) Small	Kang, Anna		Boyd's Mills Press	2008
Zelda and Ivy: The Runaways	Kvasnosky, Laura McGee	Weyant, Christopher	Two Lions	2014
			Candlewick Press	2006

Appendix F
Genre and Subject

Genre and Subject

Title	Summary	Subject Heading	Commercial Level		Non Fiction	
			Level		Fiction	
A Big Guy Took My Ball!	Piggie is upset because a whale took the ball she found, but Gerald finds a solution that pleases all of them.	Elephants-JF, Swine-JF, Whales-JF, Friendship-JF, Play-JF, Elephants-F, Pigs-F, Whales-F, Animals-F, Friendship-F, Play-F				x
A Pig, a Fox, and a Box	After finding a box just the right size to hide in, a little fox tries to play some tricks on his big friend, Pig, but things do not work out exactly as he planned. When a studious polar bear meets an inquisitive goose, they learn to be friends	Foxes--Fiction. Pigs--Fiction. Boxes--Fiction. Friendship--Fiction. Humorous stories.	Level 2 Progressing Reader			x
A Splendid Friend, Indeed	Amanda Pig and her family and friends try to find different ways to beat the heat. Friends Elephant and Piggie are playing outside when it starts to rain, and then they must decide what to do.	Friendship--F, Individuality--F, Geese--F, Polar bear--F	Penguin Young Readers Level 3 transitional Reader			x
Amanda Pig and the Really Hot Day	While searching for someone to play ball with him, a dog dreams of fantastical adventures he could have with his ball.	Pigs-F, Heat--F, Family life--F	Reader			x
Are You Ready to Play Outside?	Two mice meet their new neighbor and discover that she is not as scary as they feared.	Rain and rainfall-F, Friendship-F, Elephants-F, Pigs-F				x
BALL		Children's Stories				x
Benny and Penny in the Big No-No!		Graphic novels, Mice-F, Brothers and sisters-F, Neighbors-F.	Toon into Reading Level 2			x

Title	Summary	Subject Heading	Commercial Level	Fiction	Non Fiction
Bink and Gollie	Two roller-skating best friends-one tiny, one tall-- share three comical adventures involving outrageously bright socks, an impromptu trek to the Andes, and a most unlikely marvelous companion.	Friendship-F, Humorous stories		x	
Chicken Said, "Cluck!"	Earl and Pearl do not want Chicken's help in the garden, until a swarm of grasshoppers arrives and her true talent shines.	Chickens-F, Gardening-F, Grasshoppers-F	I can Read My first Shared Reading	x	
Cowgirl Kate and Cocoa	Cowgirl Kate and her cowhorse Cocoa, who is always hungry, count cows, share a story, and help each other fall asleep.	Cowgirls--F, Horses--F	Others in the series published by Green Light Level 2	x	
Don't Throw it to, Mo!	Mo is the youngest kid on the Robins football team. The kids on the rival team tease him for being a 'butterfingers' who's too tiny to catch the ball. But Mo's coach has a plan up his sleeve to turn Mo's little size into a big win for the Robins"-- Provided by publisher.	Football--Fiction. Size--Fiction.	Level 2 Progressing Reader	x	
First the Egg	A picture book about transformations. Simple text and illustrations introduce the endangered bumblebee bat of Thailand	Developmental biology--JL, Developmental biology			x
Hello, Bumblebee Bat	introduce the endangered bumblebee bat of Thailand	Bumblebee bat -- JL, Bumblebee bat, Bats			x

Title	Summary	Subject Heading	Commercial Level	Fiction	Non Fiction
Henry and Mudge and the Great Grandpas	When Henry and his dog Mudge go with Henry's parents to visit Great-Grandpa Bill in the home with lots of other grandpas, they lead them all on a wonderful adventure	Old age--F, Dogs--F	Level 2 Ready to Read	x	
Hi! Fly Guy	When Buzz captures a fly to enter in the Amazing Pet Show, his parents and the judges tell him that a fly cannot be a pet, but Fly Guy proves them wrong.	Flies--F, Pet shows-F	Level 2 Developing Reader 25--750 Words		x
I Broke My Trunk	Gerald the elephant tells his best friend Piggie a long, crazy story about how he broke his trunk. While playing hide-and-seeek with Buzz, Fly Guy is taken away by a garbage man.	Storytelling-F, Elephants-F, Pigs-F			x
I Spy Fly Guy!	A bear almost gives up his search for his missing hat until he remembers something important.	Flies-F, Hide-and-seeek-F	Other leveled by Scholastic Level 2		x
I Want My Hat Back	A bear almost gives up his search for his missing hat until he remembers something important.	Bears-F, Hats-F, Lost and found possessions-F			x
Jazz Baby	A group of children move and play, hum and sleep to a jazz beat.	Play--F, Jazz--F, Sorries in rhyme			x

Title	Summary	Subject Heading	Commercial Level		Non Fiction	
			Level		Fiction	
Let's Go for a Drive!	Elephant Gerald and Piggie want to go for a drive, but as Gerald thinks of one thing after another that they will have to take along, they come to realize that they lack the most important thing of all.	Automobile travel-F, Elephants-F, Pigs-F, JF/Animals/Elephants, JF/Animals/Pigs, JF/Social Issues/Friendship				x
Ling & Ting: Not Exactly the Same!	Ling and Ting are identical twins that people think are exactly the same, but time and again they prove to be different. Little Mouse gets dressed to go to the barn with his mother, brothers, and sisters. Mr. Watson's usual Saturday drive in his Cadillac with his favorite pig, Mercy, turns into an adventure when an unexpected passenger shows up in the back seat and Mercy finds herself behind the wheel.	Twins-F, Sisters-F, Individuality-F, Chinese Americans-F	Passport to Reading 3			x
Little Mouse Gets Ready		Graphic novels, Mice-F, Clothing and dress-F	Too into Reading Level 1			x
Mercy Watson Goes for a Ride		Pigs--F, Automobile driving--F, Humorous stories				x
Mouse and Mole: Fine Feathered Friends	When Spring arrives, Mole and Mouse find a unique way to bird watch.	Bird-watching-F, Birds-F, Mice-F, Moles(Animals)-F, Spring-F	Green Light Readers Level 3			x

Title	Summary	Subject Heading	Commercial Level	Fiction	Non Fiction
Move Over, Rover!	<p>When a storm comes, Rover expects to have his dog house all to himself. But finds that various other animals, including a skunk, come to join him period. It's raining cats and dogs! Good thing Rover is snuggled safe and dry inside his dog house-until, one-by-one, a soggy menagerie of creatures shows up looking for a cozy place to sit out the storm. But who's the very unwelcome surprise visitor? Skunk, of course. Suddenly, that dog house isn't quite so crowded after all!</p>	<p>Dogs-JF, Animals-JF, Animal housing-JF, Dogs-F, Animals-F, Animal housing-F, Stories in Rhyme</p>	x	x	
Mr. Putter and Tabby Turn the Page	<p>Mr Putter and Mrs. Teaberry bring Tabby and Zeke to the library for a special storytime. To an imaginative bunny, a box is not always just a box.</p>	<p>Books and reading-F, Libraries-F, Cats-F, Dogs-F</p>	x	x	
Not a Box	<p>A boy creates ten paintings in this counting book that also explores the relationship of words within words. Features die-cut pages.</p>	<p>Boxes--F, Rabbits--F</p>			x
One Boy	<p>A boy creates ten paintings in this counting book that also explores the relationship of words within words. Features die-cut pages.</p>	<p>Toy and moveable books-Specimens, Vocabulary-F. Painting-F, Counting, Toy and movable books</p>			x

Title	Summary	Subject Heading	Commercial Level	Fiction	Non Fiction
Pearl and Wagner: One Funny Day	April Fool's Day is not a happy one for Wagner the mouse because his best friend, Pearl the rabbit, and other children and adults at school keep tricking him.	April Fools' Day-F, Practical jokes-F, Schools-F, Rabbits-f, Mice-F, Animals-F	Penguin Young Readers Level 3 Transitional Reader	x	
Penny and Her Marble	Penny feels guilty after taking a beautiful blue marble that she sees in Mrs. Goodwin's grass, but gets a pleasant surprise when she goes to return it the next day.	Lost and found possessions-F, Marbles-F, Mice-F	I can Read 1 Beginning Reading		x
Pete the Cat and His Four Groovy Buttons	Pete the Cat loves the buttons on his shirt so much that he makes up a song about them, and even as the buttons pop off, one-by-one, he still finds a reason to sing.	Cats-F, Buttons-F, Singing-F, Counting.			x
Rabbit & Robot: The Sleepover	Rabbit is excited about the sleepover he has carefully planned for his friend Robot, but Robot has some different ideas about how things should go. A dog has a fun-filled day at the dog park, in this easy-to-read story.	Sleepovers-F, Rabbits-F, Robots-F, Friendship-F, Humorous stories, JF/Humorous Stories, JF/Social Issues/Friendship, JF/General	Candlewick Sparks		x
See Me Run		Dogs-F, Parks-F			x

Appendix G

Readability

Readability

Title	Readability			Counts		
		Lexile	ATOS	Pages	Chapters	TOC
A Big Guy Took My Ball!	AD	80	1	57		
A Pig, a Fox, and a Box	BR	BR	1.3	32		
A Splendid Friend, Indeed	BR	BR	0.7	32		
Amanda Pig and the Really Hot Day		290	L*	2.2	48	4 x
Are You Ready to Play Outside?		140	L	0.6	57	
Ball		770	L*	0.5 *	30	
Benny and Penny in the Big No-No!	GN	30	L	1.3	32	
Bink and Gollie		310	L	2.5	81	3 x
Chicken Said, "Cluck!"		70	L	1.6	30	
Cowgirl Kate and Cocoa		400	L	2.3	44	4 x
Don't Throw it, Mo!		270	L	1.9 *	32	
First the Egg	NP	NP		3.2 *	27	
Hello, Bumblebee Bat		430	L*	1.9	25	
Henry and Mudge and the Great Grandpas		460	L	2.6	40	4 x
Hi! Fly Guy		280	L	1.5	30	3
I Broke My Trunk		290	L*	1.3	57	
I Spy Fly Guy!		310	L	1.5	30	3
I Want My Hat Back		90	L	1	32	
Jazz Baby	AD	400	L	1.2	30	
Let's Go for a Drive!		240	L*	1	57	
Ling & Ting: Not Exactly the Same!		390	L	1.8	43	6 x
Little Mouse Gets Ready	GN	160	L	1.3	30	
Mercy Watson Goes for a Ride		390	L	2.7	74	14
Mouse and Mole: Fine Feathered Friends	AD	360	L	2.6	48	4 x
Move Over, Rover!		270	L	1.8	31	
Mr. Putter and Tabby Turn the Page		520	L	3.1	40	5 x
Not a Box	AD	50	L*	0.8 *	30	
One Boy		130	L*	2.8 *	42	
Pearl and Wagner: One Funny Day		210	L	2.1	58	3 x
Penny and Her Marble		350		2.5	48	4
Pete the Cat and His Four Groovy Buttons	AD	350	L	1.9	34	
Rabbit & Robot: The Sleepover		360	L	2.8	50	4 x
See Me Run	BR	BR		0.3	24	
Stinky	GN	170	L*	1.5	40	3
Supertruck	AD	370	L	1.5	30	
Tales for Very Picky Eaters	AD	500	L	2.3	47	5 x
The Watermelon Seed		310	L*	1	40	
There is a Bird on Your Head		210	L	1	57	
Up, Tall and High!		60	L*	0.6 *	34	3
Vulture View		70	L*	1.1	30	
Waiting		260	L	1.9	32	
Waiting Is Not Easy!	BR	BR		0.9	57	
We Are in a Book!	BR	BR		0.9	57	
Wolfsnail: A Backyard Predator		540	L	4.4	32	
You Are (Not) Small	BR	BR		0.5 *	32	
Zelda and Ivy: The Runaways		470	L	3	42	3 x

Appendix H

Lexile Codes



Also see this article (<http://www.edutopia.org/blog/graphic-novels-comics-andrew-miller>) from Edutopia about the instructional value of using graphic novels and comics in the classroom.

[Back to Top](#)

BR: Beginning Reader

Beginning Reader (BR) is a code given to readers and text that are below 0L on the Lexile scale. In some cases, for readers, a BR code is followed by a number and L (e.g., BR150L). A Lexile reader measure of BR150L indicates that the Lexile measure of the reader is 150 units below 0L. The smaller the number following the BR code, the more advanced the reader is. For example, a BR150L reader is more advanced than a BR200L reader. Unlike the reader measure, all text measures below 0L are currently reported as BR. MetaMetrics has conducted research to differentiate the BR text measures, and these measures will be available at a later date.

Note that Beginning Reader (BR) is the only Lexile code that applies to both readers and text. All other codes apply only to text.

Hop on Pop by Dr. Seuss (Random House) is a BR book.

Birds of Prey by Dr. Gerald Legg (Franklin Watts Library) is coded IG. Separate paragraphs are arranged upon the page, functioning more like multiple-sentence captions. A particular reading order is neither indicated by the layout nor important to comprehension. Thus the book measure is IG320L.

What Do Birds of Prey Eat?



Birds of prey will eat any other animal, including other birds. Insects such as wasp grubs are a favorite of the honey buzzard. Secretary birds like locusts, and kestrels will catch beetles. Some eagles eat other birds of prey.



Did You Know?

Without a large number of hawks is called a "kettle." Settles of the Saw-whet hawk in the woods. The hawks are, fields because they can find plenty of birds to capture.

mammals, tortoises, antelopes, monkeys, snakes, lizards, even dead elephants and whales, are all on the menu.

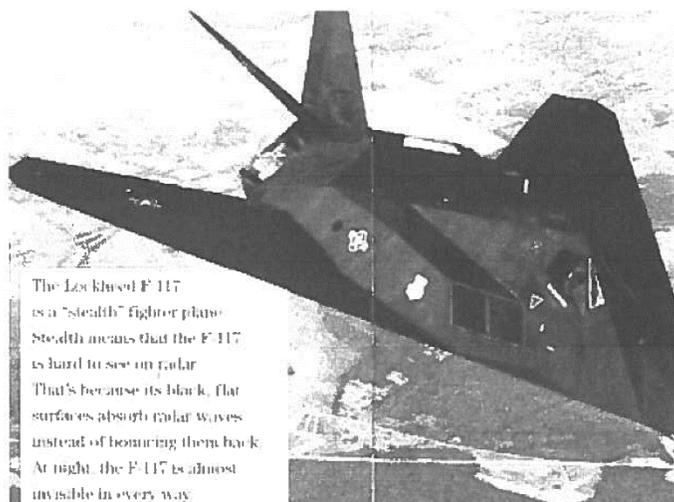


Frogs and mice are prey for many birds and mammals. These birds also eat mice, snakes, and insects.

[Back to Top](#)

GN: Graphic Novel

The GN code indicates that the book is a graphic novel or comic book. The text of GN books appears primarily in voice or thought bubbles integrated into comic book-style illustrations. Graphic novels tend to contain a larger percentage of dialogue than most other genres of books. They also typically lack some of the required text conventions of dialogue, such as putting "she said" after a quoted sentence, because illustration methods are used to indicate spoken text. The impact of picture support on reading comprehension is not captured in the Lexile measure of a graphic novel. *To Dance: A Ballerina's Graphic Novel (Aladdin)*, written by Siena Cherson Siegel and illustrated by Mark Siegel, is coded as GN610L.



The Lockheed F-117 is a "stealth" fighter plane. Stealth means that the F-117 is hard to see on radar. That's because its black, flat surfaces absorb radar waves instead of bouncing them back. At night, the F-117 is almost invisible in every way.

[Back to Top](#)

HL: High-Low

A text designated as "HL" has a Lexile measure much lower than the average reading ability of the intended age range of its readers. Librarians and booksellers sometimes refer to young adult books with disproportionately low Lexile measures as "high-low" books, meaning "high-interest" plus "low-readability." These books receive an HL code. Often fiction, HL books are useful when matching older (grade 7 and beyond) struggling or reluctant readers with text at both an appropriate difficulty level and an appropriate developmental level.

Despite their short sentences and basic vocabulary, HL books are designed to appeal to readers at a more mature developmental level. For example, Beth Goobie's *Sticks and Stones* (Orca Soundings) is classified as a young adult book and measures 430L—an average reading ability for 2nd graders. The book's characters are high-school students who struggle with the many challenges that face high-school students such as dating and gossip. Therefore, the book is coded HL430L.

[Back to Top](#)

IG: Illustrated Guide

The IG code is applied to books that consist of independent pieces or sections of text such as in an encyclopedia or glossary. These text pieces could be moved around without affecting the overall linear flow of the book. Usually nonfiction, IG books are often used as a reference resource rather than read in their entirety like a storybook. Their distinguishing text characteristics include:

- technical vocabulary, definitions, and pronunciation guides in parentheses or contrasting type
- integration of illustrations and diagrams into the text
- pull-quotes, factoids, and other categorical marginalia
- the presentation of each discrete topic on one to two pages

These text characteristics do not necessarily impact reading comprehension or developmental appropriateness. Instead, the IG code conveys an idea of the kind of book and what the book typically will be used for in the classroom or library.

"private," "gnash," and "rumpus." The parent on the couch would help the preschooler sound these words out and decipher these long sentences. Therefore the book is coded adult directed and the measure is AD740L.



of staring into all their yellow eyes without blinking once
and they were frightened and called him the most wild thing of all

Additionally, picture books can have design elements that may visually complicate reading for a child. Factors such as font size, typeface, page layout, legibility, and the relationship between pictures and text may significantly impact reading comprehension. The story and illustrations in *Where the Wild Things Are* are perfect for young children. But the lines of the text are close together and the sentences are spread over multiple pages, often in long horizontal lines. These design elements may challenge a child's ability to read the book independently even if the text difficulty is well matched. Initially, a more advanced reader may need to read the book with a child.

[Back to Top](#)

NC: Non-Conforming

The NC code is applied to books that have a Lexile measure markedly higher than is typical for the publisher's intended audience or designated developmental level of the book. The Lexile measure of a book is compared to the Lexile range of readers in the intended audience in order to make an NC code determination.

The NC code is useful when matching high-ability readers with a book that's still at an appropriate developmental level. Alternatively, some picture books with disproportionately high Lexile measures may receive an AD (Adult Directed) code.

Seymour Simon's *Amazing Aircraft* (SeaStar Books) is coded NC710L. Its spine reads "grades 1-3" but its Lexile measure is higher than a typical early elementary school student's ability range. Therefore the book is coded as Non-Conforming.

THING SING

That thing can sing!



[Back to Top](#)

NP: Non-Prose

The NP code is for any book comprising more than 50% non-standard or non-conforming prose. NP books do not receive a Lexile measure, merely the NP code. Some common examples of non-prose content are poems, plays, songs, recipes, and text with non-standard or absent punctuation. Since the Lexile Framework is based on prose analysis, Maurice Sendak's *Alligators All Around* (HarperTrophy) is coded NP. The text of the book is not in complete sentences and lacks punctuation entirely. The text difficulty of such a book cannot currently be assigned a Lexile measure.



P pushing people



Q quite quirksome

[Back to Top](#)

Appendix I
Word Analysis

Word Analysis

BOOK NAME	Words Gross	Words Net	Words Gross Non- Stopwords	Words Net Non- Stopwords	Stopwords Gross	Stopwords Net
A Big Guy Took My Ball	222	85	94	42	128	43
A Pig, a Fox and a Box	349	99	167	62	182	37
A Splendid Friend, Indeed	102	38	42	25	60	13
Amanda Pig and the Really Hot Day	1568	381	905	302	663	79
Are You Ready to Play Outside	170	82	90	52	80	30
BALL	64	6	58	5	6	1
Benny and Penny and the Big No	632	215	312	151	320	64
Bink and Gollie	855	267	523	194	332	73
Chicken said Cluck!	181	39	159	31	22	8
Cowgirl Kate and Cocoa	999	292	535	222	464	70
Dont throw it to Mo	430	140	232	95	198	45
First the Egg	76	26	33	17	43	9
Hello, Bumblebee Bat	275	139	151	99	124	40
Henry and Mudge and the Great Grandpas	675	233	385	175	290	58
Hi! Fly Guy	347	118	201	84	146	34
I Broke My Trunk	263	90	111	47	152	43
I Spy Fly Guy!	302	122	181	84	121	38
I Want My Hat Back	263	88	117	51	146	37
Jazz Baby	224	102	186	90	38	12
Let's Go for a Drive!	279	92	150	59	129	33
Ling and Ting	881	238	481	171	400	67
Little Mouse Gets Ready	278	130	142	88	136	42
Mercy Watson Goes for a Ride	2230	513	1302	421	928	92
Mouse and Mole	1889	533	1176	456	713	77
Move Over, Rover!	333	102	191	76	142	26
Mr. Putter and Tabby Turn the Page	802	227	474	169	328	58
Not a Box	77	30	21	11	56	19
One Boy	54	44	37	35	17	9
Pearl and Wagner One Funny Day	863	292	559	229	304	63
Penny and Her Marble	898	272	512	204	386	68
Pete the Cat and His Four Groovy Buttons	304	83	189	53	115	30
Rabbit and Robot	1707	425	973	350	734	75
See Me Run	77	31	33	15	44	16
Stinky	989	292	586	230	403	62
Supertruck	97	56	51	38	46	18
Tales for Very Picky Eaters	1108	349	577	274	531	75
The Watermelon Seed	146	82	74	54	72	28
There Is a Bird on Your Head!	242	75	112	43	130	32
Up! Tall! and High! (but not necessarily in that order.)	96	36	34	15	62	21
Vulture View	180	89	94	64	86	25
Waiting	274	112	124	76	150	36
Waiting Is Not Easy!	197	77	82	43	115	34
We Are in a Book!	253	93	130	56	123	37
Wolfsnail	390	186	215	142	175	44
You Are (Not) Small	92	21	32	9	60	12
Zelda and Ivy the Runaways	1042	391	588	311	454	80
AVERAGES	517	162	292	120	225	42
Median	279	102	163	76	139	37

Appendix J

Sentence Analysis

Sentence Analysis

BOOK NAME	# SENTENCES	AVG Words / Sentence	AVG Words	Ratio
			/ Sentence NS	AVGN / AVGW
A Big Guy Took My Ball	46	4.826	2.293	0.475
A Pig, a Fox and a Box	75	4.653	2.288	0.492
A Splendid Friend, Indeed	26	3.923	1.909	0.487
Amanda Pig and the Really Hot Day	242	6.479	3.755	0.580
Are You Ready to Play Outside	47	3.617	1.957	0.541
BALL	44	1.455	1.318	0.906
Benny and Penny and the Big No	151	4.185	2.277	0.544
Bink and Gollie	132	6.477	4.023	0.621
Chicken said Cluck!	46	3.935	3.457	0.878
Cowgirl Kate and Cocoa	159	6.283	3.365	0.536
Don't Throw It to Mo	70	6.143	3.314	0.540
First the Egg	32	2.375	1.500	0.632
Hello, Bumblebee Bat	49	5.612	3.082	0.549
Henry and Mudge and the Great Grandpas	87	7.759	4.425	0.570
Hil Fly Guy	74	4.689	2.716	0.579
I Broke My Trunk	53	4.962	2.413	0.486
I Spy Fly Guy!	56	5.393	3.232	0.599
I Want My Hat Back	48	5.479	2.438	0.445
Jazz Baby	60	3.733	3.153	0.844
Let's Go for a Drive!	62	4.500	2.459	0.546
Ling and Ting	144	6.118	3.387	0.554
Little Mouse Gets Ready	49	5.673	3.227	0.569
Mercy Watson Goes for a Ride	350	6.371	3.752	0.589
Mouse and Mole	311	6.074	3.781	0.623
Move Over, Rover!	63	5.286	3.081	0.583
Mr. Putter and Tabby Turn the Page	116	6.914	4.086	0.591
Not a Box	12	6.417	1.750	0.273
One Boy	22	2.455	1.682	0.685
Pearl and Wagner One Funny Day	163	5.294	3.451	0.652
Penny and Her Marble	131	6.855	3.938	0.575
Pete the Cat and His Four Groovy Buttons	73	4.164	2.589	0.622
Rabbit and Robot	268	6.369	3.686	0.579
See Me Run	22	3.500	1.737	0.496
Stinky	276	3.583	2.245	0.627
Supertruck	16	6.063	3.188	0.526
Tales for Very Picky Eaters	136	8.147	4.438	0.545
The Watermelon Seed	31	4.710	2.387	0.507
There Is a Bird on Your Head!	50	4.840	2.286	0.472
Up! Tall! and High! (but not necessarily in that order.)	26	3.692	1.619	0.438
Vulture View	55	3.273	2.293	0.701
Waiting	39	7.026	3.543	0.504
Waiting Is Not Easy!	51	3.863	1.783	0.461
We Are in a Book!	86	2.942	1.605	0.546
Wolfsnail	45	8.667	4.886	0.564
You Are (Not) Small	22	4.182	1.455	0.348
Zelda and Ivy the Runaways	131	7.954	4.523	0.569
AVERAGES	92	5.150	2.865	0.566
Median	58	5.124	2.898	0.559
Minimum	12	1.455	1.318	0.273
Maximum	350	8.667	4.886	0.906

Appendix K
Parts of Speech

Parts of Speech (Whole Text)

BOOK NAME	ADJECTIVES	NOUNS	VERB	PRONOUNS	ADVERBS
A Big Guy Took My Ball	30	56	65	34	22
A Pig, a Fox and a Box	27	79	92	57	16
A Splendid Friend, Indeed	5	12	42	29	7
Amanda Pig and the Really Hot Day	120	635	448	134	94
Are You Ready to Play Outside	9	57	56	24	13
BALL	0	51	6	0	0
Benny and Penny and the Big No	33	204	144	78	37
Bink and Gollie	53	430	252	53	44
Chicken said Cluck!	6	126	44	0	1
Cowgirl Kate and Cocoa	52	376	275	141	45
Don't Throw It to Mo	21	125	122	47	20
First the Egg	0	25	17	0	15
Hello, Bumblebee Bat	22	81	73	38	16
Henry and Mudge and the Great Grandpas	31	257	154	42	28
Hi! Fly Guy	22	154	91	26	7
I Broke My Trunk	17	59	76	47	21
I Spy Fly Guy!	7	129	76	24	14
I Want My Hat Back	7	49	85	63	21
Jazz Baby	23	97	23	6	4
Let's Go for a Drive!	13	94	61	36	7
Ling and Ting	77	319	291	77	18
Little Mouse Gets Ready	27	82	61	18	16
Mercy Watson Goes for a Ride	78	743	676	244	169
Mouse and Mole	121	660	461	166	89
Move Over, Rover!	4	133	59	3	13
Mr. Putter and Tabby Turn the Page	54	280	220	60	60
Not a Box	0	22	23	7	10
One Boy	0	41	0	0	0
Pearl and Wagner One Funny Day	34	433	242	75	41
Penny and Her Marble	45	328	261	85	67
Pete the Cat and His Four Groovy Buttons	19	99	72	44	18
Rabbit and Robot	61	551	504	186	95
See Me Run	4	10	27	16	3
Stinky	51	384	199	102	1
Supertruck	6	29	22	9	5
Tales for Very Picky Eaters	73	380	303	96	82
The Watermelon Seed	7	48	33	22	11
There Is a Bird on Your Head!	5	66	68	41	18
Up! Tall! and High! (but not necessarily in that order.)	12	10	36	20	15
Vulture View	8	52	59	6	24
Waiting	16	53	85	25	29
Waiting Is Not Easy!	8	30	72	35	21
We Are in a Book!	10	84	72	26	16
Wolfsnail	21	126	77	28	14
You Are (Not) Small	22	5	34	22	9
Zelda and Ivy the Runaways	55	384	261	103	47
(ALL the Books Combined) AVERAGES	29	184	140	52	29
Median	21	96	75	36	17
Minimum	0	5	0	0	0
Maximum	121	743	676	244	169

Appendix L

Lexical Variation

Lexical Variation

BOOKNAME	TTR	Verb TTR	Noun TTR	Adjective TTR	Adverb TTR
A Big Guy Took My Ball	0.40	0.12	0.48	0.06	0.08
A Pig, a Fox and a Box	0.27	0.10	0.28	0.07	0.05
A Splendid Friend, Indeed	0.35	0.05	0.64	0.04	0.07
Amanda Pig and the Really Hot Day	0.22	0.19	0.36	0.08	0.03
Are You Ready to Play Outside	0.47	0.12	0.60	0.07	0.07
BALL	0.19	2.00	0.12	0.00	0.00
Benny and Penny and the Big No	0.32	0.17	0.41	0.05	0.06
Bink and Gollie	0.27	0.15	0.33	0.10	0.04
Chicken said Cluck!	0.22	0.04	0.22	0.05	0.01
Cowgirl Kate and Cocoa	0.26	0.27	0.35	0.09	0.03
Don't Throw It to Mo	0.28	0.10	0.34	0.05	0.06
First the Egg	0.38	1.00	0.52	0.00	0.05
Hello, Bumblebee Bat	0.47	0.14	0.52	0.13	0.06
Henry and Mudge and the Great Grandpas	0.32	0.25	0.36	0.06	0.05
Hi! Fly Guy	0.31	0.14	0.33	0.08	0.04
I Broke My Trunk	0.32	0.05	0.35	0.05	0.07
I Spy Fly Guy!	0.37	0.13	0.44	0.05	0.05
I Want My Hat Back	0.31	0.10	0.35	0.04	0.08
Jazz Baby	0.47	0.05	0.55	0.12	0.03
Let's Go for a Drive!	0.36	0.05	0.50	0.08	0.02
Ling and Ting	0.24	0.15	0.35	0.07	0.03
Little Mouse Gets Ready	0.46	0.16	0.62	0.08	0.08
Mercy Watson Goes for a Ride	0.18	0.38	0.21	0.04	0.05
Mouse and Mole	0.23	0.66	0.27	0.05	0.04
Move Over, Rover!	0.30	0.07	0.36	0.05	0.02
Mr. Putter and Tabby Turn the Page	0.26	0.16	0.27	0.05	0.05
Not a Box	0.44	0.09	0.30	0.00	0.13
One Boy	0.82	0.04	0.92	0.03	0.03
Pearl and Wagner One Funny Day	0.32	0.31	0.37	0.10	0.04
Penny and Her Marble	0.27	0.26	0.31	0.07	0.06
Pete the Cat and His Four Groovy Buttons	0.29	0.07	0.21	0.05	0.04
Rabbit and Robot	0.20	0.38	0.30	0.04	0.03
See Me Run	0.44	0.03	0.67	0.09	0.06
Stinky	0.25	0.23	0.33	0.06	0.04
Supertruck	0.54	0.06	0.64	0.10	0.05
Tales for Very Picky Eaters	0.29	0.37	0.40	0.11	0.04
The Watermelon Seed	0.55	0.09	0.67	0.09	0.12
There Is a Bird on Your Head!	0.29	0.06	0.24	0.03	0.06
Up! Tall! and High! (but not necessarily in that order.)	0.36	0.01	0.73	0.06	0.12
Vulture View	0.48	0.20	0.60	0.07	0.04
Waiting	0.40	0.10	0.58	0.08	0.13
Waiting Is Not Easy!	0.36	0.04	0.64	0.10	0.12
We Are in a Book!	0.35	0.12	0.35	0.07	0.07
Wolfsnail	0.45	0.29	0.52	0.08	0.09
You Are (Not) Small	0.26	0.01	1.00	0.10	0.05
Zelda and Ivy the Runaways	0.34	0.50	0.47	0.07	0.04
Average	0.35	0.22	0.44	0.07	0.06
Median	0.32	0.13	0.37	0.07	0.05
Minimum	0.18	0.01	0.12	0.00	0.00
Maximum	0.82	2.00	1.00	0.13	0.13

Appendix M
Complete Word List

Word List

(No Stopwords)

able	anybody	bag	better	bop
abracadabra	anymore	baggy	big	bopping
absolutely	anyone	bags	bigger	boring
accomplishment	anything	bake	biggest	borrow
across	anyway	ball	bike	bottom
act	anywhere	banana	bill	bottomless
action	apart	bananas	billion	bounces
add	apes	bang	bin	bouncing
added	apologize	barber	binoculars	bowl
admire	apparently	barked	bird	box
admitted	appeared	barn	birds	boxes
adventure	appetite	baseball	biscuits	boy
adventures	apple	basement	bit	boys
afraid	apples	bass	bite	brake
afternoon	applied	bat	bites	branches
agree	apply	bath	bits	break
agreed	applying	bathroom	bitty	breakfast
ah	approves	baths	black	breaking
aha	arm	bats	blanket	breaks
ahem	arms	batter	blaze	breath
air	around	batteries	blech	breeze
airborne	arranged	beak	blended	bright
aisle	arrived	beans	blew	brightness
alarm	artist	bear	blizzard	bring
alarmed	aside	bears	blorp	bringing
albert	ask	beat	blows	broccoli
alert	asked	beautiful	blue	broke
allowed	asking	bebops	bluejay	broken
almost	asks	becomes	blues	brooms
alone	asleep	bed	blushed	brother
along	ate	bedtime	blustery	brothers
alongside	attacks	bee	boat	brought
aloud	auntie	beep	body	brown
already	automobiles	beeped	boinked	brush
alright	available	beg	bolts	brushed
also	aw	began	bonanza	brushes
although	award	begin	bone	bubbles
always	away	begins	bones	bucket
amaze	awful	behind	bonk	bud
amazed	babies	believe	bonnet	bug
amazing	baby	bell	bony	bugs
amount	back	belly	boo	build
Andes Mountains	backed	belong	boogie	building
anew	backing	bench	boogied	buildings
angels	backpack	bends	boogies	built
announced	backs	benny	book	bullhorn
another	backwards	bent	bookmarks	bumblebee
answer	backyard	beside	books	bundle
answered	bad	best	boom	buried
ants	baffled	bet	boots	burn

burning	cat	circus	convertible	crunchies
bush	catbird	city	cook	crunchy
bushes	catch	clank	cookbook	crust
business	catches	clanking	cooked	cry
bust	catching	clap	cookie	crybabies
busy	caterpillar	clapped	cookies	crying
butter	caught	claps	cool	cucumber
buttered	cave	class	cooling	cucumbers
butterfingers	caves	classroom	correct	cup
butterflies	certainly	claws	could	cupboard
butterfly	chair	clean	couldn't	cups
buttering	chairs	cleaned	count	curious
butterscotch	chance	clear	counted	curled
button	chapter	clearing	counting	curried
buttonless	chasing	climb	country	curtain
buttons	check	climbed	course	cushions
buy	checked	climber	cousins	cut
buzz	checker	climbs	cover	cute
bye	checkers	clip	covered	cuts
cafeteria	checking	clock	cow	dad
cake	cheeks	close	cowboy	daddy
call	cheep	closed	cowboys	daisy
called	cheer	closer	cowgirl	dance
calls	cheered	cloth	cowhorse	danced
came	cheese	clothes	cows	dancer
candies	chefs	cluck	crack	dancin
candles	cherry	coach	cracked	dancing
candy	chest	coat	crash	danger
canes	chew	cocoa	crawled	dark
cannot	chewed	cold	crayon	darker
capped	chewing	collected	crayons	darling
capsule	chick	collecting	crazy	darlings
car	chickadee	collection	cream	data
caramel	chicken	collects	creative	date
card	chickens	color	creek	dawn
cardinal	chicks	colored	crept	day
cards	children	colorful	crick	days
care	chin	colors	cried	dazed
careened	chirped	come	crisp	dead
careful	chocolate	comes	critch	deal
carefully	chomp	coming	croak	dealt
cares	choo	companion	crocodile	dear
caromed	chopstick	companions	cross	dearest
carpet	chopsticks	compromise	crosses	dears
carried	chosen	concentrate	crossover	decide
carrots	chuckled	concoction	crossword	declared
carrying	cinnamon	cone	crowded	decorated
cars	circle	contentedly	crowed	deep
case	circled	contest	cruiser	deeply
castle	circles	continued	crunch	deer

definitely	dresser	emergency	fantastic	fix
delicious	drew	empty	far	fixes
dens	dried	end	fast	flap
department	drifted	endangered	faster	flapped
desk	drink	ending	fastest	flashing
desks	drinking	ends	fat	flat
desperate	dripped	enjoyed	father	flecks
dessert	drippy	enough	fault	flesh
detection	drive	equal	favorite	flew
didn't	driver	escape	favorites	flies
different	driveway	especially	fear	flip
dig	driving	eureka	feathered	flippety
digs	droopy	even	feathers	floor
dime	drop	evening	fed	flopped
dine	dropped	ever	feed	flopping
ding	drops	every	feel	flour
dinner	drove	everybody	feelers	flower
dinos	drowsy	everyday	feeling	flowers
dinosaur	drummers	everyone	feels	flowing
dirt	dry	everything	feet	fluff
dirty	dug	everywhere	fell	fluffed
disappeared	dumb	ew	felt	flump
discovered	dump	exactly	fence	fly
disgusting	dumped	excellent	fetchd	flying
displeased	dumplings	except	fever	follow
dived	dusted	excited	fewer	followed
dizzy	ear	exciting	fifty	follows
doesn't	early	exclaimed	figured	folly
dog	ears	excuse	filled	food
doghouse	earthworms	expertly	film	foods
dogs	east	explained	finally	fooled
doll	easy	exploring	find	fools
done	eat	expose	finding	foot
dong	eater	extensions	finds	football
don't	eaters	extra	fine	footprints
door	eating	extraordinary	finest	forehead
double	eats	extremely	finger	forest
doubt	edge	eye	fingers	forever
downstairs	eek	eyelash	finish	forgot
dozy	effects	eyelids	finished	forgotten
drank	egg	eyes	fire	fork
draw	eggs	faces	firetruck	fort
drawer	eight	fact	firm	forth
drawing	eighth	fairy	first	forts
drawings	eights	fall	fish	forward
drawn	either	falling	fisheye	found
dreamed	elbows	falls	fishman	fountain
dreaming	electric	families	fist	four
dress	elephants	fancy	fit	fourteen
dressed	else	fangs	five	fox

fragrant	giving	grows	heavens	hose
free	glad	grumbled	heavy	hot
freezer	glared	grumped	hee	hotter
french	glass	guess	heh	hottest
fried	glide	gulp	held	hour
friend	glides	gum	hello	house
friends	glinted	gust	help	hover
friendship	glue	gusto	helped	however
frightening	go	guts	helping	howled
frog	goal	guy	helpings	huckleberry
front	goes	guys	herd	huff
frowned	going	gym	herds	hug
frozen	goldfinch	ha	hero	huge
fruit	gone	haiku	hey	hugged
full	good	hair	hi	hugging
fully	goodbye	haircut	hid	huh
fun	goodness	haircuts	hide	humans
funny	gosh	hairy	hiding	humorous
fur	got	half	high	hundred
furthermore	grabbed	hallway	higher	hungry
future	grabs	halt	highway	hunt
galloped	grandmother	hammer	hills	hurry
game	grandpa	hand	hip	hurt
games	grandpas	handed	hippo	hurts
garage	granny	handful	hippos	hush
garbage	grape	hands	hired	ice
garbageman	grapefruit	hanging	hit	idea
garden	grass	happen	hits	ideas
gardens	grasshopper	happened	hobbies	ignition
gas	grasshoppers	happening	hocus	ignores
gasped	grateful	happily	hog	illegal
gather	gray	happy	hold	imagine
gathered	great	hard	holding	immediately
gave	greatest	hardly	hole	implore
gazed	green	hardware	holes	imported
gears	greetings	hat	hollered	incapable
gee	grew	hatch	holy	inch
gently	grief	hatched	home	inches
get	grind	hatching	homework	indeed
gets	grip	hats	hooray	indulging
getting	groan	haven	hooted	inform
ghost	groaned	head	hooves	inquire
ghosts	groceries	heads	hop	insects
gift	groom	hear	hope	inside
gifts	groovy	heard	hoped	inspected
giggled	gross	hearing	hopefully	instead
girl	grossest	heart	hopped	intended
girls	ground	hearts	hops	interest
give	grow	heats	horse	interesting
gives	growing	heavenly	horses	interrupted

invited	knock	life	lower	messy
inviting	knocked	lift	lucky	met
isn't	knocking	lifted	lullaby	metal
itty	know	lifts	lulu	meter
ivy	knows	light	lumps	mice
jabbed	knuckles	lightning	lumpy	middle
james	label	lights	lunch	might
jar	lake	like	lunchroom	mighty
jay	land	liked	lunchtime	milk
jays	landing	likes	mad	million
jazzy	lands	limbo	made	mimicking
jealous	lap	limbos	magazines	mind
jewel	laps	line	magic	mine
jiggly	large	lines	magnetic	mint
jitterbug	lasagna	lip	magnificent	mints
job	last	liquid	maid	minute
jobs	late	list	make	miracle
join	later	listen	makes	mirror
joined	laugh	listened	making	miss
jokes	laughed	littered	mama	missed
journal	laughing	little	man	missing
journey	laughs	live	mane	mistake
judges	law	lived	many	mix
juice	lawn	lives	map	mixed
jump	laws	living	marathon	mixer
jumped	lay	located	marble	mixes
jumper	leaf	location	marbles	moaned
jumping	leaned	logical	marched	mole
jumps	leapt	logs	marry	molebird
keep	learned	lollipop	marvelous	moles
keeps	least	long	math	mom
kept	leave	longer	matter	mommy
key	leaves	longing	maybe	money
kick	led	longs	meadows	monkeys
kickball	ledge	look	meal	monster
kicked	left	looked	mean	monsters
kid	leftover	looking	means	months
kidding	legs	looks	meantime	moon
kids	lem	losing	meanwhile	morning
kind	lemonade	lost	meat	mostly
kinds	less	lot	meatball	mother
king	let	lots	melon	moths
kiss	letters	loud	melt	mound
kisses	lettuce	louder	melting	mountains
kitchen	librarian	lovable	menace	mouse
kite	librarians	love	menu	mousebird
kitty	library	loved	mercy	mouth
knees	licked	lovely	mess	move
knew	lid	loves	messed	moved
knob	lie	low	messing	moves

movie	nicer	ouch	pay	plane
moving	nick	outdone	pearl	planned
Mr.	night	outgrow	peas	planning
Mrs.	nine	outrageous	pedal	plans
Ms.	nobody	outside	peek	plant
much	nodded	oven	peeked	planted
muckiest	noise	overdue	peered	plants
mucky	none	overload	pen	plate
mucus	nonsense	overnight	pencils	play
mud	north	owe	penny	played
muddy	nose	owl	people	player
mumbled	nosed	pack	per	players
munched	note	packed	perfect	playground
mushroom	notebook	pad	perfecting	playing
mushy	nothing	padded	perhaps	please
music	notice	paddling	pests	plop
musical	noticed	page	pet	plucked
must	number	pages	pets	plumped
mustache	nurse	pail	petted	plus
muttered	nuts	pains	phooey	pocket
mysteries	nutty	paint	piano	pockets
name	oak	pair	pick	pocus
named	oatmeal	pajamas	picked	poem
names	oats	pal	pickled	poems
nap	observe	pancakes	picks	poet
naps	ocean	panic	picky	point
narrow	officer	panicked	picnic	pointed
narrowed	oh	pans	picture	pointy
narrows	oink	pants	pictures	poked
navy	oinked	papa	pie	poker
near	ok	paper	piece	polar
nearby	okay	pardon	pieces	pole
nearly	old	parents	pies	police
neat	one	part	pig	polished
neatly	onion	particular	piggy	pond
neck	onions	particularly	pigs	ponds
need	onto	partner	pile	poof
needed	oof	partners	pine	pooie
needle	oops	parts	pinecone	pool
needs	open	party	pin	poor
neighbor	opened	pass	pink	pop
neighbors	opening	passenger	pirate	popcorn
nest	opens	past	pit	popped
nests	opinion	pasted	pitcher	porch
never	orange	pasture	pizza	porcine
new	oranges	patch	pizzas	possum
news	order	patches	pjs	possums
next	ordered	patted	place	pot
nibble	organized	paused	places	potion
nice	others	paw	plan	pots

pounded	putting	reminded	rows	scribbled
poured	puzzles	remote	rubbed	scrunch
pouring	quality	remove	rubber	scurried
power	queedle	removed	ruined	scurry
pox	questions	replied	rules	sea
practice	quick	reported	rum	seals
practiced	quickly	reports	rumbles	search
preen	quiet	repulsive	rummaged	searching
prefer	quilt	rescue	run	seat
prepared	quite	responsible	runaways	seatbelt
present	rabbit	rest	runner	secret
presents	rabbits	rested	running	secretly
pressed	raccoon	result	runs	see
pretended	raced	resumes	rushed	seed
pretty	racing	retreats	rustled	seeds
prey	radula	returned	sad	seek
princess	raft	reveal	saddle	seem
printed	railroad	reviewed	saddlebag	seemed
prize	rain	rhino	safe	seen
probably	raining	ribbons	said	seeps
problem	rains	rich	salad	sees
problems	rainy	rid	salsa	selected
prodded	raised	ride	salty	sell
prodigy	raisin	rides	sandwich	selling
proof	ran	right	sandwiches	send
pros	ranch	righty	sang	served
proud	ranches	rises	sat	set
puddle	rang	rising	Saturday	settle
puff	rare	road	saved	settled
puffs	rat	robin	saw	settles
pull	rats	robins	say	seven
pulled	razz	robot	says	sewed
pulling	read	robots	scamper	shade
pulls	reader	rock	scan	shadow
pumpkin	reading	rocking	scare	shady
pumpkins	reads	rocks	scary	shakes
puppy	ready	rode	scat	shaking
purchase	real	rolled	scatter	shall
purdy	really	roller	school	shapes
pure	reason	rolls	scientists	share
purple	recess	romp	scissors	sharp
purred	recipe	romping	scoot	shazaam
push	recital	room	scrambled	shed
pushed	red	rope	scraps	sheets
pushes	reddish	rose	scratch	shelf
pushing	reek	rotten	scratched	shell
pussy	reflection	round	screeched	shh
put	remarkable	rounded	screwdriver	shining
puts	remember	rower	screwed	shiny
putter	remembered	row	screws	shirt

shivered	skating	snail	spaghetti	starts
shoes	sketch	snails	speak	stay
shoo	sketchpad	snake	special	stayed
shook	sketchpads	snap	specially	stays
short	skilled	snapping	specks	steal
shorts	skin	snappy	spectacular	steamed
shoulder	skip	snazzy	sped	steering
shouted	skit	sneaks	speed	stem
shouts	skitter	sneaky	speeding	step
show	skivvies	sneeze	spelled	stepped
showed	skunk	sneezed	spend	steps
showers	sky	sniff	spent	stew
shows	slab	sniffed	spicy	stick
shrug	sled	snip	spied	stickers
shrugged	sledding	snoozy	spies	sticks
shuddered	sleep	snore	spilled	sticky
shushed	sleeping	snored	splash	still
shut	sleepover	snorted	splat	stinky
shyly	sleeps	snout	splendid	stitch
sick	sleeve	snow	splish	stole
side	slept	snowball	split	stomach
sides	slid	snowed	splits	stomped
sidewalk	slide	snowing	splop	stomping
sigh	slime	snows	spoken	stood
sighed	slimy	snuffled	spoon	stoop
sight	slippery	snuffles	sport	stop
sign	slot	snug	spot	stopped
signed	slowly	snuggled	spots	stops
signing	slug	soaked	spotted	store
signs	slugs	soar	spread	stories
silent	slumped	sock	spreads	storm
silly	slurp	socks	spring	story
simply	sly	sofa	sprinkled	straw
since	small	soft	spy	strength
sing	smaller	soil	spying	stretch
singer	smallest	solution	squash	stretched
singing	smart	solve	squeaked	string
sings	smartest	somebody	squeaky	strip
sink	smell	someone	squeeze	strokes
sir	smelled	something	squirrel	stroller
siskin	smelliest	sometimes	squirting	strong
sister	smells	somewhere	staggered	strum
sisters	smelly	song	stand	stuck
sit	smiled	songs	standing	studied
sits	smiles	soon	star	stuff
sitting	smoke	sooner	stared	stuffed
six	smooth	sopping	stars	stumps
size	smoothly	sorry	start	sudden
skate	snack	sound	started	suddenly
skates	snag	sounds	starting	sugar

suggested	tale	throughout	trail	unremarkable
suitcase	talented	throw	trained	unusual
summer	tales	throwing	trash	unwillingness
sun	talk	throws	tree	upon
Sunday	talking	thunder	trees	upset
sunglasses	tall	tickets	triceratopses	upside
sunny	tallest	tickling	trick	us
sunrise	tap	tie	tricks	use
sunset	tapped	tied	tricky	used
supertruck	taps	tight	tried	using
suppose	taste	tilt	tripped	usual
supposed	tasted	tilted	troll	vehicle
sure	tastiest	time	trolls	ventured
surfaces	tasty	times	trouble	vinegar
surprise	tat	ting	truck	vines
surprised	taxis	tings	trucks	virtually
suspected	tea	tiniest	truly	visit
swallowed	teaberry	tiny	trunk	visitor
swam	teacher	tips	trust	visual
swamp	team	tired	try	voice
swamps	teapot	title	trying	volume
swap	tee	toad	tucked	voyage
swapped	teeny	toads	tum	vulture
swatter	teeth	toast	tumbled	vultures
sweat	tell	today	tummy	waah
sweated	telling	together	turkey	wag
sweatshirt	tells	told	turn	wagged
sweaty	tempo	tomatoes	turned	waited
sweet	ten	tomorrow	turns	wait
sweetie	tens	tongue	tv	waited
swerving	tentacles	tonight	twenty	waiting
swim	terrible	took	twig	waits
swimming	test	toolbox	twigs	wake
swing	tested	tools	twin	wakes
swinging	thank	toot	twins	walk
swings	thanks	toothy	twirled	walked
swung	therefore	toots	twitch	walking
symbol	thicket	top	twitched	wall
symbols	thing	toppled	two	wand
system	things	tore	tyrannosaurus	want
tabby	think	torn	ugh	wanted
table	thinking	toss	uh	wants
tacked	thinks	tossed	um	warm
tadpole	thirsty	tosses	umbrella	warming
tag	though	tough	umbrellas	wash
tail	thought	toward	unbuckled	wasn't
tails	thousand	town	uncle	wasted
take	thread	towns	underneath	watch
taken	three	toy	underpants	watched
takes	threw	tracks	underwear	watches

watching	wherever	wish	worried	year
water	whew	without	worrier	years
watered	whiff	wobbly	worry	yech
watermelon	whisper	woke	worrying	yelled
waved	whispered	wolfsnail	worth	yelling
waves	whistle	won	would	yellow
way	whiz	wonder	wouldn't	yelped
ways	whoa	wondered	wow	yeow
wear	whole	wonderful	wrapped	yes
wearing	whoop	wonders	wrench	yesterday
weeds	whoops	woods	write	yet
weird	whoosh	woozy	writer	yikes
welcome	whose	wop	writing	yippee
well	wide	word	written	younger
went	wider	words	wrong	yow
west	wig	work	wrote	yuck
western	willow	worked	yanked	yum
wet	win	working	yard	yummy
whale	wind	works	yawn	
whapp	window	world	yawned	
whatever	wings	worm	yay	
wheel	winning	worms	yeah	

Appendix N

Punctuaion

Punctuation Ratios

BOOK NAME	# SENTENCES	.	?	!
A Big Guy Took My Ball	46	15	8	18
A Pig, a Fox and a Box	75	52	4	16
A Splendid Friend, Indeed	26	16	9	0
Amanda Pig and the Really Hot Day	242	231	23	21
Are You Ready to Play Outside	47	13	3	41
BALL	44	1	20	10
Benny and Penny and the Big No	151	43	17	121
Bink and Gollie	132	116	25	17
Chicken Said Cluck!	46	41	0	50
Cowgirl Kate and Cocoa	159	144	12	12
Don't Throw It to Mo	70	69	0	0
First the Egg	32	0	0	1
Hello, Bumblebee Bat	49	30	9	1
Henry and Mudge and the Great Grandpas	87	75	1	11
Hi! Fly Guy	74	47	0	16
I Broke My Trunk	53	21	14	18
I Spy Fly Guy!	56	37	6	12
I Want My Hat Back	48	43	13	0
Jazz Baby	60	31	0	12
Let's Go for a Drive!	62	13	5	62
Ling and Ting	144	114	17	29
Little Mouse Gets Ready	49	9	7	30
Mercy Watson Goes for a Ride	350	396	31	34
Mouse and Mole	311	256	13	78
Move Over, Rover!	63	24	8	34
Mr. Putter and Tabby Turn the Page	116	135	1	10
Not a Box	12	3	6	3
One Boy	22	0	0	1
Pearl and Wagner One Funny Day	163	174	22	35
Penny and Her Marble	131	139	12	3
Pete the Cat and His Four Groovy Buttons	73	19	10	19
Rabbit and Robot	268	238	22	32
See Me Run	22	17	3	2
Stinky	276	39	42	196
Supertruck	16	11	2	2
Tales for Very Picky Eaters	136	116	14	4
The Watermelon Seed	31	11	0	26
There Is a Bird on Your Head!	50	13	16	25
Up! Tall! and High! (but not necessarily in that order.)	26	4	1	16
Vulture View	55	30	9	7
Waiting	39	31	4	1
Waiting Is Not Easy!	51	20	9	24
We Are in a Book!	86	13	12	61
Wolfsnail	45	40	2	2
You Are (Not) Small	22	18	4	11
Zelda and Ivy the Runaways	131	129	8	7
Averages	92.33	66.02	9.65	24.59
Median	58.00	31.00	8.00	16.00
Minimum	12.00	0.00	0.00	0.00
Maximum	350.00	396.00	42.00	196.00

Appendix O
Punctuation Counts

Punctuation Counts

BOOK NAME	,	.	?	!	'	"	:	;	...
A Big Guy Took My Ball	5	15	8	18	0	0	0	0	1
A Pig, a Fox and a Box	16	52	4	16	1	0	0	0	0
A Splendid Friend, Indeed	3	16	9	0	0	0	0	0	0
Amanda Pig and the Really Hot Day	119	231	23	21	0	0	0	0	0
Are You Ready to Play Outside	5	13	3	41	0	0	0	0	0
BALL	0	1	20	10	0	0	0	0	0
Benny and Penny and the Big No	22	43	17	121	0	0	0	0	0
Bink and Gollie	99	116	25	17	0	0	0	0	1
Chicken Said Cluck!	5	41	0	50	0	0	1	0	0
Cowgirl Kate and Cocoa	67	144	12	12	0	0	0	0	0
Don't Throw It to Mo	21	69	0	0	22	48	0	0	0
First the Egg	0	0	0	1	0	0	0	0	2
Hello, Bumblebee Bat	11	30	9	1	0	0	0	0	0
Henry and Mudge and the Great Grandpas	38	75	1	11	0	0	0	0	0
Hi! Fly Guy	16	47	0	16	0	0	1	0	0
I Broke My Trunk	14	21	14	18	0	0	0	0	8
I Spy Fly Guy!	12	37	6	12	0	0	0	0	0
I Want My Hat Back	2	43	13	0	0	0	0	0	0
Jazz Baby	14	31	0	12	0	0	0	0	0
Let's Go for a Drive!	4	13	5	62	0	0	0	0	3
Ling and Ting	50	114	17	29	0	0	0	0	5
Little Mouse Gets Ready	14	9	7	30	0	0	0	0	11
Mercy Watson Goes for a Ride	160	396	31	34	22	332	1	0	2
Mouse and Mole	60	256	13	78	7	186	13	1	11
Move Over, Rover!	42	24	8	34	0	0	0	0	1
Mr. Putter and Tabby Turn the Page	31	135	1	10	0	0	0	0	0
Not a Box	2	3	6	3	0	0	0	0	0
One Boy	0	0	0	1	0	0	0	0	1
Pearl and Wagner One Funny Day	48	174	22	35	0	0	1	0	4
Penny and Her Marble	45	139	12	3	0	0	0	0	0
Pete the Cat and His Four Groovy Buttons	30	19	10	19	0	0	5	0	0
Rabbit and Robot	94	238	22	32	13	277	7	0	0
See Me Run	2	17	3	2	0	0	0	0	0
Stinky	40	39	42	196	0	0	2	0	28
Supertruck	3	11	2	2	1	0	0	0	1
Tales for Very Picky Eaters	68	116	14	4	0	0	0	0	3
The Watermelon Seed	4	11	0	26	0	0	0	0	2
There Is a Bird on Your Head!	4	13	16	25	0	0	0	0	2
Up! Tall! and High! (but not necessarily in that order.)	3	4	1	16	0	0	0	0	2
Vulture View	26	30	9	7	0	0	0	0	8
Waiting	12	31	4	1	3	0	0	0	0
Waiting Is Not Easy!	3	20	9	24	0	0	0	0	4
We Are in a Book!	1	13	12	61	0	0	0	0	3
Wolfsnail	16	40	2	2	0	0	0	0	0
You Are (Not) Small	1	18	4	11	0	0	0	0	0
Zelda and Ivy the Runaways	58	129	8	7	0	0	3	0	0
AVERAGES	28	66	10	25	2	18	1	0	2
Median	14	31	8	16	0	0	0	0	0
Minimum	0	0	0	0	0	0	0	0	0
Maximum	160	396	42	196	22	332	13	1	28

Appendix P

Syntactic Analysis

Syntactic Analysis

BOOKNAME	S	VP	DC	CP	CN
A Big Guy Took My Ball	39	37	7	0	20
A Pig, a Fox and a Box	72	80	4	1	21
A Splendid Friend, Indeed	25	29	3	0	7
Amanda Pig and the Really Hot Day	196	318	49	54	168
Are You Ready to Play Outside	53	47	0	0	6
BALL	30	5	0	0	5
Benny and Penny and the Big No	173	138	18	4	67
Bink and Gollie	96	194	43	11	138
Chicken Said "Cluck!"	73	74	2	0	5
Cowgirl Kate and Cocoa	126	191	32	24	100
Don't Throw It to Mo	65	98	8	2	20
First the Egg	1	2	1	8	1
Hello, Bumblebee Bat	40	47	6	4	23
Henry and Mudge and the Great Grandpas	80	125	20	30	54
Hi! Fly Guy	51	78	14	4	30
I Broke My Trunk	48	48	5	3	15
I Spy Fly Guy!	41	54	5	7	25
I Want My Hat Back	56	52	3	2	11
Jazz Baby	37	53	5	5	18
Let's Go for a Drive!	75	65	12	1	18
Ling and Ting	104	189	39	19	118
Little Mouse Gets Ready	46	52	5	5	25
Mercy Watson Goes for a Ride	377	448	33	23	123
Mouse and Mole	346	405	18	30	120
Move Over, Rover!	66	74	6	2	29
Mr. Putter and Tabby Turn the Page	97	139	21	23	41
Not a Box	12	13	0	0	8
One Boy	1	2	0	1	2
Pearl and Wagner One Funny Day	142	174	14	14	62
Penny and Her Marble	114	175	33	16	71
Pete the Cat and His Four Groovy Buttons	43	51	4	4	25
Rabbit and Robot	284	381	20	30	97
See Me Run	22	21	0	6	0
Stinky	245	191	18	6	100
Supertruck	15	16	2	3	8
Tales for Very Picky Eaters	82	195	53	16	145
The Watermelon Seed	35	26	3	0	11
There Is a Bird on Your Head!	52	43	1	2	14
Up! Tall! and High! (but not necessarily in that order.)	21	21	5	0	4
Vulture View	46	40	1	2	7
Waiting	36	43	6	7	13
Waiting Is Not Easy!	52	44	2	3	6
We Are in a Book!	75	72	5	0	14
Wolfsnail	44	55	6	10	32
You Are (Not) Small	29	27	0	0	0
Zelda and Ivy the Runaways	107	188	33	25	109
Average	84.130	104.783	12.283	8.848	42.087
Median	52.500	54.500	5.500	4.000	20.500
Minimum	1	2	0	0	0
Maximum	377	448	53	54	168

Appendix Q

T-unit Analysis

T-Unit Analysis

BOOKNAME	T	VP/T	DC/T	CP/T	CN/T
A Big Guy Took My Ball	32	1.156	0.219	0.000	0.625
A Pig, a Fox and a Box	74	1.081	0.054	0.014	0.284
A Splendid Friend, Indeed	23	1.261	0.130	0.000	0.304
Amanda Pig and the Really Hot Day	195	1.631	0.251	0.277	0.862
Are You Ready to Play Outside	41	1.146	0.000	0.000	0.146
BALL	23	0.217	0.000	0.000	0.217
Benny and Penny and the Big No	141	0.979	0.128	0.028	0.475
Bink and Gollie	95	2.042	0.453	0.116	1.453
Chicken said Cluck!	72	1.028	0.028	0.000	0.069
Cowgirl Kate and Cocoa	128	1.492	0.250	0.188	0.781
Dont throw it to mo	67	1.463	0.119	0.030	0.299
First the Egg	0	0.000	0.000	0.000	0.000
Hello, Bumblebee Bat	39	1.205	0.154	0.103	0.590
Henry and Mudge and the Great Grandpas	82	1.524	0.244	0.366	0.659
Hi! Fly Guy	50	1.560	0.280	0.080	0.600
I Broke My Trunk	42	1.143	0.119	0.071	0.357
I Spy Fly Guy!	39	1.385	0.128	0.180	0.641
I Want My Hat Back	48	1.083	0.063	0.042	0.229
Jazz Baby	35	1.514	0.143	0.143	0.514
Let's Go for a Drive!	53	1.226	0.226	0.019	0.340
Ling and Ting	107	1.766	0.365	0.178	1.103
Little Mouse Gets Ready	39	1.333	0.128	0.128	0.641
Mercy Watson Goes for a Ride	366	1.224	0.090	0.063	0.336
Mouse and Mole	338	1.198	0.053	0.089	0.355
Move Over, Rover!	56	1.321	0.107	0.036	0.518
Mr. Putter and Tabby Turn the Page	99	1.404	0.212	0.232	0.414
Not a Box	12	1.083	0.000	0.000	0.667
One Boy	1	2.000	0.000	1.000	2.000
Pearl and Wagner One Funny Day	135	1.289	0.104	0.104	0.459
Penny and Her Marble	115	1.522	0.287	0.139	0.617
Pete the Cat and His Four Groovy Buttons	43	1.186	0.093	0.093	0.581
Rabbit and Robot	281	1.356	0.071	0.107	0.345
See Me Run	20	1.050	0.000	0.300	0.000
Stinky	196	0.975	0.092	0.031	0.510
Supertruck	14	1.143	0.143	0.214	0.571
Tales for Very Picky Eaters	91	2.143	0.582	0.176	1.593
The Watermelon Seed	30	0.867	0.100	0.000	0.367
There Is a Bird on Your Head!	46	0.935	0.022	0.044	0.304
Up! Tall! and High! (but not necessarily in that order.)	18	1.167	0.278	0.000	0.222
Vulture View	37	1.081	0.027	0.054	0.189
Waiting	37	1.162	0.162	0.189	0.351
Waiting Is Not Easy!	45	0.978	0.044	0.067	0.133
We Are in a Book!	66	1.091	0.076	0.000	0.212
Wolfsnail	43	1.279	0.140	0.233	0.744
You Are (Not) Small	28	0.964	0.000	0.000	0.000
Zelda and Ivy the Runaways	106	1.774	0.311	0.236	1.028
Average	79.304	1.248	0.141	0.117	0.515
Median	47.000	1.202	0.119	0.076	0.437
Minimum	0	0	0	0	0
Maximum	366	2.1429	0.5824	1	2

Appendix R

Prosodic Measures

Prosodic Measures

Title	Rhyme Scheme	Alliteration	Assonance	Internal Rhyme	
A Big Guy Took My Ball!		x			
A Pig, a Fox, and a Box					
A Splendid Friend, Indeed	couplets		x		
Amanda Pig and the Really Hot Day		x		x	
Are You Ready to Play Outside?					
Ball					
Benny and Penny in the Big No-No!		x			
Bink and Gollie					
Chicken Said, "Cluck!"					
Cowgirl Kate and Cocoa					
Don't Throw It to Mo					
First the Egg		x			
Hello, Bumblebee Bat					
Henry and Mudge and the Great Grandpas		x	x		
Hi! Fly Guy		x	x	x	
I Broke My Trunk					
I Spy Fly Guy!				x	
I Want My Hat Back				x	
Jazz Baby	ABCB/couplets	x	x		
Let's Go for a Drive!		x		x	
Ling & Ting: Not Exactly the Same!					
Little Mouse Gets Ready				x	
Mercy Watson Goes for a Ride		x	x	x	
Mouse and Mole: Fine Feathered Friends	AABB				
Move Over, Rover!	ABCA/couplets	x			
Mr. Putter and Tabby Turn the Page					
Not a Box					
One Boy	couplets		x		
Pearl and Wagner: One Funny Day					
Penny and Her Marble		x			
Pete the Cat and His Four Groovy Buttons	couplets				
Rabbit & Robot: The Sleepover		x			
See Me Run	couplets			x	
Stinky		x			
Supertruck					
Tales for Very Picky Eaters		x			
The Watermelon Seed		x			
There is a Bird on Your Head					
Up! Tall! and High! (but not necessarily in that order.)					
Vulture View	AB/couplets			x	
Waiting					
Waiting Is Not Easy!					
We Are in a Book!					
Wolfsnail: A Backyard Predator		x			
You Are (Not) Small					
Zelda and Ivy: The Runaways					
	Counts	8	16	6	9
	Percentage	17%	35%	13%	20%

Appendix S

Typography

Typography

Title	Size	Caps	Bold	Italic	Color	Font/HL
A Big Guy Took My Ball!	X	X		X		
A Pig, a Fox, and a Box		X				
A Splendid Friend, Indeed						
Amanda Pig and the Really Hot Day						
Are You Ready to Play Outside?	X	X				
BALL	X	X		X		
Benny and Penny in the Big No-No!		X	X	X	X	X
Bink and Gollie		X				X
Chicken Said, "Cluck!"						
Cowgirl Kate and Cocoa				X *		
Don't Throw It to Mo						
First the Egg		X				
Hello, Bumblebee Bat	X		X			
Henry and Mudge and the Great Grandpas				X		
Hi! Fly Guy						X
I Broke My Trunk	X			X		
I Spy Fly Guy!						X
I Want My Hat Back		X			X	
Jazz Baby	X	X				
Let's Go for a Drive!	X	X		X		
Ling& Ting: Not Exactly the Same!	X	X		X		
Little Mouse Gets Ready		X	X	X		
Mercy Watson Goes for a Ride						X
Mouse and Mole: Fine Feathered Friends						
Move Over, Rover!		X			X	
Mr. Putter and Tabby Turn the Page			X			
Not a Box						
One Boy						
Pearl and Wagner: One Funny Day						
Penny and Her Marble						
Pete the Cat and His Four Groovy Buttons	X	X	X		X	
Rabbit & Robot: The Sleepover						X
See Me Run						
Stinky	X		X	X		X
Supertruck		X				
Tales for Very Picky Eaters						
The Watermelon Seed	X	X				X
There is a Bird on Your Head	X	X				
Up! Tall! and High! (but not necessarily in that order.)						X
Vulture View	X			X		
Waiting						
Waiting Is Not Easy!	X	X		X		X
We Are in a Book!	X	X		X		X
Wolfsnail: A Backyard Predator						
You Are (Not) Small		X		X		X
Zelda and Ivy: The Runaways				X		
Feature Count	15	20	6	15	4	12
Without Feature	31	26	40	31	42	34

Appendix T

Literary Devices

Literary Devices

Title	Simile	Metaphor	Personification	Onomatopoeia	Hyperbole	Idiom
A Big Guy Took My Ball!			x			
A Pig, a Fox, and a Box						
A Splendid Friend, Indeed			x			
Amanda Pig and the Really Hot Day	x		x		x	
Are You Ready to Play Outside?			x	x	x	x
BALL						
Benny and Penny in the Big No-No!			x	x		x
Bink and Gollie				x	x	x
Chicken Said, "Cluck!"			x	x		
Cowgirl Kate and Cocoa		x	x	x		x
Don't Throw It to Mo						
First the Egg						
Hello, Bumblebee Bat	x	x				
Henry and Mudge and the Great Grandpas				x		x
Hi! Fly Guy				x	x	
I Broke My Trunk			x	x	x	
I Spy Fly Guy!						
I Want My Hat Back			x			
Jazz Baby				x		
Let's Go for a Drive!			x	x		
Ling & Ting: Not Exactly the Same!	x			x		
Little Mouse Gets Ready			x	x		
Mercy Watson Goes for a Ride		x	x	x		x
Mouse and Mole: Fine Feathered Friends			x	x		x
Move Over, Rover!				x		
Mr. Putter and Tabby Turn the Page				x		
Not a Box			x			
One Boy						
Pearl and Wagner: One Funny Day			x	x		x
Penny and Her Marble	x		x			x
Pete the Cat and His Four Groovy Buttons			x	x		
Rabbit & Robot: The Sleepover			x	x		x
See Me Run			x	x		
Stinky				x	x	x
Supertruck						
Tales for Very Picky Eaters			x		x	
The Watermelon Seed			x	x	x	
There is a Bird on Your Head			x	x		
Up! Tall! and High! (but not necessarily in that order.)			x	x		
Vulture View				x		
Waiting						
Waiting Is Not Easy!			x	x		
We Are in a Book!			x	x		x
Wolfsnail: A Backyard Predator				x		
You Are (Not) Small			x	x		
Zelda and Ivy: The Runaways	x		x	x	x	x
Device Count	5	3	27	30	9	14
% of 46 (the total)	10.87%	6.52%	58.70%	65.22%	19.57%	30.43%

Appendix U

Human Experience

Human Experience - Appendix U

Title	Name	Gender	Evidence	Animal	Ethnicity	Ability	Age
Don't Throw It to Mo!	Mo	M	name, dress, pronoun		black		C
A Pig, a Fox, and a Box	Coach Steve Pig Fox	M I I	pronoun, dress	pig fox	black		A
Supertruck	Supertruck	M	pronoun	(garbage truck/snow plow bucket truck) (fire truck) (tow truck)			
Waiting		I		owl			
		I		pig			
		I		bear			
		I		puppy			
		I		rabbit			
		F		cat			
		I		creatures			
You Are (Not) Small	Mr. Putter	M			white		E
Mr. Putter and Tabby Turn the Page	Tabby	F	pronoun	cat			
	Mrs. Teaberry	F			white		E
	Zeke	M		dog			
Waiting Is Not Easy!	Gerald	M	pronoun	elephant			C
	Piggie	F		pig			C
The Watermelon Seed		I		crocodile			
BALL		I		dog			
A Big Guy Took My Ball!	Gerald	M		elephant			C
Penny and Her Marble	Piggie	F		pig			C
	Mama	F	pronoun, dress A/I	mouse			A
	Papa	M		mouse			A
	Mrs. Goodwin	F		mouse			E
Up! Tall! and High! (but not necessarily in that order.)		I		birds			E
Let's Go for a Drive!	Gerald	M		elephant			C
	Piggie	F		pig			C
Pete the Cat and His Four Groovy Buttons	Pete	M	pronoun	cat			

Title	Name	Gender	Evidence	Animal	Ethnicity	Ability	Age
Rabbit & Robot: The Sleepover	Rabbit Robot	M I	pronoun	rabbit robot	white		C A
Tales for Very Picky Eaters	James James's Father	M M	name, pronoun father				C C
I Broke My Trunk	Gerald Piggie	M F		elephant pig			C C
I Want My Hat Back		I I		bear rabbit			C C
See Me Run		I		dog			C
Blink and Gollie	Blink Gollie	F F	pronoun dress		white white		C C
Ling & Ting: Not Exactly the Same!	Ling Ting	F F	dress A/I dress A/I		Asian Asian		C C
We Are in a Book!	Gerald Piggie	M I		elephant pig			C C
Benny and Penny in the Big No-No!	Benny Penny	M F	dress A/I behavior dress A/I behavior	mouse mouse			C C
I Spy Fly Guy!	Melinda Buzz Fly Guy	F M M	dress A/I behavior dress A/I behavior	mole fly	white		C C
Little Mouse Gets Ready	Little Mouse Mama	M F	dress A/I	mouse mouse			C A
Mouse and Mole: Fine Feathered Friends	Mole Mouse	M M	pronoun pronoun	mole mouse			C C
Pearl and Wagner: One Funny Day	Pearl Wagner	F M	dress pronoun	rabbit mouse	white		C C
Are You Ready to Play Outside?	Gerald Piggie	M F		elephant pig			C C
Chicken Said, "Cluck!"	Chicken Earl	I M		chicken			C C
One Boy	Pearl Boy	F M			black white white		C C

Appendix V

Typeface and Serif

Typeface and Serif

Title	Typeface	Serif
A Big Guy Took My Ball!	(hand)	Serif
A Pig, a Fox, and a Box		
A Splendid Friend, Indeed	Optima	
Amanda Pig and the Really Hot Day		Serif
Are You Ready to Play Outside?		Serif
BALL		
Benny and Penny in the Big No-No!		
Bink and Gollie	Humana Sans	
Chicken Said, "Cluck!"		Serif
Cowgirl Kate and Cocoa	Filosofia Regular	Serif
Don't Throw It to Mo	Adobe Caslon	Serif
First the Egg		Serif
Hello, Bumblebee Bat	Billy	
Henry and Mudge and the Great Grandpas	Goudy	Serif
Hi! Fly Guy		
I Broke My Trunk	(hand)	Serif
I Spy Fly Guy!		Serif
I Want My Hat Back	New Century School Book	Serif
Jazz Baby	Handegypt and P22Cage	Serif
Let's Go for a Drive!	(hand)	Serif
Ling & Ting: Not Exactly the Same!		Serif
Little Mouse Gets Ready		
Mercy Watson Goes for a Ride	Mrs. Evans	Serif
Mouse and Mole: Fine Feathered Friends	Adobe Caslon	Serif
Move Over, Rover!	Memphis	Serif
Mr. Putter and Tabby Turn the Page	Berkeley Old Style Book	Serif
Not a Box		
One Boy		
Pearl and Wagner: One Funny Day		Serif
Penny and Her Marble		Serif
Pete the Cat and His Four Groovy Buttons		
Rabbit & Robot: The Sleepover	Scala and Digital	Serif
See Me Run	Report School	
Stinky		
Supertruck		
Tales for Very Picky Eaters	Impressum	Serif
The Watermelon Seed		
There is a Bird on Your Head		Serif
Up! Tall! and High! (but not necessarily in that order.)		
Vulture View		
Waiting	Tarzana Wide	
Waiting Is Not Easy!	(hand)	Serif
We Are in a Book!	(hand)	Serif
Wolfsnail: A Backyard Predator	Optima	
You Are (Not) Small		Serif
Zelda and Ivy: The Runaways	Galliard	Serif
		Serif 27
		Sans Serif 19

