

1-1-2014

Genres of Children's Websites: A Comprehensive Methodology for Analyzing Digital Texts

James L. Welsh

University of South Florida, jlwelsh2@gmail.com

Follow this and additional works at: <http://scholarcommons.usf.edu/etd>

 Part of the [Education Commons](#)

Scholar Commons Citation

Welsh, James L., "Genres of Children's Websites: A Comprehensive Methodology for Analyzing Digital Texts" (2014). *Graduate Theses and Dissertations*.

<http://scholarcommons.usf.edu/etd/5605>

This Dissertation is brought to you for free and open access by the Graduate School at Scholar Commons. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Scholar Commons. For more information, please contact scholarcommons@usf.edu.

Genres of Children's Websites
A Comprehensive Methodology for Analyzing Digital Texts

by

James L. Welsh II

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Department of Teaching and Learning
College of Education
University of South Florida

Co-Major Professor: James R. King, Ed.D.
Co-Major Professor: Jenifer J. Schneider, Ph.D.
Danielle V. Dennis, Ph.D.
Liliana Rodriguez-Campos, Ph.D.

Date of Approval:
October 27, 2014

Keywords: multimodality, genre theory, literacy, new literacies, content analysis

Copyright © 2014, James L. Welsh II

Dedication

I dedicate this work to my partner, Michael Dean Raabe. His love sustains me and his remarkable commitment to his craft inspires me to try harder in everything I do.

Acknowledgments

I would like to gratefully and humbly acknowledge my committee co-chairs, Dr. James R. King and Dr. Jenifer J. Schneider. Over the years, as my mentors, advisors, collaborators, and friends, they have provided me with the inspiration and guidance to pursue a life of inquiry and scholarship. It is impossible to overstate their influence on my work and on a generation of literacy scholars to emerge from this institution. I learned much of what I know about scholarship from working alongside Jenifer and Jim as a part of the Contemporary Literacies Collaborative at USF.

I would also like to acknowledge the mentorship and support of my committee members, Dr. Danielle Dennis and Dr. Liliana Rodriguez-Campos, without whom this work would not have been possible.

I have benefitted greatly from the work of my professors. Dr. Kathryn Laframboise provided invaluable feedback on a concept paper that grew into this study. Dr. Ann Hall mentored me through my first research conference and has been a truly endless source of inspiration through the years.

I owe a debt to Dr. AnnMarie Gunn who provided feedback on my work and gave me both pep talks and kicks in the pants when each was warranted. I would also like to thank Barbara Peterson for her support, feedback, and friendship. I know no scholar more talented or more committed than she. My thanks also to my friend and collaborator Dr. Deborah Kozdras, one of the most creative scholars I know, who read drafts, discussed ideas, gave me feedback on my work.

Dr. Roy Winkelman has been my mentor and my friend for many years. Roy's patience, wit, intellect, and remarkable penchant for devil's advocacy has guided me past many rocky shoals. I am inspired and challenged by Roy's lasting contributions to education.

Thanks also to Dr. Kate Kemker, Dr. Christine Harmes, Dr. Dick Puglisi, Dr. Brent Lamons, Russell Brumfield, Jason Fortner, Regina Romeo, Nathan Wolkenhauer, Michael Ruppel, James Roth, Jason Tucker, Tara Schroeder, and Russell Hires for their encouragement and advice. Thank you to James Seaman, Dr. Chet Lyskawa, Jason Perry, Ashley Marando, and Ryan Dial who provided technical and web design information that informed my study. I gratefully acknowledge the assistance of the wonderful staff of the Office of Graduate Studies and the USF Library in the preparation of my manuscript.

I would like to acknowledge the many talented authors, illustrators, programmers, and designers whose work on children's websites goes largely uncredited. Their work brings joy to millions of children who could also be inspired by knowing what happens "behind the scenes."

Finally, I would like to acknowledge the support of my amazing family and friends. I have been extremely fortunate to encounter so many intelligent, challenging, funny, and inspiring people in my life. You have enriched my life in ways too numerous to mention. Any success that I have experienced is due in large part to the love and support of my friends, my family, and especially my mother, Ruth Gremer Fowble.

Table of Contents

- List of Tablesv
- List of Figures vi
- Abstract viii
- Chapter One: Introduction 1
 - Statement of the Problem.....3
 - Purpose.....4
 - Overview.....4
 - Research Questions.....6
 - Significance of Study7
- Chapter Two: Review of the Literature 10
 - Purpose and Overview10
 - Genre.....11
 - Introduction to modern approaches to genre12
 - Ancient Greece: Foundations of poetic and rhetorical genre.....16
 - Genre after Aristotle21
 - Genre in the Middle Ages24
 - Later thought on genre29
 - Studies of Hypertext and Online Reading30
 - Distinguishing between closed hypertext documents and online (Internet) texts30
 - Reader-focused studies of hypertext reading32
 - Text-focused studies of hypertext reading35
 - Website Content Analysis Research36
 - Methodological rigor38
 - Sampling methods.....38
 - Data collection39
 - Analysis methods42
 - Treatment of intertextuality44
 - Treatment of multimodality46
 - Treatment of multilinearity.....47
 - Treatment of advertising49
 - Conclusion50
- Chapter Three: Methods53
 - Purpose.....53
 - Overview.....53
 - Research Questions.....55

Research Design.....	56
Rationale	56
Uses of content analysis.....	56
Strengths of content analysis	57
Limitations of content analysis	58
Operationalization of website content	58
Phase 1: Instrument Development	59
Web Page Data Evaluation Checklist	59
Indexical variables	60
Web page ID (Indexical).....	60
Creation date (Indexical).....	60
Modification date (Indexical)	60
Page title (Indexical).....	60
Page address (Indexical)	60
Substantive variables	60
Subjects present (Substantive)	60
Primary subject (Substantive).....	61
Rhetorical variables	61
Functions present (Rhetorical).....	61
Primary function (Rhetorical).....	61
Formal variables.....	61
Element types present (Formal)	61
Primary element type (Formal).....	62
Element Data Evaluation Checklist	62
Indexical variables	63
Element ID (Indexical)	63
Creation date (Indexical).....	63
Modification date (Indexical)	63
Page title (Indexical).....	63
Page address (Indexical)	63
Element type (Indexical).....	63
Multimodal transcription of video content.	64
Element description (Indexical).....	64
Cluster ID (Indexical)	65
Substantive variables	68
Organization (Substantive)	68
Subject(s) (Substantive).....	68
Theme(s) (Substantive).....	69
Author position (Substantive).....	69
Audience position (Substantive).....	69
Rhetorical analysis.....	69
Rhetorical: Citation variables	70
Explicit external references (Rhetorical: Citation).....	71
Relation of explicit external reference to central claims (Rhetorical: Citation)	71

Characterization of explicit external reference (Rhetorical: Citation)	72
Implicit external references (Rhetorical: Citation)	72
Relation of implicit external reference to central claims (Rhetorical: Citation)	72
Characterization of implicit external reference (Rhetorical: Citation)	73
Rhetorical: Audience variables	73
Stated or implied audience (Rhetorical: Audience)	73
Audience prior knowledge (Rhetorical: Audience)	74
Audience attitudes (Rhetorical: Audience)	74
Persuasion (Rhetorical: Audience)	75
Call to action (Rhetorical: Audience)	75
Rhetorical: Author variables	76
Stated or implied author (Rhetorical: Author)	76
Experience claims (Rhetorical: Author)	77
Broadness (Rhetorical: Author)	77
Originality (Rhetorical: Author)	78
Idiosyncrasies (Rhetorical: Author)	78
Formal variables	79
Coh-Metrix	79
Word count (Formal)	80
Sentence count (Formal)	80
Paragraph count (Formal)	81
Syllables per word (Formal)	81
Words per sentence (Formal)	81
Sentences per paragraph (Formal)	81
Flesch reading ease (Formal)	81
Flesch-Kincaid grade level (Formal)	82
Concreteness (Formal)	82
Noun hypernymy (Formal)	82
Verb hypernymy (Formal)	83
Personal pronouns (Formal)	83
Displayed size (Formal)	83
Relative size (Formal)	83
Color (Formal)	84
Image type (Formal)	85
Image relationship to text (Formal)	85
Documentation of essential characteristics	86
Validation of the data evaluation checklists	87
Establishing reliability of data evaluation checklists	88
Phase 2: Page Level Analysis	89
Data collection	90
Background	90
Sampling method	91
Data analysis	92

Evidence of validity	94
Phase 3: Element Level Analysis.....	96
Data collection	96
Data analysis	96
Evidence of validity	97
Interpretation of Results.....	99
Role of the Researcher	100
Researcher's Relationship to the Study	101
Chapter Four: Findings	105
Websites Selected	107
PBS Kids.....	108
Nickelodeon	108
Seussville	109
FunBrain	109
National Geographic Kids.....	109
Question 1: Instruments and Methods of Study.....	109
Identification and parsing of website content	110
Text analysis tools.....	112
Color identification	112
Managing data.....	114
Question 2: Formal Characteristics of Selected Websites	114
Question 3: Substance of Selected Websites	120
Searching, browsing, and sharing	123
Question 4: Rhetorical Action of Selected Websites.....	124
Comments and message boards	130
Question 5: Evidence of Genre.....	131
Summary	134
Chapter Five: Conclusions	136
Summary	136
Conclusions.....	137
Establishing meaningful genres	137
Documenting multilinearity and multimodality.....	138
Rhetorically-based classification	139
Recommendations	139
Buried treasure	139
Knowing what changes; understanding why	140
SpongeBob is lost in time	141
Eye-tracking studies.....	143
Implications for instructional design	144
References.....	145
Appendix A: Coh-Metrix Output Fields.....	165

List of Tables

TABLE 1.1 Operational definition of terms	8
TABLE 2.1 Comparison of 3 genre-based approaches within the field of rhetoric.....	13
TABLE 3.1 Selected Indexical variables for elements in figure 3.5.....	66
TABLE 4.1 Multimodality as Demonstrated by Formal Characteristics of Selected Websites	115
TABLE 4.2 Rhetorical Characteristics of Selected Websites	125
TABLE 4.3 Participation options on selected websites	131

List of Figures

FIGURE 2.1. Fields of study that address genre.	12
FIGURE 2.2. Relationship between hypertext and the Internet.	32
FIGURE 3.1. Relationships between elements, pages, and sites.....	58
FIGURE 3.2. An example of a web page consisting of multiple elements	59
FIGURE 3.3. Overview of variables from <i>Element Data Evaluation Checklist</i>	62
FIGURE 3.4. Example of multimodal transcription of video.....	64
FIGURE 3.5. Detail from bio page featuring several elements.....	65
FIGURE 3.6. Example of cluster analysis.....	68
FIGURE 3.7. Detail from an illustration of the character “Leatherhead”	74
FIGURE 3.8. Image element depicting the character Donatello	84
FIGURE 3.9. Three image elements of different image types	85
FIGURE 3.10. A portion of a character biography page.....	86
FIGURE 3.11. First page of automatically generated site map.....	92
FIGURE 3.12. Non-linear structure.....	93
FIGURE 3.13. Hierarchical structure	93
FIGURE 4.1. Percentage of pages by primary element type on five sites	116
FIGURE 4.2. On PBS Kids, an extended informational text reading passage	117
FIGURE 4.3. On PBS Kids, a brief and decontextualized scientist biography	118
FIGURE 4.4. Typical arrangement of a property home page on nick.com.....	120
FIGURE 4.5. Nickelodeon content volume by property	121

FIGURE 4.6. Organizational pattern found on three of five sites	122
FIGURE 4.7. Organizational pattern illustrated with examples from Nickelodeon.....	122
FIGURE 4.8: Seussville character menu displaying links to activities and books to purchase	127
FIGURE 4.9. An example of cross-property promotion	128
FIGURE 4.10. Cross-marketing of properties	129

Abstract

This study establishes a comprehensive methodology for analyzing children's website content, based on both linguistic and rhetorical data, by employing defensible criteria to evaluate both qualitative and quantitative data. By employing genre theory as a prism for examining form, substance, and rhetorical action within children's websites, this study applies that methodology to a purposeful sample of five children's websites. Results of the analysis document the complex multimodal and multilinear nature of the websites studied and identify a possible new genre, the pop culture carousel website.

Chapter One: Introduction

When approaching a new form of digital communication with the assumption that one has encountered a new genre, the researcher is faced with a dual conundrum. As has been noted by many scholars, not only is the object of study elusive and unstable – what constitutes a “good” instantiation of a particular genre, where does a particular genre begin and end, what are constitutive features and how can they be measured? – but the term genre itself is ripe with ambiguity and broadness like few other descriptive labels. It seems that genre means something to both researchers and non-academics; to film critics, journalists, scholars of literature, information retrieval experts and computer scientists alike. But, problematically, it appears to mean something different to all of these experts. (Puschmann, 2009, p. 49)

Reading is a complex process. Successful reading draws upon the reader’s knowledge of the structure of the language, vocabulary, and background knowledge of the subject (Clay, 1991; Snow, 2002). Successful reading also draws upon multiple meaning systems, multiple areas of the brain, and multiple strategic approaches (Cope & Kalantzis, 2000). Interpretation of meaning is negotiated in ways that include text, images, color, and layout (Callow, 2008; Kress, 2009).

Reading is impacted by the genre of the text (Cope & Kalantzis, 1993; Devitt, 1996). Effective readers employ different strategies and activate different kinds of background knowledge based on the genre of the text (Hicks, 1997; G. Kamberelis, 1999; Pappas, 1991; Wollman-Bonilla, 2000). For example, one would not read a romantic novel in the same way one would read a bank statement, although both require comprehension and involve the same discrete

decoding skills. Similarly, one would not read a Shakespearean sonnet in the same way one would read search engine results online.

Reading is even more complex in online environments (Coiro, 2003; D. J. Leu & Kinzer, 2000). Online texts involve multiple possible paths through a text, multiple means for navigation, an ever-changing corpus, and a functionally infinite amount of text (Duke, 2012). In online reading environments, it is more difficult for readers to ascertain the author of a given work and more difficult to separate advertising from content (Coiro & Dobler, 2007).

Online texts have their own unique set of genres that are richly multimodal (Miller & Shepherd, 2004; Ratliff, 2003). Multimodal refers to the co-deployment of multiple modes in communication. Modes include written text, image, audio, video, color, gesture, and speech (Kress, 2003a, 2009). Websites exist in an environment that allows for seamless integration of images, color, audio, and video. Unlike traditional print environments, images and color can be incorporated into websites with no additional cost to the author and using no more specialized tools than a word processor. In the early days of website development, materials that were already developed for print distribution were converted to web formats. A print brochure, or a product information sheet, or an encyclopedia article was repackaged as a webpage, which was no more multimodal than the original. In fact, due to limitations of file size and download speed, frequently these webpages were less multimodal than their print siblings. As designers and developers have become accustomed to the multimodal affordances of web-based text, and as more of their audiences move online, and as technology has improved, websites have grown more multimodal. Online texts are not incidentally multimodal; they are integrally multimodal.

The systematic study of online genres reveals the forms, functions, and rhetorical contexts of these genres (Miller & Shepherd, 2004). Studies have examined such Internet-

specific genres as personal home pages (Dillon & Gushrowski, 2000), blogs (Herring, Scheidt, Wright, & Bonus, 2005; Miller & Shepherd, 2004; Puschmann, 2009; Ratliff, 2003), and online newspapers (Ihlström & Henfridsson, 2005; So, 2005; Wise, Bolls, & Schaefer, 2008). These genres have their own affordances and respond to unique exigencies (Bazerman, Bonini, & Figueiredo, 2009; Devitt, 2009). As the Internet matures, Internet genres grow to take advantage of the affordances of the medium and grow less like their antecedent forms (Bazerman et al., 2009).

Although many online genres have been studied, no study to date has examined an online genre geared toward children. Researchers who have investigated children reading online have confined their work to informational texts read for academic purposes. Although this is important work that directly informs classroom practice, it excludes the majority of what children read online. The spaces children choose to visit on the Internet remain largely unexamined by literacy research.

Statement of the Problem

Most of the time children spend online is spent in ways other than reading informational texts for academic purposes. Yet, current research on reading online focuses on reading informational texts for academic purposes. The websites most often visited by children remain unexamined by literacy research.

There is also little literacy research that applies the rich body of literature about reading multimodal texts to children's websites. The Internet is a multimodal medium, allowing the incorporation of text, audio, video, photographs, illustrations, and animations. Consequently, any complete analysis of website content must include multimodal analysis. Due to the complex nature of websites, some current website content analyses exclude the rich range of content

displayed on websites and instead include analysis of text without considering images (e.g., Gow, Lydecker, Lamanna, & Mazzeo, 2012; Opoku, Pitt, & Abratt, 2007) and no published website content analyses address multimodality from a social-semiotic perspective (Kress, 2009).

Research on the genres of children's websites is needed to fully understand the unique challenges and affordances facing children when reading in online environments. This research will support new pedagogical and analytic approaches to children's online reading behavior and experience.

Purpose

The purpose of this study is to establish a comprehensive methodology for analyzing children's website content, based on both linguistic and rhetorical data, by employing defensible criteria to evaluate both qualitative and quantitative data. By applying genre theory through a mixed methods content analysis, I will address gaps between research into online genres and research into children's reading on the Internet. This study is developed around the supposition that understanding the texts that children choose to read online, not exclusively informational or academic texts, is necessary in order to fully understand reading in online environments designed for children.

Overview

I will select a sample of five children's websites and conduct a content analysis (Krippendorf, 2004) focused on four kinds of variables for each website: indexical, substantive, rhetorical, and formal. Genre can be defined by substance, rhetorical action, and form (Bazerman, 1988; Devitt, 1996; Miller, 1984). My data categories correspond to these three with

the addition of indexical data¹. Substantive data refers to the content or subject matter of the page. Rhetorical data describes the rhetorical action being accomplished or contributed to with the element, as defined by Bazerman (1988) to include lexicon, citation, author, and audience. Formal data refers to measures of formality, complexity, and coherence. Following a relatively broad page-level analysis of five websites, I will proceed to a more in-depth element-level analysis.

Prior to selection of the websites for evaluation, I will develop two data collection and analysis instruments, the *Web Page Data Evaluation Checklist* and the *Element Data Evaluation Checklist*. I will establish the validity and reliability of the instruments using 50 purposefully sampled web pages from children's websites. Therefore, the study comprises three distinct phases: phase one, instrument development; phase two, page level analysis of five websites; phase three, element level analysis of one or more websites.

In this study, I proceed from an understanding of genre as socially-constructed, flexible, situated, and dynamic discourse forms (Chapman, 1995; Devitt, 1996; Jamieson, 1975; Kamberelis & Bovino, 1999; Miller, 1984). Although formal, structural elements are an essential part of defining a genre, a genre is most significantly defined as the site of social action (Miller, 1984). Genres are deployed by people to take action through communication with other people. Genres change constantly in response to the needs of the people who use them. Still, as rhetorical exigencies recur, rhetorical solutions to those situations also recur. To the extent that those individual rhetorical solutions resemble one another in substance, rhetorical action, and form, genres exist. As Schryer expressed it, genres are "stabilized-for-now or stabilized-enough sites of

¹ Indexical data, such as a unique sequential number, identifies the page or element and is not meant to be interpretive.

social and rhetorical action” (1993, p. 204). The concept of genre is significant to the study of reading in contemporary contexts precisely because online texts represent a rapidly shifting literacy landscape in which “stabilized enough” may last for a very brief period of time.

Research Questions

I will address the following research questions:

1. What methodology allows for the comprehensive study of genre of children’s websites?
 - a. What instruments can facilitate the systematic evaluation of children’s websites?
 - b. What methods of analysis can be applied to understand the genre of children’s websites?
2. What is the nature of the formal elements of selected children’s websites?
 - a. What are the formal textual elements associated with selected children’s websites?
 - b. What are the formal visual elements associated with selected children’s websites?
 - c. What text-image relationships are found within selected children’s websites?
3. What is the nature of the substance of selected children’s websites?
 - a. What are the organizational and thematic units of within selected children’s websites?
 - b. What are the topics addressed by selected children’s websites?
 - c. How are the author and reader positioned by the text?
4. What is the nature of the rhetorical action of selected children’s websites?
 - a. What stated or implied communicative purposes are served by selected children’s websites?

- b. Within what context or community do selected children's websites operate?
- c. Who is the intended or purported audience of selected children's websites?
- d. What attitudes and prior knowledge does the author presume the reader to have?
- e. How is authorship of selected children's websites identified?
- f. How are implicit and explicit external references used to position the website by the author?
- g. What evidence points toward antecedent genres that fulfilled the rhetorical purposes of the genres represented within selected children's websites prior to the existence of children's websites?

5. What evidence supports or refutes the identification of a genre or genres within the body of children's websites examined within this study?

Significance of Study

This study addresses gaps in two bodies of research (online genres and children's reading on the Internet) and provides a foundation for a more subtle and powerful understanding of the reading tasks faced by children in online settings. By understanding the content children commonly face in online settings, educators will be better positioned to prepare children for the reading tasks associated with contemporary literacy. The study of the genres of children's websites will reveal the complexities that confront children when reading online.

In addition, this study will support other research into literacy in online environments. Regardless of the perspective of those researching literacy in online settings, we can agree that children are spending time in online environments and that what children encounter in online environments matters. From a cognitive perspective, one might be interested in reading level, proposition density, or some other proxy for text complexity. A consideration of reading from a

sociocultural standpoint may prompt an interest in the societal pushes and pulls evident in the text or in how authority is established, or how authorship is represented. A researcher who adopts a New Literacies perspective may be interested in the degree to which one text or another is multimodal or multilinear. In any case, whether studying readers or writers and regardless of theoretical orientations, the nature of authentic online texts should be reflected within the research design and systematically characterized in the methods and instruments deployed. One researcher may be interested in studying how children read independently and then they subsequently test for their comprehension, while another may be interested in children reading collaboratively and constructing multimodal communal representations of their lived literacy experience. In either case, some concern and focus needs to be placed on the nature of online content available to children. Researchers will be able to use the methods developed in this study to establish the similarity of their research conditions to authentic online environments, or to aid in the selection of online environments for study. In other words, whether a researcher is interested in studying readers or writers, or online contexts, the current study will proffer a method of accounting for the properties of online text(s).

TABLE 1.1
Operational definition of terms

Antecedent genre	A genre that preceded the genre of interest and which fulfilled the same rhetorical purposes or responded to the same exigencies prior to the existence of the genre of interest. (Devitt, 2009; Jamieson, 1975)
Audience	The viewers or readers of a rhetorical action (Jamieson, 1975). In this context, viewers or readers of a website.
Children's website	For the purposes of this study, a children's website is a website with children as the intended audience.
Exigency (Rhetorical)	A need or purpose that can be addressed by human interaction, specifically communication (Bitzer, 1968).

TABLE 1.1 (Continued)

Generic	Adjective meaning of or related to genre. The common definition of generic is characteristic of a class, group, or genus, although it is used informally to mean not specific or even of low quality. Within genre studies, generic is used to describe or compare characteristics of a genre or genres, as in “generic distinctions,” “generic contexts,” “generic features,” and “generic resources.”
Genre	A flexible, dynamic, socially-constructed grouping of texts that share rhetorical purposes, formal elements, and subject matter that is understood by a group of rhetors.
Multimodal text	Text broadly construed to include multiple meaning-making systems, such as text, images, video, color, and layout.
Non-textual elements	Specifically referring to meaning-making elements not constructed from alphabetic text, such as images, video, color, and layout.
Reading	Interpretation of texts, including multimodal texts, to make meaning
Rhetor	The speaker or author. One who takes rhetorical action (Bitzer, 1968).
Rhetorical situation, action	The rhetorical situation includes the exigency, the rhetor, the rhetorical action, and the audience. Rhetorical action refers to speaking or writing to accomplish a purpose (Bazerman et al., 2009; Bitzer, 1968; Devitt, 1996).
Website	A set of related web pages, usually sharing a common server

Chapter Two: Review of the Literature

That children are engaged with the Internet is well known. What is less known is the nature of the literacy landscape in which they choose to spend time. Although many studies have examined the reading children do online in the context of academic reading, little is known about the texts with which they choose to engage on the Internet. Moreover, many would not consider the websites children choose to visit to be texts at all. Nonetheless, children read and write in online environments that they choose to navigate, within what are arguably literacy's new texts.

Purpose and Overview

The purpose of this study is to establish a comprehensive methodology for analyzing children's website content, based on both linguistic and rhetorical data, by employing defensible criteria to evaluate both qualitative and quantitative data. By applying genre theory through a mixed methods content analysis, the study addresses gaps between research into online genres and research into children's Internet reading. This study is developed around the supposition that understanding the texts that children choose to read online, not exclusively informational or academic texts, is necessary in order to fully understand reading in online environments designed for children.

The analytical approach employed in this study relies heavily on the concept of genre as it has developed in Western culture in a number of different disciplines. As such, it is informative to explore briefly the history of genre, which I do in the next section. Following the history of genre, I will present an overview of research related to reading hypertext and reading online. Finally, I will present an overview of content analysis research focused on websites.

Genre

The story of genre is the story of tension between tradition and invention. On one hand, intelligibility of language depends upon adherence to a common system of meaning. Without a common understanding of the meaning of words—and the possible patterns of words within a language—an attempt to communicate will fail. From a semiotic perspective, communication requires a common understanding of a system of symbols. A completely novel system of symbols would result in no communication. Genres can be understood as recurring patterns of communication that emerge in response to similar rhetorical situations (Bazerman, 1988). As Charles Bazerman put it, "Genres arise in social processes of people trying to understand each other well enough to coordinate activities and share meanings for their practical purposes." A common system of symbols by itself is insufficient. Without adherence to generic norms in the use of language, communication may fail.

In contrast to the limits imposed by language structure and expectations for it, the use of language is itself animated by the expressive desires of the language user. The speaker/writer uses language in an attempt to express feelings, thoughts, and experiences. Language users also innovate. Languages grow and change to accommodate new meaning, new contexts, new understandings. The use of language is inherently creative and that creativity shapes language; and yet, paradoxically, must respect the functions imposed by the structure of language.

So, language exists as a negotiation between novelty and tradition, between creativity and adherence to established forms, between expression and communication. In addition, Kress characterized a binary between "authenticity (a question of fit with personal truth) to appropriateness (a question of fit with social truth)" (Kress, 1999, p. 464). Therefore, language use must be a negotiation between intentions for fidelity and the aims of expression. The aim of

fidelity, which attempts to maintain the integrity of a shared system of meaning, animates the broad historical trend toward faithful reproduction of existing forms and the emphasis on accuracy and indoctrination of rhetors. Literary and social movements consistent with this animus are neo-classicism, Medievalism, and Scholasticism, among many others. By contrast, the aim of expression, representation of the rhetor's vision, is primary to Romanticism and Aestheticism, again, among many others. Some rhetorical situations—and some literary movements—reward and encourage innovation. Some rhetorical situations—and some literary movements—reward and encourage fidelity to existing forms.

My purpose in this section of the literature review is to illuminate the theoretical traditions that undergird current methodologies deployed in genre-based research, to define the theory behind my definition of genre, and to explain how genre relates to online digital texts. In this section, I will also explore the commonalities, divisions, and interstices between reading studies, writing studies, and studies of literary criticism, factors which complicate any study of genre.

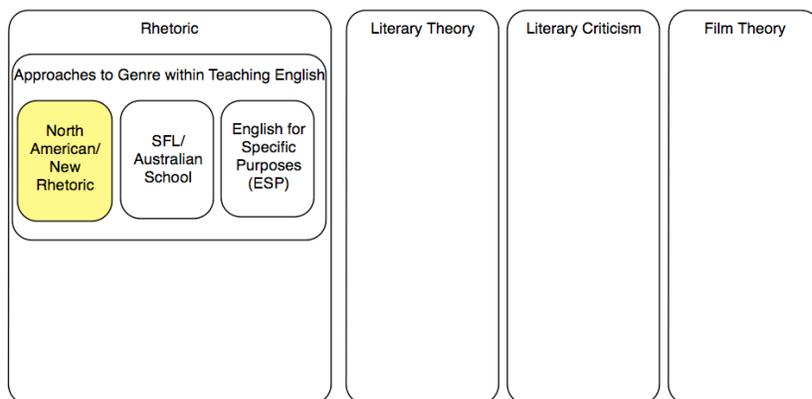


FIGURE 2.1. Fields of study that address genre. This study is situated within the North American/New Rhetoric theoretical perspective on genre.

Introduction to modern approaches to genre. The genre theorists with whom this study is most closely aligned are Carolyn Miller (Miller, 1984; Miller & Shepherd, 2004) and

Charles Bazerman (Bazerman, 1980, 1988, 1997; Bazerman et al., 2009). According to Hyon (1996), both Miller and Bazerman are associated with North American New Rhetoric studies. New Rhetoric (NR) is one of three predominant approaches to genre theory as it concerns the teaching and learning of English. The other two are Systemic Functional Linguistics (SFL), or the Australian school (Halliday, 1985, 1989), and English for Specific Purposes (ESP) (Flowerdew, 2000; Swales, 2009, 2011). These three approaches are all situated within, or related to, the broader field of rhetoric. Beyond rhetoric, my understanding of genre theory is informed by scholarly work in many other fields, including literary theory (Fry, 2012), literary criticism (Habib, 2008), and film theory (Altman, 1984, 1999).

TABLE 2.1

Comparison of 3 genre-based approaches within the field of rhetoric.

	Non-fiction genres, workplace genres	Social Justice, marginalized, non-mainstream	Primarily Children	Primarily Adults	ELLs
NR	X			X	
ESP	X			X	X
SFL	X	X	X		X

All three of these approaches, NR, SFL, and ESP, focus on students writing successfully within professional and academic genres. As Hyon (1996) states, all three of these approaches to genre have an “overarching concern with helping students become more successful readers and writers of academic and workplace texts” (p. 700). Table 2.1 displays details about the three approaches. While Systemic Functional Linguistics (SFL) focused on children, English for Specific Purposes (ESP) and New Rhetoric (NR) were both developed with adults and/or

graduate students in university settings. SFL's focus on issues of social justice and helping students acquire "genres of power" springs from SFL's development to meet the needs of children from marginalized, non-mainstream cultures. As Kress described it, "If there was a predictability and recognizability of text-forms, then, so the argument went in Australia, these were things that should be made available as explicit knowledge for all learners in school. This was a political and a pedagogic move... an explicit curriculum was seen as the essential underpinning of an equitable curriculum" (Kress, 1999, p. 463). Language arts instruction was used to attempt to balance access to "cultural capital" (Bourdieu, 1991) within a pluralistic and diverse society. ESP and SFL were both developed to improve writing instruction for English language learners (ELLs). Both of these approaches were developed in the early 1980s in response to the concern that writing instruction was overly focused on developing personal identity and narrative style and that writing instruction was not preparing students to write "factual", non-fiction texts—genres that they would need in the workplace. According to Kress, the shift in the focus of writing instruction in the 1980s centered on a shift "from desire to power" (1999, p. 463). In other words, the new genre paradigms of writing instruction de-emphasized the desire of an individual toward self-expression, such as the approach advocated by the New Hampshire school, and instead focused on the ability of the individual to access the genres of power within the dominant culture.

Why was so much of writing instruction prior to the 1980s focused on personal expression? The movement to make writing instruction more authentic to students' lives, to teach students to use writing to express identity and to develop personal reactions to literature was most clearly embodied by the Anglo-American Seminar on the Teaching of English, held in 1966 at Dartmouth College in Hanover, NH, and otherwise known as the Dartmouth Seminar.

Comprised of forty-seven influential scholars, the Dartmouth Seminar most significantly addressed the definition of English as a subject, the place of literature in the curriculum, and the training of English teachers. In subsequent years, two reports were published based on the work of the seminar participants: *The Uses of English*, by Herbert J. Muller (1967), intended for general audiences, and *Growth Through English*, by John Dixon (1967), intended for the professional community. Those texts, in combination with six monographs about the findings of the conference published by the National Council of Teachers of English (NCTE), had enormous impact on English instruction for decades. The Dartmouth Seminar identified three broad historical eras of writing instruction that matched the needs of their respective periods (Dixon, 1967). During the first period, writing instruction was skills-based to meet the needs of “an era when initial literacy was the prime demand” (p. 1). The second period “stressed the cultural heritage, the need for a civilizing and socially unifying content” (pp. 1-2). Finally, the current era should focus “on personal growth: on the need to re-examine the learning processes and the meaning to the individual of what he is doing in English lessons” (p. 2). The Dartmouth scholars endorsed a general shift in writing instruction away from a focus on product toward a focus on the processes of writing. Rather than teaching students to slavishly reproduce canonical genres, the Dartmouth participants offered that writing instruction should focus on developing students’ understanding of writing as a creative or expressive process (Sublette, 1973). At the same time, the Dartmouth Seminar endorsed a move away from the traditional canon of Western literature, e.g. the works of Tennyson, Dickens, Shakespeare, Milton, and others, and toward “books embodying diverse visions of life and beliefs about values... moral complexity, ambiguity, and paradox” (Muller, 1967, p. 93). It is worth noting that there was contemporary disagreement about these directions. For example, Sublette decried the substitution “of Malcolm X for

Shakespeare, of sociology for literature and art, of films for writing, of intensity for lucidity, and of personal gut reactions for scholarship” (1973, p. 354). More recent criticism includes the idea that the sort of curriculum advocated by the Dartmouth Seminar presupposes a culturally and economically homogenous and advantaged student population (Kress, 1999, p. 463) that presumably does not exist in most classrooms.

From the previous perspective, it should be apparent that the New Rhetoric school of genre studies, which immediately followed the New Hampshire school, was a reaction to that previous predominant scholarship, exemplified by the Dartmouth Seminar, which was in turn a reaction to the preceding era. As Leu and Kinzer put it, “the forms and functions of literacy, as well as literacy instruction itself are largely determined by the cultural forces at work within any society” (2000, p. 111). Genre theory is connected to and is shaped by the political, social, and philosophical traditions of Western civilization. In order to grasp the depth and significance of the roles of genre in reading and writing, it is informative to take an even broader historical view and consider the ways in which our current concepts of genre rely on the historical contexts, which gave them form. In order to understand the more current assertions of Bazerman or Miller, it is necessary to consider more than just the pushes and pulls of the past fifty years. As one historian of philosophy put it, “The historian who isolates ‘philosophical’ arguments of the past from their contexts, studying them without reference to the presuppositions and aims of their proponents, will not understand them” (Marenbon, 1988, p. viii). The current treatment of genre is better understood in relation to broad, even ancient, movements in literature and rhetoric that establish genre as the site of tension between invention and tradition, both then and now.

Ancient Greece: Foundations of poetic and rhetorical genre. In Western civilization, the origins of genre can be traced back to ancient Greece. The earliest known expression of the

concept of genre comes from Aristotle's *Poetics* (c. 335 BCE). At the outset of that work, Aristotle proposed the "various kinds" of poetry and the "essential quality of each" (Telford, 1961, p. 1). That earliest list of genres is composed of "Epic poetry and Tragedy, Comedy also and Dithyrambic poetry." Aristotle also described examples of works within each genre, defining each genre by the common characteristics of its members. As noted by Devitt (2000), Aristotle's concept of genre was predicated on the assumption that generic distinctions exist as an attribute of a text. Aristotle's *Poetics* formed the foundation for literary genre theory, as well as genre theory in drama, art, film, photography, and music (Altman, 1999).

It is tempting to consider the ancient Greeks as a singular cultural-historical unit because their civilization is so remote from our own. However, for the most part, Aristotle was not referring works contemporary to him when he developed poetic genres. Aristotle lived from 384 BCE to 322 BCE. The poetic works upon which his classification system was based were written from about 800 BCE to about 400 BCE, including the epic poetry of Homer (7th or 8th century BCE), the Cyclic poets (7th to 5th century BCE), and the works of the great dramatists Aeschylus, Sophocles, Euripides, and Aristophanes (525-406 BCE). In other words, in Aristotle's time, he was evaluating a well-established poetic and dramatic tradition produced by previous generations, including works that were, in some cases, hundreds of years old.

Separately from this tradition, Aristotle also formed the foundation of rhetorical genre theory in his delineation of kinds of oratory in *Ars Rhetorica* (Aristotle, Buckley, & Hobbes, 1900). Again, Aristotle was working with texts from preceding generations. The discipline of rhetoric was founded by Corax in 476 BCE on the island of Syracuse (Habib, 2008). Rhetoric was further developed by a group of thinkers known collectively as the Sophists—Protagoras, Gorgias, Antiphon, Lysias, and Isocrates—working mostly in the century preceding Aristotle

(Habib, 2008). Ancient rhetoric addressed the effectiveness or persuasiveness of communication between speaker (or writer) and audience. Aristotle identified the three genres of rhetoric as deliberative (focused on future actions of the state), forensic (focused on past actions, and the basis for our modern legal system), and display (focused on the present, typically either praising or denigrating a subject). He went on to describe the parts and subparts of speeches and classification thereof.

Reading the earliest descriptions of rhetorical forms, one is struck by the similarities to various modern systems of writing instruction and evaluation. For example, from the time of Aristotle (4th century BCE) until the time of Cicero (1st century BCE), the characteristics or traits of rhetoric were commonly identified as 1) invention, 2) arrangement, 3) style, 4) memory, and 5) delivery (Habib, 2008). These concepts are not so far removed from characteristics identified in the Six Plus One Traits of Writing (Culham, 2003): ideas, organization, word choice, voice, conventions, sentence fluency, and presentation.

Aristotle's teacher Plato had condemned the techniques of rhetoric as dangerous because rhetoric could be used to convince an audience of an argument that was untrue and because effective use of rhetoric did not require the rhetor to be an expert in the field of the argument, but merely a talented speaker. Plato argued "that the rhetorician is a non-expert persuading other non-experts. He never need know the actual facts of a situation; he needs no expertise, merely a persuasive ploy" (Habib, 2008). Plato believed that rhetoric's focus on communication and argumentation, and not on truth, was dangerous.

By contrast, Aristotle presents rhetoric as useful because it can serve the cause of truth. Aristotle argued that rhetoric is the counterpart of the logical argument. With regards to Plato's criticism that rhetoric could be used to further an untrue argument, Aristotle said that the true

position is the naturally stronger position. Aristotle posited that rhetoric does not belong to any specific discipline; it can be applied regardless of the field of study. Where Plato places absolute importance on truth, Aristotle places importance on *communicating* truth, and therefore finds value in rhetoric.

Therefore, in many ways Aristotle's work must be understood as a reaction to the work of his teacher Plato. That reaction is exemplified both in Aristotle's position on rhetoric and in his position on poetics. Plato had posited two important binaries. First, Plato viewed poetics as the opposite of philosophy; poetics appealed to emotion where philosophy appealed to logic. In simplest terms, poetry presents a chaotic view of the world while philosophy presents an orderly view of the world. Most importantly, Plato viewed philosophy as the only path that could lead one to discover truth. Famously, Plato said that a poet should be respected and revered – and cast out because, though worthy of respect, he had no place in an orderly society (McPherran, 2010). By contrast, Aristotle argued that poetics (artistic expression) is another valid way to pursue understanding. Aristotle believed that all arts imitated the actions of men, and by presenting these imitations to audiences, artists could reveal universal truths (Habib, 2008).

In the second binary, Plato viewed the physical world as an imperfect projection of a separate world of perfect Forms (Dancy, 2004; Habib, 2008). In a classic example of this distinction, one can imagine a perfect triangle, but any triangle that one creates in the real world is bound to be imperfect in some way due to the limitations of the tools or the skill of the person creating it. Likewise, all things in the physical world are imperfect versions of perfect ideals, or Forms. By extension of this principle, Plato denigrated the work of the poet. In Plato's example of the three beds, the first bed is the perfect Form of the bed, made by god. The second is the physical bed made by the carpenter. The third is the imitation of the physical bed made by the

poet or painter. The second is an imitation of the first, and the third is an imitation of the second – an imitation of an imitation, twice removed from truth and perfection. For Plato, reality and truth exist beyond our world and philosophy is the only way to pursue that truth. However, for Aristotle, reality and truth can be found within the specifics of our world. This is the basis of Aristotle's observation and classification of the natural world. Accordingly, for Aristotle poetics and rhetoric should be classified like every other area of human knowledge. For Aristotle, everything has its place and purpose and function within the divine order. The purpose of art is to obtain knowledge of universals, the same as the purpose of any other act of creation. So, in keeping with his efforts to categorize the natural world, Aristotle described categories, or genres, of poetry and of rhetoric. Like all other things he classified, he looked for common characteristics. Aristotle handed down a definition of genre as a category that could be defined by characteristics of a text, but also by effect on audience, as in the case of tragedy.

Aristotle's philosophy is based on inductive reasoning, making generalizing statements based on many specific observations. Plato's philosophy is based on deductive reasoning, following general statements to findings of specific truths.

When applying Aristotle to the past fifty years of genre theory, several possible connections become apparent. Modern rhetorical approaches to writing instruction that emphasize genre (SFL, NR, ESP) take up an Aristotelean position with regard to rhetoric that emphasizes the importance of understanding forms and strategies of effective presentation of information, not merely on understanding truth. The findings of the Dartmouth Seminar are evocative of the Aristotelean idea that artistic expression is a valid way to seek truth. Conversely, they could also be viewed as evocative of the Platonic idea that expression of truth is a more important goal than rhetorical elements. These extrapolations, however, would leap-frog over the

intervening two thousand years of developments in rhetorical and poetic theory and practice, as the following sections illustrate.

Genre after Aristotle. Greek concepts about art and philosophy, including the key concepts of Plato and Aristotle regarding poetics and rhetoric, were taken up and spread across the vast empire of Alexander the Great, at one time a student of Aristotle. Upon the death of Alexander, the empire was fractured, but key foundational concepts of Greek culture remained within the shattered pieces of the former empire and those concepts continued to develop. Centuries later, the lands of Alexander's former empire were absorbed by the growth of the Roman Empire and the Greek ideas were taken up and spread even further across the ancient world.

Roman philosopher Cicero (106 BCE - 43 BCE) posited that the function of rhetoric was to establish a society based on justice and common welfare, not on strength (Habib, 2008). Unlike Plato, he believed that rhetoric and philosophy were compatible and even complementary, and that rhetoric could be used not only to seek truth, but to seek justice. Horace (65 BCE - 8 BCE) was a Roman lyric poet and critic. During his lifetime and for centuries afterward, Horace was enormously influential, more so than Plato, at times more so than Aristotle (Habib, 2008). Horace wrote *Ars Poetica* as advice from an established poet to young poets, in the form of poetry. In it, he discussed the relation of a poet to his work, the job of poetry to entertain and amuse but also to educate and provide moral guidance, and he explored the role of the audience. He also discussed the changes in genre and language throughout history. Horace recommended that poets carefully scrutinize their own compositions before releasing them to the public, because once a poem is released to the public, the poet no longer has control of his work. It belongs to the public and it can never be taken back. A poet can only decide to

destroy his work before it is released. Horace's argument about artists' destruction of their work provides an early perspective on the ways in which artists relate to their published works and to their audiences. Horace's argument also provides evidence that the intent of the author was only one consideration in the interpretation of poetic works at the time and that meaning was acknowledged to be socially determined, at least to some extent.

In *Ars Poetica* (Horace & Wilkins, 1964), Horace developed Aristotle's primarily descriptive work, which by that time (18 BCE) was treated as "received truth" (Altman, 1999), into a set of rules for poets to follow strictly in the faithful reproduction of forms. As Horace explained, "simula quelque locum tenant sortita decentem", translated as "Once you've got a genre, you've got to do it right" (Shechter & Schechter, 1987, p. 2). In terms of orientation toward genre, Horace's approach emphasizes faithful reproduction of existing forms as opposed to creative innovation. Horace's influence over the succeeding generations cannot be overemphasized. Our inheritance from Aristotle and Horace includes the concepts that genres exist, that genres have firm and definable boundaries, and that genre is determined, at least in part, by characteristics within the text. These ways of conceptualizing genre were maintained by genre theorists for more than a thousand years, and in apparent contrast to more current conceptions of genre as instantiations of specific social action.

Education, and early childhood education in particular, continues to feel the influence of the work of Cicero's successor Quintilian (35 CE - 100 CE), a Roman rhetorician from Hispania. Quintilian emphasized the good character of the ideal rhetorician and went on to describe the ideal education for children. He said that children should learn to read and write, then learn grammar.

In analyzing poetry, the student must be taught to read closely, to specify the parts of speech, the feet and meter, to identify the correct usage of words, to know the various senses of a given word, to recognize all kinds of tropes, figures of speech, and figures of thought, to be acquainted with relevant historical facts, and, above all, to understand the merit in the way the whole work is organized (IO, I.viii.5-18). In general, the stories told by poets should be used to increase students' knowledge rather than simply treated as models of eloquence (IO, I.xi.2-6) (Habib, 2008).

Quintilian also specifically outlined the traits of a good teacher of children: moral, patient, clear, not reproachful, not too affable, practical, and adaptable.

In his influential text *On the Sublime*, the Roman philosopher known as Longinus argued that the sublime is genius in the use of language that does not merely convince an audience, but shatters what came before and transports them (Habib, 2008; Longinus, 1st c./1964). The sublime is defined by emotional rather than logical elements. Longinus noted five genuine sources of the sublime that separately have no value but when used together in a unified and organic way, achieve perfection: 1) strong ideas, 2) emotional appeal, 3) use of figures of speech and thought, 4) diction and metaphor, and 5) general dignity and elevation. Longinus, like so many others before and since, was describing criteria for great use of language. Longinus' main point was that the sublime is something that causes humanity to strive beyond its boundaries, and to approach the divine. According to Longinus, the highest purpose of mankind is to strive beyond human nature, a desire that is fundamental to humanity. Longinus believed in the power of language to transform our perception of the world and help us approach divinity. According to Habib (2008), *On the Sublime* became a major influence on literary criticism in the 17th century,

notably influencing French poet and critic Nicolas Boileau-Despréaux and German philosopher Immanuel Kant. Like Plato, Longinus affirmed a separation between the sensible world and the invisible world of perfection. Markedly unlike Plato, Longinus embraced emotional appeal and the tenets of rhetoric in seeking the divine. It is also interesting to note that nearly two thousand years ago, Longinus lamented that truly great literature was no longer being produced because people were too focused on money and possessions. Reasoning on genre by Longinus demonstrates that the concern over the commercialization of art is not a modern invention.

Genre in the Middle Ages. In the third and fourth centuries, Roman philosopher Plotinus revived, adopted, and expanded Plato's notion that the physical world is separate from and takes its meaning from the spiritual world (Habib, 2008). Plotinus founded an intellectual movement that was later termed neoplatonism. Plotinus and the neoplatonist philosophers that followed in his tradition tried to reconcile Plato and Aristotle, Plato and Homer, and more broadly, philosophy and poetry. Neoplatonism represents an important moment of transition between ancient philosophy and the ascent of the Christian church.

For Plotinus, the physical world, including matter, time, and everything that can be perceived, existed on the lowest level of perfection, which he referred to as “Here.” Above the physical plane, Plotinus’ “There” represented the world of intellect and thought, which also contained the “World-soul” from which proceeds all human souls. The “divine mind” presides over the intellectual realm. Above the plane of There, Plotinus identified “the One,” the eternal source of all things, the embodiment of truth and good. Plotinus asserted that everything that exists is an emanation of the divine, patterned to return to the divine. A philosopher's job is to help souls ascend to the divine. It is important to note that Plotinus never mentions Christianity. Any similarities to Christian doctrine are due to the fact that the works of Plotinus were

enormously influential on early Christian thinkers, such as Augustine (354-430). Augustine took from the work of Plotinus ideas that were complementary to Christian doctrine and discarded philosophical elements incompatible with Christianity. The work of Plotinus is an elaboration on the foundation set by Plato. Plotinus also believed that through art we are able to express the eternal better than we can through intellect, a view that seems to be in direct contrast to Plato's formulation that art is not a useful way to pursue knowledge. Neoplatonism became the conduit through which concepts of Greek philosophy gave form to Christian philosophy.

In addition to the work of Plotinus, the works of neoplatonists Macrobius (385-430), who compared Greek and Latin grammar and literary and philosophical language, and Boethius (480-524), an early Christian whose focus was the preservation of classic Greek philosophy, were seminal texts in the transition to Christian philosophy in the Middle Ages. Centuries later, as the Roman Empire dissolved, the works of the neoplatonists played an important role in transmitting Greek scholarship of antiquity to the Latin west throughout the Middle Ages because of the translation of these ideas from Greek into Latin (Gerson, 2013; Marenbon, 2013). Original works of Plato, Aristotle, and most other Greek and early Roman philosophers were not available during much of the medieval period and their ideas were only accessible through reference and reaction in the writings of later Roman and early Christian philosophers (Marenbon, 1988).

The Roman Republic died with the defeat of Antony in 27 BCE and the Roman Empire was born, ruled by emperors until the fall of Rome in 410. In the Empire, public speaking was increasingly restricted. Rhetoricians began to focus more on style than on the content of arguments. Christianity became the religion of the Empire in 380 and the classical methods of the Greeks were widely rejected. In 426, Augustine (354-430) argued for the importance of

rhetoric for spreading Christianity. Augustine used Cicero's work as a guide for priests in effective communication.

The early Middle Ages saw the rise of the Christian (Catholic) church and the consolidation of power in the person of the Pope. The Catholic Church preserved classical knowledge through monastic traditions where monks were responsible for copying and translating texts. The interest of the Church was to preserve unity and order over across geography, culture, and time. The Germanic tribes that overtook the Roman Empire had a legal, economic, and political structure that centered on local institutions and de-emphasized centralized authority and focused on local institutions, such as the village, the monastery, and the local church. The central administrative structure of the empire developed into feudalism. The Catholic Church encouraged a static and hierarchical view of life in which every person had their place in the divine order established by God (Habib, 2008). In literate activity, this emphasis favored an adherence to tradition and faithful repetition of established forms.

Most intellectual work was driven by classical (Greek and Roman) thought or Christian reaction to classical thought. Christian scholars continued developing surviving classical works on grammar and poetic genres. The works of Cicero and Virgil were basic texts of education in the early Middle Ages. The foundation of education was the “trivium,” or the three roads: rhetoric, grammar, and logic. At this juncture, the earlier boundaries between poetics and rhetoric were blurred. Later, under Thomas Aquinas (1224-1274) and his contemporaries, rhetoric was subordinated to logic. Art and literature were subordinated by the Church to saving souls. Scholars within the Catholic tradition during the Middle Ages took one of two very different approaches to literature: to distance Christianity from the “pagan arts” of antiquity or to continue to appropriate the philosophy and rhetoric of antiquity to serve the aims of Christian

faith (Habib, 2008). The former viewpoint at times strongly endorsed the rejection of all art as a distraction from the truth of God, a position that parallels Plato's dismissal of poetry and rhetoric in favor of logic.

Augustine of Hippo (354-430), better known as Saint Augustine, was enormously important as a bridge from classical philosophy to Christian philosophy. Augustine was born a Roman citizen into a wealthy family in the Roman province of Africa. His mother was Christian, his father Pagan. Augustine was well educated and chose to follow the Manichaean religion, a popular Gnostic religion at the time. Christianity became the official religion of the empire when Augustine was twenty-six and he converted to Christianity seven years later at the age of thirty-three. By the time Augustine was forty-one, he had risen to the level of bishop within the Christian church. The Roman Empire fell when Augustine was fifty-six. He defended Christianity from those who associated it with the fall of Rome. Augustine served as the Bishop of Hippo until his death at the age of seventy-six.

Like other Christian thinkers at the time, Augustine's philosophy could be described as metaphysical realism (Habib, 2008), "insisting that reality is spiritual and that sense-perception and observation of the world were not reliable avenues to truth" (p. 68). This philosophy is consistent with Plato, as Augustine was influenced by neoplatonist philosophers. To Augustine, classical literature, and indeed the observable world, was worthy of study only for what it could reveal to us about metaphysical truth, i.e. about God. Spiritual knowledge is the only justifiable end and all other knowledge pursuits are a means to that end. The physical world should be interpreted as a sign system. As such, Augustine distinguished between natural signs (e.g. smoke signals fire) and conventional signs, which are created for human communication. "No one uses words except as signs of something else; and hence may be understood what I call signs: those

things, to wit, are used to indicate something else. Accordingly, every sign is also a thing; for what is not a thing is nothing at all. Every thing, however, is not also a sign” (Shaw, 1886). Augustine wrote guidelines for interpreting scripture, including instructions for determining those elements within texts which should be read literally and those elements within texts that should be read figuratively. A literal interpretation of a symbolic passage represented a “miserable slavery of the soul,” (Shaw, 1886) according to Augustine. Man can use reason and science to understand the “Word of God” only to a certain degree. For Augustine, “ultimately, this Word stands above human language and reason, and men must ascend allegorically from a literal understanding of their world to a symbolic view of it as a small part in a vast scheme which both subsumes and gives meaning to it. The world must be understood as the Word of God” (Shaw, 1886). This kind of metaphorical view of the physical world can lead to the confiscation of property, the subjugation of will, and ultimately to the destruction of people. If the physical world is primarily allegorical, then a leader, a government, or a church can impose “the will of God” to communicate a message. According to Augustine, pagan knowledge should be appropriated if it is useful and condemned if it is hurtful.

Geoffrey of Vinsauf (ca. 1200) wrote prominent pieces on rhetoric and poetry that drew heavily on classical concepts, including a hierarchy of genres. Geoffrey identified poetry as a branch of rhetoric. Unlike many of his contemporaries who saw poetry merely representing nature, Geoffrey thought that nature could be transformed by poetry and that poets needed to be original and not retread old ground.

Grammar was far more important and a broader concept in the Middle Ages than it is today (Marenbon, 1988). The study of grammar gave one access to the international intellectual conversation, to law, to theology, and in sum, to power. The study of grammar was the

foundation of learning and knowledge (Habib, 2008). Medieval grammar was concerned with a world of signs rather than a world of things. No element had independent, isolated meaning.

Later thought on genre. In the Renaissance (c. 1350-1700), rhetoric was revived by humanists, who focused on the works of Cicero and Quintilian. In the later Renaissance (c. 1600), the study of rhetoric was reduced to the study of style; rhetoric was subsumed under poetics, and essentially lost as a discipline. Rhetoric became about style, self-expression, introspection, and not about relating to an audience.

The Romantics (c. 1800-1850) reacted against Aristotelian poetics and rhetoric. The rise of written culture, brought about by the printing press and by the subsequent rise of secular education and national governments, meant invention and creativity were private and solitary. The Romantics emphasized creative expression over the reproduction of existing forms.

In the late 18th and 19th centuries rhetoric disintegrated because of increasing specialization of disciplines or areas of knowledge and because of a focus on literal language. Rhetoric became a negative term meaning emptiness and ornamentation. Yet, rhetoric remained central in politics and law, and subsequently, in business, particularly advertising and marketing. In education, writing instruction is one legacy of rhetoric.

Because rhetoric views text as a performance, not as a representation of the private thoughts of the author, rhetoric has experienced a revival in the 20th century through speech act theory, psychology, linguistics, and literary theory.

Imposed binaries can lead to overarching meta-narratives that distort the concepts one is trying to understand. However, binaries can be used, as in strategic essentialism (Spivak, 1990), to illuminate relationships between concepts and contexts. All models are limited in some way. No model represents all aspects of the thing it seeks to explain. Allowing for those limitations,

binaries can be used provisionally to propose relationships that may be useful. Bearing that in mind, I propose that the binary between Plato and Aristotle can be deployed to better understand the role of genre in communication throughout history. Plato, because he saw a dualism in the nature of reality, because that dualism was essential, and because the perfect world was separate from the real world, is more closely aligned with adherence to tradition, faithful reproduction of existing forms. Aristotle, because he saw the act of creation (of art) as a way of seeking truth and studying reality, is more closely aligned with the creative urge, innovation, creativity.

Studies of Hypertext and Online Reading

The body of literature that can broadly be described as pertaining to reading in online environments can be divided into two areas: the first area of inquiry is focused on readers of online texts with a goal of understanding and/or improving *readers' strategies* for reading online texts; the second is focused on the text structures of hypertext with the goal of improving and/or understanding the *design* of online texts. This dichotomy is consistent with a division noted in earlier hypertext research between user-centered and system-centered research (Rouet, Levonen, Dillon, & Spiro, 1996). All of these studies involve both readers and text to varying degrees, but the first group emphasizes the reader and the second group emphasizes the writer or the writer's text. This section begins by explicating the technical terms necessary to understand the various types of documents involved in these studies. Next, I will discuss the reader-focused research, the goals of this research, and the various approaches to generating or selecting text-bases for these studies. Finally, I will discuss research that focuses on the text structures of online texts.

Distinguishing between closed hypertext documents and online (Internet) texts. The earliest studies of online reading are studies of hypertext reading, not reading on the Internet, so it is necessary to distinguish between these environments before proceeding. Hypertext is

defined by *Encyclopedia Britannica* as a method of “linking of related pieces of information by electronic connections in order to allow a user easy access between them” (n.p., 2014). Hypertext can be composed in a number of different computer languages, the most popular language being HyperText Markup Language (HTML). The Internet is a worldwide network of computer networks. Although most people think of the Internet and the World Wide Web as synonymous, there are a number of other communication technologies that use the Internet, including email and instant messaging. Hypertext documents that are shared from computers connected to the Internet make up the World Wide Web, i.e. web pages and web sites. However, hypertext documents can also be used in environments that are not connected to the Internet. Many hypertext-based CD-ROMs were developed to take advantage of hypertext before widespread access to the Internet or before Internet connection speeds were fast enough to support the transfer of large numbers of images and other media. For example, as early as the 1980s, encyclopedias were developed in electronic versions available on CD-ROM. Likewise, an electronic version of a textbook and a children’s book on CD-ROM are examples of networks of hypertext documents that are not Internet-based. With the increase of access to high-speed Internet connections, most of these resources are now available online. Hypertext documents that are not online are sometimes referred to as “closed” hypertext documents to distinguish them from the hypertext documents that make up the World Wide Web, also known as online texts. Figure 2.3 illustrates these relationships.

Although hypertext is now primarily encountered on the Internet, some reading research continues to use closed hypertext documents to simulate Internet reading environments. Closed hypertext documents provide the researchers with a greater degree of control of the content

readers will encounter. However, that very control reduces the ecological validity of any inquiry that attempts to model online reading.

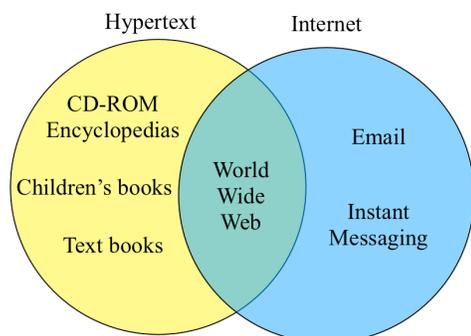


FIGURE 2.2. Relationship between hypertext and the Internet.

Reader-focused studies of hypertext reading. Not surprisingly, most studies that focus on readers reading online or in closed hypertext come from literacy researchers. I have organized these studies into three categories based on the theoretical approach of the authors: cognitivist (Salmeron, Kintsch, & Canas, 2006), constructivist (Cho, 2011; Coiro, 2007; Coiro & Dobler, 2007; Donald J. Leu et al., 2005; Zhang & Duke, 2008), and social semiotic (Chandler-Olcott & Mahar, 2003; Zammit, 2011).

A cognitivist approach (Salmeron et al., 2006) to questions of reading online focuses on specific questions of cognition during reading in an online environment. In a 2006 study, Salmerón and his colleagues took an expository text on air pollution and broke it up by subtopic into multiple separate hypertext pages. Participants in the study were told that after reading the document, they would be asked to answer questions. The implied construction of the reading task focuses on recall of basic facts within the text. After reading a passage, participants were able to choose one of two passages to read next and were later asked the reason for each selection. Participants either followed an interest strategy or a coherence strategy in selecting reading passage sequence. The authors reported that results from application of the strategies differed between “low-knowledge” readers and “high-knowledge” readers, with a coherence

strategy favoring those with little background knowledge. The authors do not report the source of the non-fiction text used in the experiment and only describes it as “expository.” While they do report the Flesch-Kincaid grade level and the Flesch Reading Ease calculations, they do not discuss ways in which this text is similar to or different from authentic website texts. Care is taken in constructing hyperlink titles that are indicative of the semantic content found on linked pages, but no information is available about how the links or the organizational structures found in this study compare with texts that readers actually encounter online. The goal of the Samerlon study was to isolate one aspect of online reading, different strategies for selecting reading path. Validity would have been strengthened either by more closely approximating actual Internet texts or by carefully describing the ways in which study texts differed from actual Internet texts.

Other literacy-focused studies adopting a social semiotic approach used a naturalistic design to observe readers and document literacy practices. Researchers in an observational study of two adolescent girls (Chandler-Olcott & Mahar, 2003) provided a detailed examination of reading and writing within affinity groups in online spaces, including an analysis of artifacts from online documents written by the participants. In a study involving year 4, 5, and 6 students (approximately 9 – 11 years old) in an Australian school, Zammit (2011) videotaped sessions within the classroom context. Students were completing academic tasks not specifically constructed for the study. Using multimodal transcription, Zammit was able to document literacy “moves” executed by the students in constructing their reading paths.

Studies that take a constructivist approach (Cho, 2011; Coiro, 2007; Coiro & Dobler, 2007; L. A. Henry, 2007; Donald J. Leu et al., 2005; Zhang & Duke, 2008) frequently involve the use of Internet search strategies. In some cases, this involves open searching on the Internet (Cho, 2011; Coiro, 2007; Zhang & Duke, 2008), and in other cases the reader is restricted to a

single website or a small set of predetermined websites (Cho, 2011; Coiro, 2007; Coiro & Dobler, 2007).

The reading tasks defined in these studies all involve participants reading informational texts for academic or academic-like purposes. The Online Reading Comprehension Assessment (ORCA) used by Coiro (2007) and the Digital Divide Measurement Scale (DDMS) used by Henry (2007) both grew from the work of the New Literacies Research Team at the University of Connecticut. Working with middle school students, the project documented successful online literacy practices and used that basis for developing reliable measures of online reading comprehension (Leu et al., 2009; Mokhtari, Kymes, & Edwards, 2008; O'Byrne & McVerry, 2009). One of the great challenges of this work has been the development of an online reading model that accounts for variation observed in different settings. Noting unanticipated differences in the relationships between “offline” reading comprehension measures and two different online reading comprehension measures, Coiro (2011) noted,

Understandably, one might wonder how two presumably valid and reliable instruments designed to assess aspects of the same online reading processes (e.g., online locating, evaluating, and communicating) could have such dramatically different relationships with the same standardized measure of offline reading comprehension. One possibility is that conceptions of online reading comprehension may very well be dependent on the complexity of the task. For example, as the structure of information in online reading environments moves from simple paragraphs on a page to complex lists of search engine results, and from multileveled information websites to highly populated blog interactions, it may be that each new text structure increases the likelihood that specialized sets

of online reading strategies (i.e., new literacies) might be applied in ways that appear less and less similar to offline reading strategies. If this were the case, it might very well serve to explain how measures of similar online reading processes situated in differentially complex online reading environments (e.g., a blog versus an online quiz interface) could have very different relationships with the same standardized measure of offline reading achievement (pp. 370-371).

For this reason, we can better understand online reading comprehension by understanding online text structures, or more importantly, online genres, and more specifically, children's online genres. Print texts from widely different genres are not expected to be read in the same ways. For example, reading a textbook is a different task with different demands and expectations than reading a novel (Afflerbach, 1990; Bazerman, 2004; Carter, Ferzli, & Wiebe, 2004). Treating online texts as a single genre yielding to a single set of reading strategies presents analogous problems.

Additionally, assessments, online reading instruction, and research would be improved by a better understanding of the ways in which study contexts systematically differ from real world reading contexts. While current online reading assessments provide valuable information, they may not adequately account for the role of multimodality, multilinearity, and intertextuality in online texts. In addition to the need for more observational studies of online reading called for by Leu and Kinzer (2000), there is a need to document the literacy dimensions of online spaces to use as reference when conducting observational studies and constructing similar tasks for assessment purposes.

Text-focused studies of hypertext reading. Studies that focus on hypertext come from literacy, computer science, and cognitive psychology. I have organized text-focused studies into

three categories based on the specific focus of the research: studies of reader performance with varying text structures (Ant Ozok & Salvendy, 2003; Barab, Bowdish, & Lawless, 1997; Barab, Young, & Wang, 1999; Bilal, 2000, 2001; Hofman & van Oostendorp, 1999; Lawless, Mills, & Brown, 2002; McDonald & Stevenson, 1998; Mohageg, 1992), studies of reader performance with varying available choices (Calisir & Gurel, 2003; Wise et al., 2008), and reader performance with hypertext versus traditional text (Kramarski & Feldman, 2000).

Within studies of the effects of varying text structure, several researchers used closed hypertext documents (Barab et al., 1997; Barab et al., 1999; Hofman & van Oostendorp, 1999; Lawless et al., 2002; McDonald & Stevenson, 1998; Mohageg, 1992). While this allows researchers to control for potentially confounding variables, it also introduces a serious threat to validity. If the text used in a study does not adequately model the characteristics of actual websites or other authentic hypertext documents, the study design may not be able to reveal significant information about the behaviors of interest. In Hofman and Oostendorp (1999) and Barab, Young, and Wang (1999), researchers created closed hypertext documents from a pre-existing informational text.

Taken as a whole, the focus of these studies is to direct website creators in methods of improving the organization of text structures on websites to aid in comprehension. Studies based in computer science tend to have a limited application of the reading task. Participants read a hypertext document and are tested on recall of basic facts.

Website Content Analysis Research

In this section, I will describe currently published content analysis research focusing on websites. The reason for this is two-fold. First, I seek to document and critique current practices in website content analysis. Second, I hope to demonstrate that current practices in website

content analysis would benefit from a literacy-focused methodology, the establishment of which is the aim of the current study.

Coiro and Dobler (2007) identified four differences between informational hypertext and traditional print information text noted within the theoretical and research literature: 1) Hypertext environments require the reader to be more active in making choices about the text path. 2) Hypertext provides less support to navigation and previewing of text. 3) Hypertext readers are expected to decode images as well as text. 4) Hypertext documents are more intertextual and have more immediate connections to other documents. As Coiro and Dobler noted, little empirical evidence has been gathered in these areas. These traits represent the considered and informed opinion of experts, but each should be systematically investigated. For the purposes of this review, I have operationalized these traits as multilinearity (Dresang, 1999; Portela, 2007; Tyrkkö, 2007) (the degree to which a text supports multiple reading paths), multimodality (Bezemer & Kress, 2008; Kress, 2009) (the degree to which the text involves multiple semiotic modes), and intertextuality (Lankshear & Knobel, 2007; Leu et al., 2009) (the degree to which a text implicitly or explicitly refers to other texts). Another important trait of Internet-based content is the ubiquity of advertising and the contiguity and embeddedness of advertising within website content. Finally, I considered the role of a theory of design within the studies. This is particularly important because distortion can result from confusing content with the constraints of a given medium. If a researcher does not consider the affordances and constraints of a medium, one can mistake noise for signal.

Therefore, I examined website content analysis studies with regard to the following characteristics: methodological rigor, treatment of intertextuality, treatment of multimodality, treatment of multilinearity, and the treatment of advertising. Complete results of this analysis are

reported in Supplemental Files. What follows is a discussion of what represents best practice for each of these factors.

Methodological rigor. The studies examined were conducted for a wide range of purposes and some of the research design choices, which I characterize as lacking in rigor, may have been reasonable choices given the context and purpose of the research pursued. Nevertheless, the standards which I have applied are also reasonable and necessary to establish the confidence with which one can view truth claims from these studies. To characterize the rigor of these studies, I compared the following methodological choices: sampling, data collection, and analysis methods.

Sampling methods. For the quality of sampling methods, I rated the studies on two criteria. First, did the authors clearly identify a population of interest? Second, did the authors specify how the sample was chosen from the population of interest? For each of these questions, I also considered whether or not the authors articulated a rationale for their choices. I assigned one of four levels: 3 Yes, with rationale; 2 Yes, no rationale; 1 Unclear; 0 No. A complete listing of outcomes can be found in Supplemental Files.

These criteria were applied regardless of research paradigm. For example, consider two of the studies that were assigned a 3 for both criteria. In their analysis of corporate websites in China, Xi, Zhuang, Huang, She, and Zhang (2007) assigned random numbers to the 600+ firms on the Shanghai Stock Exchange, sorted the list by the random numbers, and chose the first 74 companies. (Sample size was determined by representation within industries and practicality of analysis.) The authors clearly specified a population of interest and a rationale for defining that population. They detailed a quantitative selection procedure justified by their study design. In another study, Opoku, Pitt, and Abratt (2007) examined the ways best-selling authors

communicate brand personality on their websites. The population of interest was clearly identified in the study. To select an appropriate sample for analysis, the authors examined best-seller lists identified in industry sources (New York Times and Publisher's Weekly) and chose ten author-websites common to both. Although these two studies differed greatly in their methods, both clearly specified a population of interest and justified the selection methods for building their sample.

On the other end of the spectrum, several studies lacked either a clear definition of the population of interest (Henry & Story, 2009; Villanueva-Alfonso, Luzón-Marco, & Ruiz-Madrid, 2008), a clear description of selection criteria (Saiki, 2010), or lacked both (Gow et al., 2012).

The selection of which content to include in a content analysis varies with the research paradigm. Quantitative analysis methods that are aimed at generalizing findings to a larger population should include random sampling from a clearly defined population. Qualitative analysis may focus on cases that will aid in addressing a particular concern, for example, extreme cases, typical cases, confirming or disconfirming cases. Regardless of research paradigm, sampling methods and rationale for using those methods should be systematic and explicit.

Data collection. A researcher conducting a content analysis of a website faces unique challenges for data collection. The specific content of any given website is subject to change at any time. By contrast, any content analysis of documents or other artifacts that exist in a static format (newspapers, novels, letters, television shows, commercials, etc.) at least has the advantage of a static corpus. Website content constitutes a moving target. Methods for addressing that fluidity vary greatly. Based on my analysis of the published studies, I have

organized methods of addressing variability among websites across three variables: depth of data collected, fluidity within a site, and fluidity between sites.

Within the published literature, *depth of data collected* can range from a study that collects all content (including all pages) from a given website to a study that collects content from only the homepage or a specific area of the website. Researchers in one study individually reviewed over 4,000 web pages on 96 web sites (Moore, 2006), which represented every page on the sites included. One common research practice is to begin with the home page and specify a level of depth away from the home page for inclusion. For example, one study included the home page and all pages on the website directly linked from that home page (Culp, Bell, & Cassady, 2010), i.e. one level down or one click away. Another specified four levels down from the home page (Opoku et al., 2007), although the researchers included only typographic text (see section on *Treatment of Multilinearity* below). I coded the included studies at four levels of depth: *3 All data*, *2 Home page plus number of links*, *1 Homepage only or subset*, and *0 Not specified*.

Fluidity within the site refers to a study design's ability to describe the content on a site notwithstanding the possibility of change over time. This problem is specific to websites. For example, the content of a traditionally-printed newspaper article does not change over time, which allows the researcher to return to the same dataset over the course of analysis, whereas the content of a web-based news article can and does change over time, sometimes substantially. Researchers must account for the fact that any part of a site may change at any time. The most complete solution is to use a piece of software called a website archive tool. A website archive tool creates a copy on the user's computer of all of the files that make up a website, including HTML documents, images, audio, video, and formatting information. In studies that use this technique (Giesbers, Verdonck-de Leeuw, van Zuuren, Kleverlaan, & van der Linden, 2010;

Wohn, 2011), researchers essentially freeze the website content at the point of data collection in a way that preserves its multimodal, multilinear, networked nature. Another technique used in some studies (Alvy & Calvert, 2008; Juarascio, Shoaib, & Timko, 2010; Moore, 2006; Opoku et al., 2007; Yu, King, & Yoon, 2010) involves copying the website content in some manner that preserves some information while obscuring other information. “Lossy” collection includes taking screenshots of webpages, printing webpages, copying and pasting text from websites into offline documents, and videotaping while a person navigates a website. In each case, some aspect of website content is not recorded and thus not available for later inspection. The method least resistant to fluidity within a website is “live coding,” simply viewing live web pages online while completing a coding sheet. Using this method, the researcher has no way of revisiting the source data after initial coding has taken place. Additionally, if multiple coders are being used, the researcher using live coding must ensure that all coders view each live site within the same period of time or face a serious threat to both reliability and internal validity. Researchers in one study, Moore (2006), engaged in live coding but increased reliability by documenting every webpage with a screenshot during the live coding. I assigned one of three levels to each content analyses: *2 Duplicated content in the original format*, *1 Documented content in another format*, and *0 Coding live sites or not specified*. For complete results, see Supplemental Files.

For content analyses that included at least two websites, *fluidity between sites* refers to the resistance of the research design to changes that may occur between websites over time. For example, in a hypothetical study that intended to compare the treatment of a given topic on several different websites in which data collection occurred serially over an extended period of time (e.g. site 1 is observed in January and site 10 is observed in June), differences observed may be attributable to changes that have taken place over time. The most obvious and complete

solution to this problem employed by studies (Weber, Story, & Harnack, 2006; Wohn, 2011; Yu et al., 2010) is to collect data from all sites within the same brief period of time, between one day and one week. A researcher using live coding of data almost certainly cannot complete data collection within an abbreviated period. Several included studies identified longer data collection periods (Alvy & Calvert, 2008; Culp et al., 2010; Gerodimos, 2008; Shin & Huh, 2009).

Notably, Moore (2006) identified the exact data collection date for each website included in her content analysis. Although Moore's data collection took place over approximately six months, reporting these dates allows the reader to apply greater scrutiny to her findings.

Analysis methods. The highest standards in analysis involved an instrument that was designed for the study in question or from previous studies, piloted, and revised based on pretesting (Giesbers et al., 2010; Kelly, Bochynska, Kornman, & Chapman, 2008; Moore, 2006; Weber et al., 2006). Giesbers et al. (2010) is a particularly good example. The authors began with a well-established instrument from prior research, adapted it for the current research purposes, pilot tested the adapted survey, and then made revisions based on their pilot testing. Studies that adapted an instrument from previous studies without reporting pilot testing (Culp et al., 2010; Gerodimos, 2008; Haneefa & Nellikka, 2010; Juarascio et al., 2010; Tsiotsou, Rigopoulou, & Kehagias, 2010; Xi et al., 2007) were judged to be more rigorous than studies that developed an original instrument and used it without piloting (Caballero-Luque, Aragonés-Beltrán, García-Melón, & Dema-Pérez, 2010; Shin & Huh, 2009; Yu et al., 2010). In some cases, authors did not describe their instruments at all (Gow et al., 2012; Opoku et al., 2007; Saiki, 2010; Villanueva-Alfonso et al., 2008; Wohn, 2011).

Regarding the number of raters/coders and the rigor of the methods reported, a great range exists. The highest standard is reported in studies such as Moore (2006) and Alvy and

Calvert (2008). Moore used a total of five judges who were trained and randomly assigned to sites. Almost every site (92%) was evaluated by two coders. (Moore reports that 8% of her sample changed before a second coder could visit it, which highlights the importance of using a web archive tool in data collection, as discussed in *Data Collection* above.) The training of coders and the calculation of intercoder reliability are also important to establishing the validity and reliability of Moore's analysis. Alvy and Calvert reported that their three raters received 30 hours of training prior to the study. After calculating interrater reliability, all sites were coded a second time. Other studies that reported the number of coders, training of coders, and intercoder reliability were Yu, King, and Yoon (2010) and Shin and Huh (2009).

This level of rigor contrasts with website content analysis studies that did not report intercoder/interrater reliability (Giesbers et al., 2010; Gow et al., 2012; Xi et al., 2007), did not report training of the raters/coders (Culp et al., 2010; Giesbers et al., 2010; Gow et al., 2012; Tsiotsou et al., 2010; Weber et al., 2006; Wohn, 2011), or did not report any of these elements (Haneefa & Nellikka, 2010; Henry & Story, 2009; Juarascio et al., 2010; Opoku et al., 2007; Saiki, 2010; Villanueva-Alfonso et al., 2008).

Finally, with regard to the information reported from the analysis, some studies provided statistical analysis (Kelly et al., 2008; Moore, 2006; Wohn, 2011; Yu et al., 2010) and some provided simple counts of presence/absence of elements (Haneefa & Nellikka, 2010; Henry & Story, 2009; Tsiotsou et al., 2010). An analysis technique that seems very informative (and one not included in most website content analyses) was a holistic, narrative impression of the website. For example, in Gerodimos (2008), after applying the coding instrument to each site, each coder wrote a detailed case description to capture the overall experience and to locate problems, glitches, and anything else not covered by the coding sheet. Villanueva et al. (2008)

provided a detailed description of the visual appearance of each website. Caballero-Luque et al. (2010) is a rigorous evaluation of a single company website to establish the consistency of website content with the company goals which includes a narrative description of many aspects of the website.

No single website content analysis I observed included all of the best practices that were included in the overall sample. For example, Gerodimos (2008) employed a rigorous and thorough explanation of criteria used to identify and select websites for inclusion, but could have improved reliability by creating offline content archives rather than or in addition to live-coding websites. Henry and Story (2009) archived website content for offline analysis, but could have improved their design by reporting the number of raters, their training, and inter-rater reliability. Alvy and Calvert (2008) trained their coders and tested inter-coder reliability, but could have improved their design by shortening the six month period of website data collection or otherwise accounting for fluidity of content between websites. The methods of analysis planned for my study are informed by the methodological choices of the studies analyzed here.

Treatment of intertextuality. Intertextuality is the degree to which a text implicitly or explicitly refers to other texts and/or derives meaning from relationships with other texts (Bazerman, 1988; Lankshear & Knobel, 2007; Leu et al., 2009). Intertextuality exists in all written communication (Bakhtin, Holquist, & Emerson, 1986; Hartman, 1992; Kristeva, 1984). Intertextuality is pushed to the fore in hypertext because, in addition to rhetorical or poetic references to other texts, direct navigation links can be embedded in a text. In addition, a vast number of different texts can be accessed instantaneously on the Internet. This is an aspect that has not received enough attention. Authors of all kinds of texts can now rely on interested readers to locate the support needed to comprehend even an oblique and passing reference to an

obscure text. Using an Internet search engine and any one of a growing series of online compendia, readers can uncover layers of meaning within a text. An element as central to hypertext environments as intertextuality should be considered within any study that focuses on websites.

However, within the body of website content analyses reviewed for this study, most did not explicitly discuss intertextuality (Alvy & Calvert, 2008; Culp et al., 2010; Giesbers et al., 2010; Gow et al., 2012; Haneefa & Nellikka, 2010; Henry & Story, 2009; Juarascio et al., 2010; Opoku et al., 2007; Saiki, 2010; Weber et al., 2006; Wohn, 2011). Some studies counted hyperlinks to other websites (Kelly et al., 2008; Tsiotsou et al., 2010), although they did not describe the rhetorical implications of these links. Shin and Huh (2009) counted links to other websites, but only to determine the degree of similarity or difference between pairs of U.S. and Korean corporate websites. Moore (2006) discussed the presence of marketing partnerships, including media tie-ins, promotions, and sponsorships in her examination of advertising aimed at children, although intertextuality is not explicitly discussed. Villanueva et al. (2008) discusses “digital genres,” but does not specifically mention intertextual connections. In the evaluation conducted by Caballero-Luque et al. (2010), the evaluation client defined conceptual links between pages on the client website and those relationships informed the evaluators’ assessment of consistency on the website.

Perhaps the absence of a thorough discussion of intertextuality in website content analyses can be explained by the authors’ understanding of the purpose and role of content analysis as a method. Gerodimos (2008) states that content analysis is not meant “to establish the intentions of a message’s producers or its effects on the users,” (p. 972) but instead to provide a comprehensive assessment of the *capacity* of a website to affect readers. Gerodimos goes on to

say that in order to determine either the actual intentions of the website authors or the effects on website readers, authors or readers should be studied directly. This point is well taken. However, I believe that describing intertextual dimensions of website elements is necessary to understanding rhetorical capacities and that such description stops well short of trying to determine author intentions. In fact, such documentation of intertextuality on websites can inform and support future studies of authors and readers in online spaces. For this reason, my study design includes an assessment of intertextuality represented within website content.

Treatment of multimodality. Multimodality is the degree to which the text involves multiple semiotic modes (Bezemer & Kress, 2008; Kress, 2009). Websites typically include content in many modes, including text, images, audio, and video. A content analysis of any multimodal text that does not address all modes is limited by definition. For example, the photograph that an editor chooses to include with a newspaper story may subtly or drastically affect the meaning of the text it accompanies. An analysis that looks only at text and ignores images may be leaving vital meaning-making material unexamined.

While most studies that addressed multimodal content simply counted the number of images, videos, etc., within a site, some studies primarily considered the purpose or meaning of multimodal content. For example, Alvy and Calvert (2008) treated images and other multimodal content the same way that they treated text, first categorizing all elements on a site based on their advertising purpose, then categorizing them based on the mode. In their study of marketing techniques, Culp et al. (2010) focused primarily on the strategies employed in their study of marketing. They identified the frequency of elements meant to encourage exposure, branding strategies, and healthy messages. Similarly, Xi et al. (2007) first looked at meaning and purpose of each element. The researchers then identified its mode (i.e., text, graphics, audio, and video).

Moore (2006) described the prevalence of images, specifically brand identifiers, and games. Music and colors were also mentioned in relation to customizable game spaces. The use of other modes to communicate is implied in Moore's analysis of brand identifiers that are images, although this is far from an in-depth analysis of multimodality. Finally, some studies explicitly limited themselves to an examination of typographic text (Giesbers et al., 2010; Juarascio et al., 2010; Opoku et al., 2007). Although limiting analysis to typographic text may distort the overall interpretation of a website, making the limitation explicit at least makes the analysis methods clear to the reader. In summary, although communication in multiple modes is an important fact of website content, it is rarely studied in website content analyses. Those studies that analyze non-typographic text for meaning and rhetorical purpose are more likely to accurately represent those multimodal texts. My study design includes extensive analysis of each element, inclusive of all modes.

Treatment of multilinearity. Multilinearity can be defined as the degree to which a text supports multiple reading paths (Dresang, 1999; Portela, 2007; Tyrkkö, 2007). Hypertext documents are typically described as multilinear, in that the reader can determine his or her own reading path through a document. Hypertext documents are sometimes referred to as nonlinear (Culp et al., 2010), however I would argue that the term multilinear is a much better description. Multilinearity is not new. Referring to traditional printed text, Hartman (1992) goes as far as to suggest that all texts are constructed by the reader as he or she chooses among possible reading paths. This suggests that the difference between hypertext and traditional printed text with regards to linearity is a difference of degree rather than a difference of type. The question is the degree to which an author attempts to restrict a reader's choice of path, or restated, the degree to which a single path is determined for the reader, or restated yet again, the degree to which the

text supports multiple paths. In many traditional print genres, multilinearity is expected. An encyclopedia, for example, is not intended to be read in a linear fashion from beginning to end. Other examples of multilinear printed texts include newspapers, magazines, and information picture books, such as *Castle* by David MacAulay. For contrast, a novel is an example of a highly linear text, whether read in electronic or paper format. Because websites are typically meant to support multiple reading paths, and in fact the writer is able to suggest multiple reading paths through the use of hyperlinks, site maps, and increasingly through the use of suggestions based on prior browsing behavior, website content analyses should account for multilinearity.

Unfortunately, most website content analyses do not address multilinearity. The notable exception among the included studies is Villanueva et al. (2008) which identified hypertext as “one of the most important affordances” (n.p.) of digital genres. The authors analyzed the structure and design of the websites in question. They documented the hyperlinked structure using a concept, or network, map. This important step allowed them to consider the possible reading paths supported by the author. In their study of food advertising, Henry and Story (2009) documented all internal links going to the designated children’s area (DCA) on the websites they studied. Their method, which was an extension and improvement of the methods of Weber et al. (2006), noted the presence or absence of certain types of internal links. In their description of website ease-of-use, Xi, Zhuang, Huang, She, and Zhang (2007) approached a description of the multilinear affordances of websites. Moore’s (2006) references to the customizability of website content also can be considered consistent with multilinearity. Tsiotsou et al. (2010) used a combination of multi-dimensional scaling (MDS), concept mapping, and hierarchical cluster analysis (HCA) in their study of marketing techniques on websites. This technique presents an interesting empirical mapping of related goals and orientations in website content and would be

an interesting complement to a study of website designers. Based on my analysis of the published literature, my methodology includes the documentation of website structure through the use of concept maps.

Treatment of advertising. On the Internet, advertising is ubiquitous. Frequently, the boundaries between advertising and non-advertising content are blurred to the extent that no boundaries exist, and strong motivation exists for website creators to include compelling advertising content. In their investigation of “adver-gaming,” Culp et al. (2010) stated that the Internet may have “more potential than television to influence children’s purchasing habits” (p. 199). In general, studies that specifically and exclusively focused on advertising were the only studies that included any kind of treatment of advertising. Most website content analyses did not consider advertising as a factor. Among those studies that did consider advertising, it sometimes receives a shallow treatment.

In their study of food brand websites, Henry and Story (2009) listed the percentages of sites that included each of 35 “marketing techniques.” These techniques included elements such as “Incentives” which might be construed to describe some type of rhetorical persuasion, but not so with other elements such as “Safeguards”, “Downloads”, or “Web site features.” This study would have benefitted from employing a set of categories of rhetorical persuasion or marketing and fitting the listed characteristics within the categories. Weber et al. (2006) has similar limitations. Tsotsou et al. (2010) specifically studied the marketing capacities of sixty-four retail websites. Their extensive analysis revealed the need of companies to align their websites to company goals, a recommendation shared by Caballero-Luque et al. (2010). Both of these studies were oriented toward improving outcomes for the companies that owned the websites, and understandably neither took the critical approach to online advertising content adopted by

proponents of critical literacy (Luke, 1996). A critical approach to advertising can be found in other included studies (Alvy & Calvert, 2008; Kelly et al., 2008; Moore, 2006), although none categorize or otherwise describe the content of advertising appeals. Based on my reading of the literature, my methods include analysis of advertising content without separating it from other website content.

Conclusion

This review of the literature encompassed three distinct bodies of inquiry. First, a historical survey of rhetoric and literature helps situate the concept of genre. Genre exists as a tension between tradition and revolt against tradition. It is an essential concept to understanding changes in literacy and literate practices as different modal genre traditions collide in new literacy environments, such as children's websites.

Second, a survey of published research that addresses reading in online environments revealed a dichotomy between reader-focused and writer-focused studies of online reading. In both cases, inadequate attention is paid to understanding the features of online spaces in which unstructured reading activities take place. Primary focus has been on studying informational reading for academic purposes, which is an admittedly important goal. However, this research would be greatly improved with a better and more systematic understanding of the nature of the online texts with which children choose to engage.

Finally, this review focused on content analysis methods as applied to websites. Content analyses of websites have little consistency in methodological rigor. Good examples can be found of selection, data collection, and data analysis methods, but seldom in the same study. In addition, it is all too rare for a website content analysis to adequately account for intertextuality, multimodality, multilinearity, and advertising. Each of these elements have been identified as

critical to understanding online texts in previous theory, but each has been regularly overlooked in content analysis of websites.

Each of these areas of review leads directly to the methods described in the next chapter. The current study will provide a comprehensive method of website content analysis from a literacy perspective with a particular emphasis on the role of genre. Studies of online reading will benefit from a better understanding of the literacy dimensions typically encountered by children reading online. Content analyses will benefit from a comprehensive approach that considers traditional literacy elements in addition to web-related factors like multilinearity, multimodality, and intertextuality.

Eventually, the exploration of children's website genres should be informed by and should inform genre theory work involving other multimodal media. Film theory has a rich history addressing the interaction of words and images within various genres (Altman, 1999; Bordwell, 2008; 1989; Langford, 2005). Indeed, both public and scholarly film genre definitions have evolved over decades of analysis, and as Bordwell (1989) notes, "Genres, and genre, function as open-ended and corrigible schemata" (p. 148). The interpretation of picture books provides another rich example of scholarship in a multimodal genre that may inform the examination of children's websites (Sipe, 1998; Nikolajeva & Scott, 2006; Bang, 2000).

However, as Kress (2003) asserts, "...we cannot transport mode-specific theories from one mode to another without producing severe distortions... meanings, in the broad sense, can be realized in any mode, but that when they are, they are realized in mode-specific articulations." The relationship between images and text in picture books is synergistic (Sipe, 1998) in ways that may be similar to the images and text on children's websites, but those similarities should be

examined in light of independent examinations from each area of inquiry. The current study is an effort to develop a method for documenting the multifaceted nature of children's websites.

Chapter Three: Methods

That children are engaged with the Internet is well known. What is less known is the nature of the literacy landscape in which they choose to spend time. Although many studies have examined the reading children do online in the context of academic reading, little is known about the texts with which they choose to engage on the Internet. Moreover, many would not consider the websites children choose to visit to be texts at all. Nonetheless, children read and write in online environments that they choose to navigate, within what are arguably the new texts of literacy.

Purpose

The purpose of this study was to establish a comprehensive methodology for analyzing children's website content, based on both linguistic and rhetorical data, by employing defensible criteria to evaluate both qualitative and quantitative data. By applying genre theory through a mixed methods content analysis, I addressed gaps between research into online genres and research into children's reading on the Internet. This study was developed around the supposition that understanding the texts that children choose to read online, not exclusively informational or academic texts, is necessary in order to fully understand reading in online environments designed for children.

Overview

I selected a sample of five children's websites and conducted a content analysis (Krippendorff, 2004) focused on four kinds of variables for each website: indexical, substantive, rhetorical, and formal. Genre can be defined by substance, rhetorical action, and form

(Bazerman, 1988; Devitt, 1996; Miller, 1984). My data categories corresponded to these three with the addition of indexical data, which identified the page or element and was not meant to be interpretive. Substantive data referred to the content or subject matter of the page. Rhetorical data described the rhetorical action being accomplished or contributed to with the element, as defined by Bazerman (1988) to include lexicon, citation, author, and audience. Formal data referred to measures of formality, complexity, and coherence. Following a relatively broad page-level analysis of five websites, I proceeded to a more in-depth element-level analysis.

Prior to selection of the websites for evaluation, I developed two data collection and analysis instruments, the *Web Page Data Evaluation Checklist* and the *Element Data Evaluation Checklist*. I established the validity and reliability of the instruments using 50 purposefully sampled web pages from children's websites. Therefore, the study comprised three distinct phases: phase one, instrument development; phase two, page level analysis of five websites; phase three, element level analysis of one or more websites.

In this study, I proceeded from an understanding of genre as socially-constructed, flexible, situated, and dynamic discourse forms (Chapman, 1995; Devitt, 1996; Jamieson, 1975; Kamberelis & Bovino, 1999; Miller, 1984). Although formal, structural elements are an essential part of defining a genre, a genre is most significantly defined as the site of social action (Miller, 1984). Genres are deployed by people to take action with other people and genres change constantly in response to the needs of the people who use them. Still, as rhetorical exigencies recur, rhetorical solutions to those situations also recur. To the extent that those individual rhetorical solutions resemble one another in substance, rhetorical action, and form, genres exist. As Schryer expressed it, genres are "stabilized-for-now or stabilized-enough sites of social and

rhetorical action” (1993, p. 204). Genre is significant in this context because genres are essential to understanding the shifting literacy landscape of online texts.

Research Questions

1. What methodology allows for the comprehensive study of genre of children’s websites?
 - a. What instruments can facilitate the systematic evaluation of children’s websites?
 - b. What methods of analysis can be applied to understand the genre of children’s websites?
2. What is the nature of the formal elements of selected children’s websites?
 - a. What are the formal textual elements associated with selected children’s websites?
 - b. What are the formal visual elements associated with selected children’s websites?
 - c. What text-image relationships are found within selected children’s websites?
3. What is the nature of the substance of selected children’s websites?
 - a. What are the organizational and thematic units of within selected children’s websites?
 - b. What are the topics addressed by selected children’s websites?
 - c. How are the author and reader positioned by the text?
4. What is the nature of the rhetorical action of selected children’s websites?
 - a. What stated or implied communicative purposes are served by selected children’s websites?
 - b. Within what context or community do selected children’s websites operate?
 - c. Who is the intended or purported audience of selected children’s websites?

d. What attitudes and prior knowledge does the author presume the reader to have?

e. How is authorship of selected children's websites identified?

f. How are implicit and explicit external references used to position the website by the author?

g. What evidence points toward antecedent genres that fulfilled the rhetorical purposes of the genres represented within selected children's websites prior to the existence of children's websites?

5. What evidence supports or refutes the identification of a genre or genres within the body of children's websites examined within this study?

Research Design

Rationale. The research questions outlined above require the researcher to make inferences based on the content of children's websites. Consequently, these questions reflect the need for a complex combination of qualitative and quantitative data to form a replicable basis for those inferences. Prior research seeking to establish genres in online contexts has relied either on a rhetorical, interpretive model (Herring et al., 2005; Miller & Shepherd, 2004; Puschmann, 2009) or on a quantitative, linguistic model (Dillon & Gushrowski, 2000; Emigh & Herring, 2005) to establish and define an online genre. I employed a mixed methods content analysis design to take advantage of techniques from both the rhetorical and linguistic traditions of genre analysis.

Uses of content analysis. Content analysis is defined by Krippendorff as, "a research technique for making replicable and valid inferences from texts (or other meaningful matter) to the contexts of their use" (2004, p. 18). As a methodological framework, content analysis allows for the incorporation of both quantitative and qualitative methods. The choice of specific

methodologies is governed by the research questions and contexts in which those questions are asked (Busch et al., 2012; de Sola Pool, 1959; Krippendorf, 2004).

The roots of content analysis can be traced through theological studies of religious and secular texts as far back as 17th century Europe. In the early 20th century, the earliest content analyses conducted in English were adopted as an objective method of analyzing newspaper content within the new field of journalism studies (Krippendorf, 2004). Content analysis grew during the 1930s and 1940s in response to four factors: the American public's growing skepticism about mass media, growing recognition of the fundamental differences between newspapers and new media (radio and television), acknowledgement of the influence of radio in the rise of fascism, and increasing public acceptance of behavioral and social sciences (Krippendorf, 2004). During World War II, Allied intelligence analysts used content analysis to interpret internal Nazi propaganda (George, 2009). Although prior content analysis focused primarily on word counts, by the mid-1950s, content analysis was expanded to examine the relationships of concepts within a text rather than simply the presence of words (de Sola Pool, 1959). Today, content analysis is used in marketing and media studies, literary analysis, rhetorical analysis, ethnography, gender studies, sociology, political science, psychology, and cognitive science (Busch et al., 2012) because it provides a rational basis for making inferences based on many different kinds of content and allows for the inclusion of many different forms of evidence.

Strengths of content analysis. The strengths of content analysis lie within the researcher's ability to directly examine the texts or transcripts of communication within specific contexts. It is a flexible framework that allows for the examination of many different types of data and the employment of many different quantitative and qualitative methodologies, including

the statistical analysis of coded text. It is an unobtrusive means of analyzing communication, and it can illuminate complex patterns of interaction (Busch et al., 2012).

Limitations of content analysis. Content analysis is limited by its status as a descriptive methodology. Although content analysis can be used to document the existence of certain rhetorical and linguistic structures, it may be insufficient for explaining how and why those structures exist. Further, content analysis is based on the interpretation of the analyst, acknowledging that other equally valid interpretations also exist.

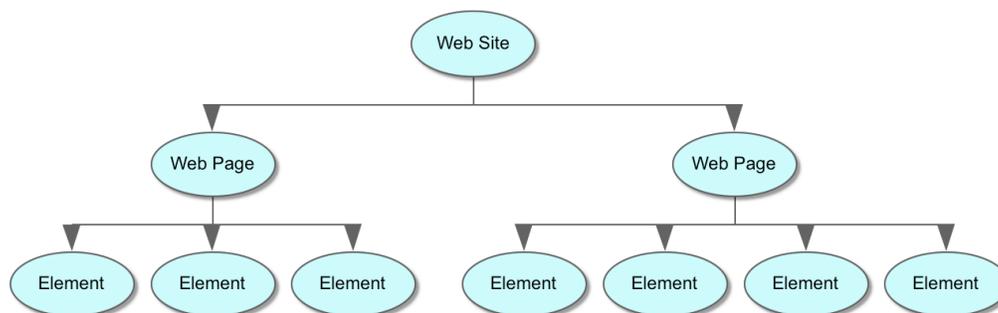


FIGURE 3.1. Relationships between elements, pages, and sites. A website is a combination of web pages. Web pages are a combination of elements.

Operationalization of website content. For the purposes of this analysis, I defined an *element* as a component of a web page that may convey meaning or contribute to conveyance of meaning. An element may be a text passage, a text label, an illustration, a photograph, a video, a map, a diagram, an internal link, an external link, or a cluster of other elements. A web page may include any combination of elements and a website may contain any combination of web pages (see figure 3.1).

For example, a section of the children’s website nick.com comprised 116 web pages. One of those pages was the biography page for character Kendall on the show “Big Time Rush,” shown in figure 3.2. This page contained approximately 100 separate elements, including a text

passage element (the two paragraph biography), multiple text label elements (picture captions, navigation buttons), several photograph elements (pictures of each band member, pictures of Kendall), multiple illustration elements (within the content of advertisements and navigation), internal links (such as “HOME,” “GAMES,” “VIDEOS,” “TOUR MAP,”), external links (advertising content), and a number of clusters (each navigation bar functions as a cluster of text, link, and image elements, for example). From this example, it is evident that text will be considered in several different ways as it contributes to the communicative purposes of the page.

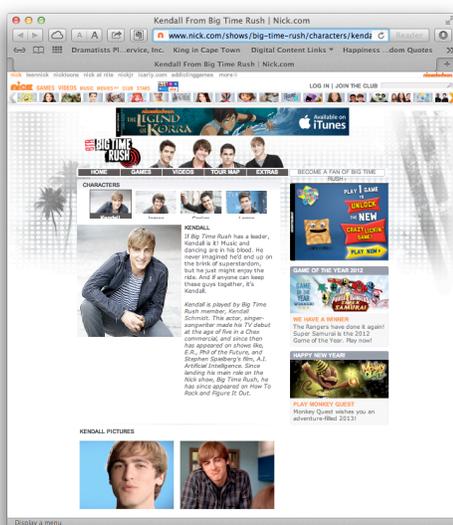


FIGURE 3.2. An example of a web page consisting of multiple elements. Reprinted from “Kendall from Big Time Rush”. (2014). Nickelodeon. Retrieved from <http://www.nick.com/big-time-rush/kendall/>.

Phase 1: Instrument Development

During this phase, I developed (1) a data evaluation checklist to be applied to *web pages* and (2) a data evaluation checklist to be applied to *web page elements*. Each checklist included a series of variables to facilitate consistent description of each element or page.

Web Page Data Evaluation Checklist. I evaluated four kinds of variables for each web page: indexical, substantive, rhetorical, and formal. Genre can be defined by substance,

rhetorical action, and form (Bazerman, 1988; Devitt, 1996; Miller, 1984). My data categories correspond to these three with the addition of indexical data, which identifies the page and is not meant to be interpretive. Substantive data refers to the content or subject matter of the page. Rhetorical data describes the rhetorical action being accomplished or contributed to with the element, as defined by Bazerman (1988) to include lexicon, citation, author, and audience. Formal data refers to measures of formality, complexity, and coherence. Based on my understanding of genre, I included the following variables on the data evaluation checklist for web pages. The category of data appears parenthetically following the variable name.

Indexical variables. The following variables were used for identification, organization, and general description and were generally not intended to be interpretive.

Web page ID (Indexical). This was a unique numeric identifier for each web page.

Creation date (Indexical). This identified the date the data record was created. This did not refer to the creation date for the web page itself.

Modification date (Indexical). This identified the last date the record was modified.

Page title (Indexical). The title for the web page was a named element within the code of the page.

Page address (Indexical). The uniform resource locator, or URL, for the page on which the element appears was the web address for the page, e.g. <http://news.google.com/>.

Substantive variables. The following variables addressed the subject matter or content of the web pages.

Subjects present (Substantive). This variable referred to the semantic content of the web page. Values were generated inductively using the initial sample of 50 web pages and validated

with subject matter experts prior to beginning data analysis. The values for this variable were highly idiosyncratic since they are derived from the particular pages under examination.

Primary subject (Substantive). This variable referred to the primary semantic content of the web page, selected from the list of *Subjects Present*. Selection of a primary subject from the subjects present included consideration of frequency of occurrence, proportion of semantic content represented, rhetorical emphasis, and other indicators of relative importance.

Rhetorical variables. The following variables addressed the rhetorical action or function of the web pages. A more detailed discussion of the rhetorical framework appears below, in the description of the *Element Data Evaluation Checklist*.

Functions present (Rhetorical). This variable referred to the rhetorical function of the web page. The values were generated inductively using an initial sample of 50 web pages and validated with subject matter experts prior to beginning data analysis. Values included *description, narration, exposition, argumentation, and instruction*.

Primary function (Rhetorical). This variable referred to the primary rhetorical function of the web page, selected from the list of *Functions Present*. Selection of a primary function from the functions present included consideration of both frequency and importance.

Formal variables. The following variables described formal aspects of the page. A more detailed discussion of form appears below, in the description of the *Element Data Evaluation Checklist*.

Element types present (Formal). This variable referred to the element types present on the web page. The list of possible values for this variable were generated inductively using an initial sample of 50 web pages and validated with subject matter experts prior to beginning data

analysis. Values for this variable included text passage, text label, illustration, photograph, video, map, diagram, internal link, external link, or cluster of other elements.

Primary element type (Formal). This variable referred to the primary element type on the web page, selected from the list of *Element Types Present*. Selection of a primary element type from the element types present included consideration of frequency of occurrence, proportion of semantic content represented, rhetorical emphasis, and other indicators of relative importance.

Indexical	Substantive	Rhetorical	Formal
<i>These variables allow basic identification and description of the elements.</i>	<i>These variables describe the substance or content of the elements.</i>	<i>These variables describe the rhetorical action of the elements in terms of lexicon, citation, audience,</i>	<i>These variables depict the formal dimension of the elements.</i>
Element ID Date Created Data Edited Page Title Page Address Element Type Element Description Cluster ID	Organization Subject(s) Theme(s) Author Position Audience Position	<i>Citation:</i> Explicit External References Relation of EER to Central Claims EER Characterization Implicit External References Relation of IER to Central Claims IER Characterization <i>Audience</i> Stated or implied audience Audience Prior Knowledge Audience Attitudes Persuasion Call to Action <i>Author</i> Stated or implied Author Experience Claims Broadness Originality Idiosyncrasies	Word Count Sentence Count Paragraph Count Syllables per Word Words per Sentence Sentences per Paragraph Flesch Reading Ease Flesch-Kincaid Grade Level Concreteness Noun Hypernymy Verb Hypernymy Personal Pronouns Displayed Image Size Relative Image Size Dominant Color(s) Image Type Image Relation to Text

FIGURE 3.3. Overview of variables from *Element Data Evaluation Checklist*.

Element Data Evaluation Checklist. Four kinds of variables were evaluated for each element: indexical, substantive, rhetorical, and formal. Genre can be defined by the substance,

rhetorical action, and form (Bazerman, 1988; Devitt, 1996; Miller, 1984). My data categories corresponded to these three with the addition of indexical data, which identifies the page and is not meant to be interpretive. Substantive data referred to the content or subject matter of the page. Rhetorical data described the rhetorical action being accomplished or contributed to with the element, as defined by Bazerman (1988) to include lexicon, citation, author, and audience. Formal data referred to measures of formality, complexity, and coherence. Based on my understanding of genre, I included the variables described in the following subsections on the *Element Data Evaluation Checklist*, after validation with subject matter experts.

Indexical variables. The following variables were not meant to be interpretive and were used to identify and organize the elements.

Element ID (Indexical). This was a unique numeric identifier for each element.

Creation date (Indexical). This identified the date the record for this element was created.

Modification date (Indexical). This identified the last date the record was modified.

Page title (Indexical). The page title for the page on which the element appeared was a named element within the code of the page on which the element appears and was not interpretive.

Page address (Indexical). The uniform resource locator, or URL, for the page on which the element appeared was the web address for the page, e.g. <http://news.google.com/>. The page address was not interpretive.

Element type (Indexical). The element type was selected from a controlled list of element types validated by subject matter experts. Element types included text passage, text label, illustration, photograph, video, map, diagram, internal link, external link, or cluster of other elements.

Multimodal transcription of video content. In addition to the standard variables analyzed for all elements, video elements received special attention. I completed a multimodal transcription (Baldry & Thibault, 2006) for selected elements. Multimodal transcription allows the researcher to describe changes in multiple elements in a video in parallel and indexed by time. An example of multimodal transcription of video content is included as figure 3.4. Transcription began with a straightforward accounting of each second of the video. Column one referred to seconds, the visual frame was described in columns two and three, movement accounted for in column four, and audio described in column five. Column six was used to describe the phases and subphases of the video from a functional perspective. I used multimodal transcription to unpack meaning from video elements, which were then examined using the full set of variables.

T	Visual Frame	Visual Image	Kinetic Action	Soundtrack	Metafunctional Interpretation, Phases and Subphases
:09		↓	"GALACTIC HOT DOGS" logo continues slowly growing in size approaching the edge of the frame.	Up tempo music	Movement, image, and music reinforce dynamism.
:10		Empty gray background. Black and white image of Cosmo with text "THAT'S ME". Motion lines around image.	Cut to blank gray background. Cosmo moves in from left. Accent lines appear and disappear. Text fades and scales in.	↓	PHASE 2: "THAT" has no other referent than the image. The character is the implied author/narrator, otherwise the text would read "THAT'S COSMOE." Cosmo is speaking directly to the viewer.
:11		B/W image of Cosmo with text "THAT'S ME". Motion lines around image. Curved arrow from text to image. "COSMOE THE EARTH-BOY"	Arrow draws in. Second text appears character by character. Cosmo slowly drifts right.	↓	Animation that depicts the drawing of the arrow and the typing of the text reinforces the idea that "Cosmoe" is speaking informally to the viewer.
:12		↓	↓	↓	↓
:13		Text: "THE KID WITH THE FUNKY HAIR."	Cut to blank gray background. Text appears character by character.	Guitar riff	Animation that depicts text typing in reinforces "Cosmoe" as the narrator. Informal language "kid" and "funky" implies comical tone.
:14		↓	↓	As guitar riff concludes, moves back into up tempo music bed, continuing under successive effects.	Guitar riff punctuates this text, establishes a rhythm similar to a setup and punchline.
:15		Close-up image of Cosmo. Text "FUNKY HAIR". Arrow pointing from text to image.	Image of Cosmo moves in from right to left in front of the text. The text disappears as he wipes across. "FUNKY HAIR" appears when the image of Cosmo stops moving. Arrow appears and flashes on and off, pointing from the text to the image of Cosmo's hair.	Woosh sound effect as the image of Cosmo moves in. "Boing" sound effect plays each time the arrow appears.	Comical tone reinforced. "Cosmoe" as our narrator owns the "funky" label, embracing uniqueness.

FIGURE 3.4. Example of multimodal transcription of video. Thumbnail images from "Galactic Hot Dogs: Meet Cosmoe & the Crew!". (2014). Pearson. Retrieved from <http://www.funbrain.com/galactichotdogs/index.html>.

Element description (Indexical). This was a brief text description of the element. As much as possible, I avoided interpretive comments here. The purpose of this variable was to allow the researcher to refer to individual elements during interpretation. Examples of element

descriptions included “biography of Kendall from Big Time Rush” and “photograph of Kendall from Big Time Rush” (see figure 3.5). Criteria for writing element descriptions were specified in the data coding manual during the validation and reliability phases.

Cluster ID (Indexical). The Cluster ID was a unique number that allowed for the grouping of elements into clusters. Cluster analysis (Baldry & Thibault, 2006) involved identifying several elements that work together to create meaning in addition to any individual meaning. An example of a group of elements that included clusters is shown in figure 3.5, a detail from the “Big Time Rush” band biography page from nick.com (<http://www.nick.com/shows/big-time-rush/characters/kendall.html>). The individual elements depicted here included four photographs, five text labels, and one background color box illustration. Each photo and its corresponding text label (e.g., photo of Kendall + text label “Kendall”) functioned as a cluster element, providing an additional four elements. The dark box element around the “Kendall” cluster indicated to the audience which biography was displayed or selected. All four text/image clusters combined with the text label “CHARACTERS” and the dark box element to create a navigation cluster for the adjacent section of the page.

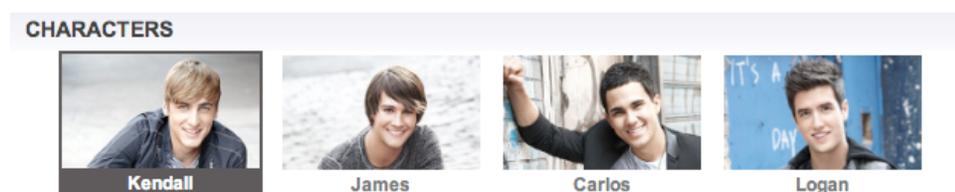


FIGURE 3.5. Detail from bio page featuring several elements. Reprinted from “Big Time Rush”. Nickelodeon. (2014). Retrieved from <http://www.nick.com/big-time-rush/#about>.

Table 3.1 displays values for three variables from the Indexical section of the *Element Data Evaluation Checklist*. Although each individual element was evaluated against the entire range of substantive, rhetorical, and formal variables, each cluster element also was evaluated against the same variables. Thus, the photograph of Kendall would be evaluated individually

(element ID 1002), as part of the photo/text cluster (element ID 1006), and as part of the navigation cluster (element ID 1014). As with the interpretation of other elements, the

TABLE 3.1.
Selected Indexical variables for elements in figure 3.5.

Element ID	Element Type	Element Description
1001	Text label	“CHARACTERS” navigation title
1002	Photograph	Medium close-up photo of Kendall
1003	Photograph	Medium close-up photo of James
1004	Photograph	Medium close-up photo of Carlos
1005	Photograph	Medium close-up photo of Logan
1006	Text label	“Kendall” photo caption
1007	Text label	“James” photo caption
1008	Text label	“Carlos” photo caption
1009	Text label	“Logan” photo caption
1010	Cluster	Kendall photo + text label
1011	Cluster	James photo + text label
1012	Cluster	Carlos photo + text label
1013	Cluster	Logan photo + text label
1014	Illustration	Dark box around Kendall cluster
1015	Cluster	Navigation cluster including four image/text clusters, title, and dark box

identification of clusters was an interpretive act. I identified cluster elements by considering the ways in which individual elements were organized, arranged, ordered, systematized, harmonized, correlated, synchronized, brought together, matched, complemented, set off, and/or blended. The specific criteria for the establishment of clusters was explicated during the validation phase.

The process of cluster analysis (Baldry & Thibault, 2006) promotes a non-linear understanding of multimodal texts by allowing the analyst to examine organizational principles other than sequence and proximity, such as periodicity, the patterned repetition of structures or variation within a framework (Baldry & Thibault, 2006). During cluster analysis, the researcher identifies groups of multimodal items on the page that function in tandem to create meaning. For example, a series of images on a web page may work together to set an expectation and then subvert that expectation in an ironic way for comedic effect. In such a case, the irony does not reside in any single image but in the images functioning in tandem. This does not preclude the interpretation of component items independently from the cluster in other parts of the analysis. Instead, the cluster analysis complements the analysis of elements in isolation. For example, while a particular image on a webpage may carry a certain meaning on its own, it may communicate an additional message when considered as part of a cluster that includes adjacent text and other images. In fact, the meaning of a cluster may exist in tension with the meanings of some of its constituent elements. An example of cluster analysis method is shown in figure 3.6. Element 1017 in this example is the “Main Navigation Cluster.” Although each image, each text label, and each hyperlink is a separate element evaluated for independent rhetorical, substantive, and formal functions, clustering these elements allows the researcher to examine how they function in tandem (or in opposition) to create meaning. Another example from figure 3.6 is element 1021, a cluster that includes text labels from two different sections of the page. These

elements are not adjacent yet still function in tandem. Cluster analysis allows us to consider additional possible rhetorical functions by clustering non-adjacent items.

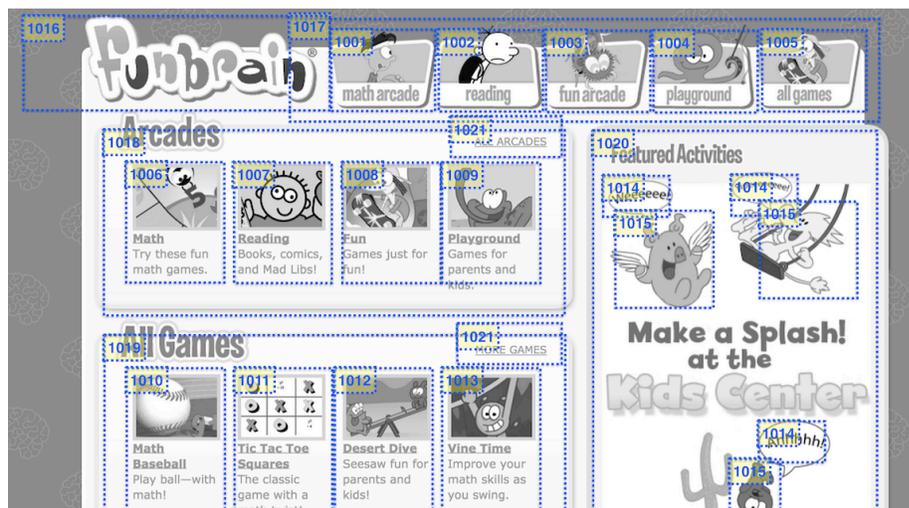


FIGURE 3.6. Example of cluster analysis. Source material reprinted from “FunBrain.com - The Internet’s #1 Education Site for K-8 Kids and Teachers”. (2014). Pearson. Retrieved from <http://www.funbrain.com/>.

Substantive variables. The following variables addressed the subject or content of the elements. All substantive variables were subject to validation with subject matter experts and reliability testing.

Organization (Substantive). This variable referred to the ways in which meaning was organized within an element. For example, within a text passage sub-headings may be used to organize meaning. The list of possible values for this variable was generated inductively using my initial sample of 50 elements and validated with subject matter experts prior to beginning data analysis.

Subject(s) (Substantive). This variable referred to the semantic content of the element. The list of possible values for this variable was generated inductively using my initial sample of 50 elements and validated with subject matter experts prior to beginning data analysis. Because of the great variety of possible subjects that can be addressed, this variable did not involve a

controlled list of values. During data validation and reliability testing, I specified the specific criteria for identification of subjects.

Theme(s) (Substantive). This variable referred to the unifying or dominant idea expressed using this element. Theme is a broader category than subject. The list of possible values for this variable was generated inductively using my initial sample of 50 elements and validated with subject matter experts prior to beginning data analysis. As with the subject variable, evaluation of this variable involved a specific process rather than a controlled list of possible values.

Author position (Substantive). I identified the rhetorical stance taken by the author through this element. The author position was an interpretive element that described the attitude, argument, tone, purpose, and/or opinion the author adopts through the use of this element (Bitzer, 1968). The list of possible values for this variable was generated inductively using my initial sample of 50 elements and validated with subject matter experts prior to beginning data analysis.

Audience position (Substantive). I identified the position of the audience indicated by the author through this element. The audience position was an interpretive element that described the ways in which the author characterized the audience and/or oriented the audience toward the subjects or themes through the use of this element (Bitzer, 1968). The list of possible values for this variable was generated inductively using my initial sample of 50 elements and validated with subject matter experts prior to beginning data analysis.

Rhetorical analysis. The evaluation of rhetorical variables was guided by Bazerman's (1988) method described in his study of disciplinary genres, which is itself a modification of Kinneavy's (1971) communication triangle. Kinneavy's triangle describes the rhetorical situation as the interdependent relationship between the writer, the audience, and the subject matter.

Bazerman's approach considers four interdependent contexts, or dimensions, in rhetorical analysis of a genre: the lexicon of the text, the nature of citation, treatment of audience, and authorial representation. Bazerman's model is appropriate for this study because it recognizes that texts participate in genres as a function of interaction between and amongst these different contexts. Bazerman's four contexts provide a systematic way in which to map and model these interactions. The evaluation checklists included codes for evidence of each of these four contexts.

The first context, lexicon, or "language and reality" (Bazerman, 1988, p. 26), refers to the type of information conveyed, how that information is framed, and the type of symbolic systems used to depict that information. Lexicon also considers the precision of language applied to the context. Lexicon defines the object or objects under consideration and the quality of connection between the text and the world. For each page under consideration, I determined the subject, the symbolic systems used to convey information about that subject, the ways in which the information is framed, and the precision of the language used. Because there was overlap between each of the four contexts, these issues were addressed within the variables for substantive, formal and other rhetorical contexts, therefore there are no variables specifically labeled *Rhetorical: Lexicon*.

Rhetorical: Citation variables. The following variables reflect a rhetorical analysis of Bazerman's notion of citation. Citation (or "language and tradition") referred both to the implicit knowledge and the explicit references the author makes to previous literature on the subject. For explicit references, Bazerman contends that the researcher must consider exactitude of meaning of the reference, the relation of the reference to the central claims of the text, the use and manner of discussion of the reference within the text. For example, citations can be used as symbols for

knowledge within texts under consideration. I coded elements for evidence Bazerman's concept of *citation*. I identified both explicit and implicit references the author made to other sources of information on the same topic, the way in which explicit and implicit references were related to central claims, and how references were characterized by the author. I operationalized these issues with the variables that follow. My operationalization was validated with subject matter experts.

Explicit external references (Rhetorical: Citation). The value of this variable was either 'yes' or 'no' depending on whether or not the element refers explicitly to a source outside of the website on which the element appears. A 'yes' value indicated that the element involved both an unequivocal reference to another idea and that the reference was to a source external to the website on which the element appears. Explicit external references included outbound hyperlinks to other websites. I specified criteria for identification of explicit external references during the validation and reliability-testing phase.

Relation of explicit external reference to central claims (Rhetorical: Citation). Using this variable, I identified the ways in which an external reference is related by the author to the subject(s), theme(s), author position, and/or audience position of element. Values for this element include agreement, opposition, evidentiary, and neutral. The list of possible values for this variable was generated inductively using my initial sample of 50 elements and validated with subject matter experts prior to beginning data analysis. Explicit external references included outbound hyperlinks to other websites. Outbound hyperlinks may tell us with which websites the author of a target website would like the website associated, or which websites the author sees as furthering or otherwise supporting its goals and purposes.

Characterization of explicit external reference (Rhetorical: Citation). This variable referred to the way in which the external reference was discussed or characterized by the author through the element. Both the *relation* variable and the *characterization* variable described the author's depiction of the external reference. However, the *relation* focused on the description relative to the current website and *characterization* focused on description without reference to the current site. The difference between these variables is analogous to the difference between relative and absolute positioning. Possible initial values included "authoritative," "affiliative," and "marginalizing." The list of possible values for this variable was generated inductively using my initial sample of 50 elements and validated with subject matter experts prior to beginning data analysis.

Implicit external references (Rhetorical: Citation). The value of this variable was either 'yes' or 'no' depending on whether or not the element referred implicitly to a source outside of the website on which the element appears. By its nature, an implicit reference involved a greater degree of interpretation than an explicit reference. A 'yes' value indicated that the element implied reference to another idea and that the reference is to a source external to the web site on which the element appears. I specified criteria for identification of explicit external references during the validation and reliability-testing phase.

Relation of implicit external reference to central claims (Rhetorical: Citation). Using this variable, I identified the ways in which an implicit external reference is related by the author to the subject(s), theme(s), author position, and/or audience position of element. Similar to *relation of explicit external reference to central claims*, values for this element included agreement, opposition, evidentiary, and neutral. The list of possible values for this variable were generated

inductively using my initial sample of 50 elements and validated with subject matter experts prior to beginning data analysis.

Characterization of implicit external reference (Rhetorical: Citation). This variable referred to the way in which the external reference is discussed or characterized by the element. Similar to *characterization of explicit external reference*, values included “authoritative” and “marginal”. The list of possible values for this variable was generated inductively using my initial sample of 50 elements and validated with subject matter experts prior to beginning data analysis.

Rhetorical: Audience variables. Treatment of audience (or “language and society”) included assumptions made by the author about the prior knowledge, experience, and attitudes of the reader, the methods of persuasion, calls to action, and the structure of argumentation. I coded the data set for evidence to address Bazerman’s concept of *audience*. For each website, I identified the knowledge and attitudes the author assumed the reader to have. I operationalized these questions as the variables that follow. My operationalization was validated with subject matter experts.

Stated or implied audience (Rhetorical: Audience). I identified the stated or implied audience for the element. Audience can be implied by association with other elements that explicitly identified audience. For example, on the “Thomas the Tank Engine” website (www.thomasandfriends.com), a link at the bottom of the home page was labeled “For Parents” and clicking it took the user to a web page with the banner “Welcome Parents!” Based on the content of the page, the stated audience for these elements was parents of young children. On the same site, the home page navigation includes the link “Activities.” There was no stated audience. The implied audience for elements on this page was young children. Some elements had multiple

audiences. For this interpretive element, I developed evaluation criteria using my initial sample of 50 elements and established validity and reliability prior to beginning data analysis.

Audience prior knowledge (Rhetorical: Audience). I used this variable to identify knowledge that the author assumes the audience to have. For example, on the “Thomas the Tank Engine” website (www.thomasandfriends.com), the home page featured a large illustration of the “Thomas the Tank Engine” character with no caption or label of any kind. As an element, the illustration carried an assumption of prior knowledge of the character on the part of the audience for the site. For this interpretive element, I developed detailed evaluation criteria using my initial sample of 50 elements and established validity and reliability prior to beginning data analysis.



FIGURE 3.7. Detail from an illustration of the character “Leatherhead”. Reprinted from “Teenage Mutant Ninja Turtles”. (2014). Nickelodeon. Retrieved from <http://www.nick.com/pictures/ninja-turtles/tmnt-mutation-gallery.html#&slider1=7>.

Audience attitudes (Rhetorical: Audience). I used this variable to identify attitudes that the author assumes the audience to have (Bazerman, 1988). For example, on the web page for the character “Leatherhead” on Nickelodeon website (see figure 3.7), Leatherhead the alligator was depicted in a cartoon illustration with arms outstretched, narrowed eyes, dark colors, and an open mouth featuring rows of sharp teeth. The caption described Leatherhead as “terrifying.” The author attempted to portray the character as threatening through details such as the sharp teeth

and assumed that the audience adopted an attitude consistent with this interpretation. For this interpretive element, I developed evaluation criteria using my initial sample of 50 elements and established validity and reliability prior to beginning data analysis.

Persuasion (Rhetorical: Audience). This variable referred to the type of persuasion attempted using this element. Values for this variable included “emotional connection”, “credibility”, and “logical argument” (*Aristotle et al., 1900*). For example, on the curriculum page for the Nickelodeon show *Dora the Explorer* (<http://www.nickjr.com/dora-the-explorer/about-dora-the-explorer/dora-the-explorer-curriculum.html>), the following text appeared:

Dora the Explorer teaches children how to observe situations and solve problems as they explore Dora's world with her. Along the way, kids learn basic Spanish words and phrases, as well as math skills, music, and physical coordination. The show is highly interactive, and Dora's young viewers are encouraged throughout the show to respond to Dora and to actively participate in the adventure through physical movement. (n.p.)

This text passage element employed logical argumentation to build the case for adults that watching *Dora the Explorer* is beneficial to children in several ways. In this case, the argumentation was not accompanied by evidence; nonetheless, the form of persuasion attempted by the author was logical argument. For this interpretive element, I developed detailed evaluation criteria using my initial sample of 50 elements and established validity and reliability prior to beginning data analysis.

Call to action (Rhetorical: Audience). Using this variable, I identified what the author used the element to ask the audience to do. For example, on the home page of the website for

Barney the Dinosaur (<http://www.barney.com/usa/index.asp>), a vertical rectangle on the right side of the page under the heading “ADVERTISEMENT” encouraged visitors to “Dress Up With Angelina Ballerina iPhone App.” At the bottom of the rectangle, a button proclaimed, “Buy Now.” The author used the cluster element of the advertisement to encourage the audience to purchase the Angelina Ballerina iPhone App. For this interpretive element, I developed detailed evaluation criteria using my initial sample of 50 elements and established validity and reliability prior to beginning data analysis.

Rhetorical: Author variables. Authorial representation (or “language and mind”) referred to all of the ways the author is represented in the text, including style (or voice), self-reference, and originality of the author’s claims. As Bazerman states it, “The human mind stands between the reality it perceives and the language it speaks in; statements reflect the thoughts, purposes, observations, and quirks of the individual” (1988, pp. 25-26). I coded the data set for each website for evidence to address Bazerman’s concept of *author*. For each website, I identified the broadness and originality of claims, idiosyncracies, values, experience claims, and observation claims. I operationalized these issues as the variables that follow. My operationalization was validated with subject matter experts.

Stated or implied author (Rhetorical: Author). I identified the stated or implied author for the element. Authorship was sometimes implied by association with other elements that explicitly identified the author. An element can have multiple authors or no apparent author. Many web page elements had no stated author leaving the audience to assume that the publisher of the website claims authorship for the content. For example, on the Crayola Kids website home page (<http://www.crayola.com/kids-playzone.aspx>), no author is listed. The “About Us” page (<http://www.crayola.com/about-us/company-profile.aspx>), described the corporation in almost

human terms (“Crayola is passionate about helping parents...” and “Crayola has called Lehigh Valley, Pennsylvania, its home...”, n.p.) and listed only the highest level executives in the company. In this example, the implied author for all content on the site was Crayola, the company. For this interpretive element, I developed detailed evaluation criteria using my initial sample of 50 elements and established validity and reliability prior to beginning data analysis.

Experience claims (Rhetorical: Author). I identified the experience claims of the author through the use of this variable. In the example text quoted above from the “Dora the Explorer” curriculum page, the author made an implicit claim of experience with evaluation of academic progress by viewers of the program. For this interpretive element, I developed detailed evaluation criteria using my initial sample of 50 elements and established validity and reliability prior to beginning data analysis.

Broadness (Rhetorical: Author). This variable identified the broadness of claims made by the author using this element. Values included *narrow*, *somewhat narrow*, *somewhat broad*, and *broad*. This element focused on how the author identified broadness, universality, and boundaries of claims. For example, in a passage on the PBS Sesame Street page for parents and educators (http://www.pbs.org/parents/sesame/about_ss_overview.html), the following paragraph appeared:

While Sesame Street's educational approach is key, it leaves plenty of room for humor. Sesame Street's iconic comedic style appeals to children and their parents alike, encouraging intergenerational co-viewing.

This text passage element excerpt demonstrated a broad rhetorical claim to the universality of the comedic elements of the television show, appealing to all ages. For this

interpretive element, I developed detailed evaluation criteria using my initial sample of 50 elements and establish validity and reliability prior to beginning data analysis.

Originality (Rhetorical: Author). This variable identified the originality claimed by the author with this element. Values included *common*, *somewhat common*, *somewhat original*, and *original*. This element focused on how the author identified the originality of claims. The following partial transcript is from a video on the Nickelodeon page promoting the television show “Life with Boys” (<http://www.teennick.com/shows/life-with-boys>):

TESS: When you’re a 14-year old girl, living with your dad and three brothers, it can seem like life is just all boys, boys, boys...

ALLIE: Um, Tess, if you’re any 14 year old girl, life is pretty much boys, boys, boys!

This element was coded as *common* for the *Originality (Rhetorical: Author)* characteristic because through it the author claimed commonality and universality (Bazerman, 1988) of the attitudes described. For this interpretive element, I developed detailed evaluation criteria using my initial sample of 50 elements and establish validity and reliability prior to beginning data analysis.

Idiosyncrasies (Rhetorical: Author). This variable identified to what extent the rhetorical claims of this element could be considered distinctive, peculiar, or abnormal. I operationalized Bazerman’s (1988) notion of idiosyncrasies using a Lickert scale (Dillman, 2000) to facilitate consistent application. This variable was scored using the following scale: customary, somewhat customary, somewhat idiosyncratic, idiosyncratic. The following example was taken from the character biography page for the character Tess from the Nickelodeon page for the show “Life with Boys” (<http://www.teennick.com/shows/life-with-boys/characters/tess-foster.html>).

Tess has the self-confidence to make a run at most challenges that come her way. But make no mistake; she also has the vulnerabilities that show she cares what people think of her. Well-liked, but not popular, Tess, as with most teens, wants to fit in. But, unlike most teens, it is just not in her DNA to bend too much to accomplish that... and that will definitely complicate her life.

In this text passage element, the author made a rhetorical claim that the character Tess was idiosyncratic in her unwillingness to compromise despite her desire to fit in. For this interpretive element, I developed evaluation criteria using my initial sample of 50 elements and established validity and reliability prior to beginning data analysis.

Formal variables². The following variables represent aspects of the elements closely related to form or structure. The formal variables included many of the linguistic descriptors drawn from a text analysis tool called Coh-Metrix. The formal variables also included descriptors of images and text image relationships not included in Coh-Metrix.

Coh-Metrix. Coh-Metrix is a tool developed at the University of Memphis as an automated measure of text cohesion. The tool accesses several other databases and text analysis tools to provide a broad spectrum of text analysis data from a single tool. In addition to simple measures of text readability, such as word count, average word length, and average sentence

² Many of the indexical, substantive, and rhetorical variables contribute to an understanding of the formal characteristics of the websites.

Although I do not equate genre with formal characteristics of a text, formal characteristics provide important evidence of the genre. Formal elements are the tracings of discursive action. In other words, the formal elements of a genre are artifacts of the regular patterns of communication that occur within discourse communities; they are evidence of the discourse patterns, not the discourse patterns themselves. A careful examination of the formal elements provides one source of evidence for differentiating genre. Text characteristics such as complexity, formality, and coherence help mark a text as being produced within a particular genre.

length, Coh-Metrix looks for characteristics within the text that are likely to help the reader create a coherent mental model of the ideas represented in the text (Graesser, McNamara, & Louwerse, 2003; Graesser, McNamara, Louwerse, & Cai, 2004). Syntax and parts of speech are analyzed using an automated parser developed by Charniak (2000). Linguistic and semantic information about words is provided by a database called WordNet (Fellbaum, 1998; G. A. Miller, Beckwith, Fellbaum, Gross, & Miller, 1990). Latent semantic analysis (LSA) is conducted using other tools (Landauer & Dumais, 1997). A complete list of indices reported by Coh-Metrix is included as Appendix A.

Although Coh-Metrix provided several indices to describe the structure of a given text, I used these measures of linguistic complexity as a means for demonstrating the degree of similarity between and among several text passages because previous research has indicated that text complexity can be used to establish genre (Biber, 1988, 1995; Chafe, 1982; Emigh & Herring, 2005; Heylighen & Dewaele, 1999). The subject of this study is genre, not readability. I did not use these measures of linguistic complexity to attempt to seek implications related to the readability level or complexity level of the texts.

Word count (Formal). For text elements, this variable reflected the word count provided by the Coh-Metrix online tool. For video or audio elements that include language, I generated transcripts and applied *Word count* and all of the other text specific variables to the transcript. For non-text elements, the value was N/A.

Sentence count (Formal). For text elements, this variable reflected the sentence count provided by the Coh-Metrix online tool. Sentences are identified by Coh-Metrix based on terminal punctuation, i.e. period, question mark, exclamation point. For non-text elements, the value was N/A.

Paragraph count (Formal). For text elements, this variable reflected the paragraph count provided by the Coh-Metrix online tool. Paragraphs were identified by hard returns, not by indentation. For non-text elements, the value was N/A.

Syllables per word (Formal). For text elements, this variable recorded the average number of syllables per word, as reported by Coh-Metrix. The number of syllables is determined by Coh-Metrix by applying a list of human syllable counts for English words. For non-text elements, the value was N/A.

Words per sentence (Formal). For text elements, this variable recorded the average number of words per sentence, as reported by Coh-Metrix. For non-text elements, the value was N/A.

Sentences per paragraph (Formal). For text elements, this variable recorded the average number of sentences per paragraph, as reported by Coh-Metrix. For non-text elements, the value was N/A.

Flesch reading ease (Formal). This variable reported the Flesch Reading Ease score as calculated by the Coh-Metrix engine. Readability formulas were first developed in the 1920s and were popularized beginning in the 1950s. Readability formulas provide a simple, direct estimate of the difficulty or complexity of a text passage. The accuracy of readability formulas depends on the premise that the number of syllables in a word functions as a proxy for vocabulary difficulty. The formulas also rely on the number of words in a sentence to function as a proxy for sentence complexity. In most written English, these are reasonable assumptions that provide useful estimates, although unusual passages will cause unreasonable estimates and over-reliance on readability formulae can be problematic. Readability is an appropriate measure to use in this study because readability will be calculated to note similarities and differences between different

text passages, not to ascertain the appropriateness of text for students in specified grades. Flesch Reading Ease is calculated with the following formula: $\text{Score} = 206.835 - (1.015 \times \text{Average Sentence Length}) - (84.6 \times \text{Average Number of Syllables per Word})$. This formula results in a number between 1 and 100 with higher numbers representing easier text.

Flesch-Kincaid grade level (Formal). This variable reported the Flesch-Kincaid Grade Level score as calculated by the Coh-Metrix engine. In the 1970s, the U.S. military commissioned Kincaid to revise the Flesch Reading Ease score to provide an estimate of the U.S. school grade level necessary to read text at the specified difficulty. The Flesch-Kincaid formula is as follows: $\text{Score} = 0.39 (\text{average words per sentence}) + 11.8 (\text{average syllables per word}) - 15.59$.

Concreteness (Formal). This variable recorded the concreteness of vocabulary, as reported by Coh-Metrix through the Medical Research Council (MRC) Psycholinguistic Database (McNamara, Louwerse, McCarthy, & Graesser, 2010). The database is a human-rated tool that produces scores, which range from 100 to 700, with higher scores representing more concrete vocabulary. For example, the word “table” produces a concreteness score of 604 while the word “this” produces a concreteness score of 240. “Table” represents a more concrete concept than “this”.

Noun hypernymy (Formal). This variable was calculated using a database containing a conceptual hierarchical taxonomy of English nouns. At the highest levels of the taxonomy, nouns are very broad and general. The hypernymy score for a noun reflects the number of levels above that noun in the hierarchy. A noun with a lower numeric score is more general. For example, the relatively non-specific noun “food” has a noun hypernymy score of 4.333. Moving down through the hierarchy, “vegetable” is a more specific noun with a score of 7, and “broccoli” is more

specific still with a score of 8.5. As reported by Coh-Metrix, this variable reflected the average number of levels in a conceptual taxonomic hierarchy above nouns that appeared in a given passage. This was a rating of the concreteness of nouns in a text passage.

Verb hypernymy (Formal). Similar to Noun Hypernymy, this variable represented the average number of conceptual levels above verbs that appear in a given passage and was reported by the Coh-Metrix tool.

Personal pronouns (Formal). An additional measure of text formality reported by Coh-Metrix, this variable was a ratio of personal pronouns to total word count in a text passage.

Displayed size (Formal). For image elements, this variable reflected the size of the image in pixels as displayed on the page. The size was a straightforward value not subject to interpretation. Size was determined by examining the element attributes within the web browser. The size of an image provides an indicator of its importance to the overall message of the page. Consideration of images was essential because web pages are governed more by the mode of the image than the mode of text (Kress, 2003b). The number and size of images has been shown to differ by genre (Marsh, 2002). Exclusion of images from evaluation could distort or misrepresent the content of the websites.

Relative size (Formal). The relative size of the images provided a measure of the emphasis or importance a given author places on the image mode. That relative emphasis was a formal structural element of a website and may help distinguish genre. For image elements, this variable reported the approximate percentage of the web page covered by the image in question. Modern audiences access web pages on a variety of devices with a wide range of screen resolutions and display preferences, from mobile devices to very large television displays. For this reason, *displayed size* was not sufficient for describing the prominence of an image on a web

page. *Relative size* provided the necessary additional information. For each image element, I estimated the relative width and the relative height to calculate the relative size. The proposed benchmark values for relative width were *up to 25% of page width, up to 50% of page width, up to 75% of page width, and up to 100% of page width*. The proposed benchmark values for relative height were *up to 25% of page height, up to 50% of page height, up to 75% of page height, and up to 100% of page height*. The relative size was calculated by multiplying the two percentages. For example, an image element that was estimated at 50% of page width and 50% of page height represented approximately 25% (50% X 50%) of the total area of the page. The benchmarks and methods of estimation were validated with subject matter experts prior to data collection.



FIGURE 3.8. Image element depicting the character Donatello. Reprinted from "Donatello - Ninja Turtles - TMNT Characters". (2014). Nickelodeon. Retrieved from <http://www.nick.com/ninja-turtles/donatello>.

Color (Formal). The dominant colors in each image were recorded. Colors are used to communicate meaning within visual media (Wierzbicka, 1990). For this reason, information about the dominant colors may help identify purpose, audience, and topic during interpretation of data. Figure 3.8 displays an image element with green as the dominant color, because approximately two thirds of the image is colored with several shades of green. The list of colors

used for coding were derived with the initial 50 element sample and validated with subject matter experts.

Image type (Formal). For image elements, I recorded the type of each image. The possible values for image type were derived during my initial coding of 50 elements and were validated by subject matter experts. Initially, image types included realistic illustration, photograph, cartoon image, and diagram. Figure 3.9 displays examples of image types from the Yahoo Kids home page (<http://kids.yahoo.com>).

Realistic illustration



Photograph



Cartoon image



FIGURE 3.9. Three image elements of different image types. Retrieved from the Yahoo Kids home page (<http://kids.yahoo.com>).

Image relationship to text (Formal). According to Lemke (2005), "Texts and images in genuinely multimodal genres are mutually organizing" (p. 49), therefore describing that relationship is important to understanding the relative purposes of image and text within the genre. The relationship between images and text was described using Marsh's (2002) coding scheme for image-text relationships on web pages. According to this three-faceted typology, a) *image functions express little relation to text*, b) *image functions express close relation to text*, or c) *image functions go beyond the text*. Figure 3.10 displays a portion of the biography page for a character on the Thomas the Tank Engine website.

In this example, the relationship between the image and the text was close because both the image and the text function to define the character of Toby. The reader learns that Toby is

brown and square from both the image and the text. The data-coding manual contains definitions, as described below. A single image may have distinct relationships with more than one text passage. I used clusters (see cluster analysis) to parse out multiple text-image relationships involving the same image.

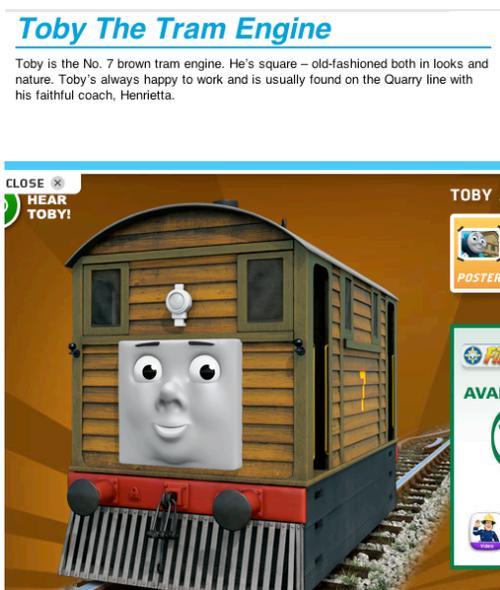


FIGURE 3.10. A portion of a character biography page. Reprinted from “Toby the Tram Engine.” (2014). Gullane Thomas Limited. Retrieved from <http://www.thomasandfriends.com/usa/Thomas.mvc/EngineDetail/Arthur>.

Documentation of essential characteristics. The instruments and methods developed were designed to document the essential characteristics of websites, i.e., multimodality, multilinearity, instability, and intertextuality. Multimodality was documented through the application of the *element type* variable at the element level and the *element types present* and *primary element type* variables at the page level. The process documented modes present on each page and the centrality of each mode to the communicative goals of the overall websites. Multilinearity was addressed by documenting the organizational patterns of the websites through the use of site maps and graphical representations of the site structures. In addition, cluster analysis promoted further exploration of multilinearity by providing a basis from which to

examine the interaction of non-linear elements. Although, instability must necessarily be studied over the course of time, this methodology provided a basis from which to examine instability. The careful documentation of a website at a point in time will enable future research into changes to the websites in question and to the genres in which they participate. Finally, the methods provided a foundation for examining intertextuality by documenting apparent references exemplified by explicit and implicit external references. It is important to note that although this methodology provided a necessary foundation for understanding multilinearity and intertextuality, producing a full description of either characteristic would require directly studying readers, writers, or both. Likewise, the methodology described in the current study provided a necessary foundation for a study of the instability of children's websites, but the full description that characteristic requires data beyond the scope of the current study, namely, data collection at multiple points in time.

Validation of the data evaluation checklists. As previously stated, preliminary values for the *element type* variable included text passage, text label, illustration, photograph, video, map, diagram, internal link, external link, or cluster of other elements. For the purposes of ensuring that the instruments allowed adequate description of each element type, I included at least five examples of each element type in my instrument development sample. Five examples of each of the ten element types led me to select 50 elements, and 50 web pages, for my instrument development sample. I conducted an initial coding of 50 web pages from children's websites using the *Web Page Data Evaluation Checklist*. I conducted an initial coding of 50 elements from children's websites using the *Element Data Evaluation Checklist*. The 50 web pages were selected beginning with children's websites with which I was familiar, then choosing from posted lists of children's websites. I conducted multiple web searches using the phrase

“children’s websites” and culled the results for websites that I have observed children using or heard children mention. At each stage, I also tried to choose sites that appeared to be different (in form or content) from the sites I had already chosen. Although this was a convenience sample and not meant to be representative of a larger population of websites, I believed that diversity of content would help ensure that the tools developed in this phase would be useful on a variety of different websites. The 50 elements were selected from within the selection of 50 pages. The sampling process for the 50 pages and 50 elements contrasted with the purposeful sampling method I employed in phases two and three. The purpose of these data was to provide a starting point for the evaluation checklists prior to the establishment of validity and reliability. During this preliminary evaluation, I adjusted variables and evaluation scales on checklists using an emergent process to ensure that the evaluation criteria appropriately and adequately described the observed data.

I then consulted individually with experts in web design, literacy, and multimodality to validate my selection of these variables and evaluation criteria. With each expert, I provided my written description of the variables in question, my working definitions of the possible values, and a sample of elements to be measured using the variables and values. I asked each expert to address the appropriateness of the variable and the appropriateness of the possible values for the variable according to their area of expertise. I asked for alternative variables and value scales from subject matter experts. Based on the recommendations of these experts, I edited the list of variables and evaluation criteria. I documented the revision process thoroughly to provide an audit trail.

Establishing reliability of data evaluation checklists. Through the process described in the preceding paragraphs, I established and refined operational definitions for each of the

variables and created a data-coding manual for this study. The data coding manual included specific guidance on evaluating each variable along with examples. I asked two different coders to use the data-coding manual to code a random selection of ten pages and ten elements from my initial sample of 50 pages and 50 elements. Each coder held a Ph. D in Education or was an advanced doctoral student in Education. Intercoder agreement was calculated using a percentage of agreement method with a goal of 85% agreement or better. I adjusted the manual and codes as needed and recoded until at least 85% agreement was achieved. The following variables were subject to reliability testing: Organization, Subject(s), Theme(s), Author Position, Audience Position, Explicit External References (EER), Relation of EER to Central Claims, EER Characterization, Implicit external references (IER), Relation of IER to Central Claims, IER Characterization, Stated or Implied Audience, Audience Prior Knowledge, Audience Attitudes, Persuasion, Call to Action, Stated or Implied Author, Experience Claims, Broadness, Originality, Idiosyncrasies, Relative Image Size, Dominant Image Color, Image Type, Image Relation to Text.

Phase 2: Page Level Analysis

Following instrument development, I proceeded to a page level analysis of five purposefully selected children's websites. The page level analysis provided an overview of the subjects, rhetorical actions, and formal elements present on the five selected websites. One of the reasons for conducting a page level analysis was to guide the selection of a web site for the more in-depth element level analysis to follow in phase three. The page level analysis gave an indication of the typicality of the general features of children's websites and was be important in determining the next steps of possible genre identification. Another reason for conducting a page level analysis was that it proved to be a far more parsimonious method than the element level

analysis. Because this study was focused on producing a new methodology, it was important to include both page level and element level analyses.

Data collection. This section presents the rationale and the method for identifying and selecting specific websites for analysis within this study. With regards to selection, it is important to understand the challenges in defining what is and what is not a children's website.

Background. For the purposes of this study, children's websites are websites that identify children as their target audience. It is important to note that many websites visited by children are not identified as "children's websites". In a 2005 survey that sought to identify the most popular websites of 8-11 year old children, in addition to popular children's websites like Disney and Nickelodeon, the Kidsay survey reported that google.com, yahoo.com, and espn.com were among the top 15 websites visited by the children surveyed (as reported in Alvy & Calvert, 2008).

General ratings information easily tells us which websites on the Internet are the most visited overall. For example, a recent check of data from Alexa, an Internet information company, shows that the top five most visited websites were Google (48.84% of global traffic), Facebook (45.39%), YouTube (32.45%), Yahoo (20.88%), and Amazon (5.54%) (Alexa, 2012). One approach to determining the most visited children's websites using this data is to consider the overall ranking data for all websites and then to subjectively choose the websites that have children as the intended or expected audience. Taking this approach, the most visited website *for children* in this example would be Disney (disney.go.com) which is ranked 131st overall, with 0.158% of overall global visitors. Although Alexa disaggregates website visitor data by age groups, the lowest group for which they report data is 18-to-24-year-olds. A further examination of the statistics reported by Alexa reveals some of the assumptions underlying these statistics:

relative to the general population of web users, women between the ages of 25 and 44 who have children are greatly overrepresented among visitors to Disney's website. Therefore the rating information does not reveal how many children within different age groups choose to visit the website independently, or indeed if any children visit the website, independently or otherwise. While the ratings company Nielsen employs more extensive methods in determining web user data (Nielsen, 2012) including surveys and nation-wide user panels along with web traffic data, Nielsen's public releases of this data are sporadic. The most recent data on this topic published by Nielsen is a 2008 study of the habits of children in the United Kingdom (Nielsen, 2008). In addition, limited representation of sub-groups (such as children) within sampling populations (Belkin, 2006) may limit the usefulness of this data in determining which websites are most popular with children. More research into the websites that children choose to visit on the Internet, regardless of the intended audience of the website, is needed. In the absence of that research, this study addressed websites with children as an intended audience, as identified by considering multiple sources from a number of expert audiences. In all of these analyses, it is important to note that children are not a homogenous group and that treating them as such may tend to marginalize non-majority children of all kinds. More research into the Internet usage among diverse groups of children is needed.

Sampling method. To select a sample of children's websites for the second phase of the study, I examined several lists of popular children's websites. I obtained market research data from the two largest firms that provide such data for the United States, Nielsen and Alexa. I also used lists of top children's websites produced within the past five years by the American Library Association, Kidsites.com, Common Sense Media, Kido'z, and Digital Trends. Each of these lists meet the following criteria: a) a relatively current list of websites for children, b) a list

compiled by a national organization with a stated interest in addressing the needs of children and families, and c) an organization that is not promoting their own website with content for children. I looked for websites that were common to most or all of these lists and selected a sample of five websites for the second phase of analysis. Five is a sufficiently large number to examine similarities and differences within the group of websites. For this study, which was focused on developing a methodology, I did not seek statistical generalizability to a larger population of websites.

Data analysis. For each website selected for analysis, I began by attempting to generate a map of the entire website, using an online site map creation tool (<http://www.xml-sitemaps.com/>). A site map is a hierarchical text outline that lists every page included within a given domain. The tool is essentially a computer program that starts at the home page and follows every possible internal link on every page and generates a map of the functional structure of the website. An example of the first page of a text site map is included here as figure 3.11.

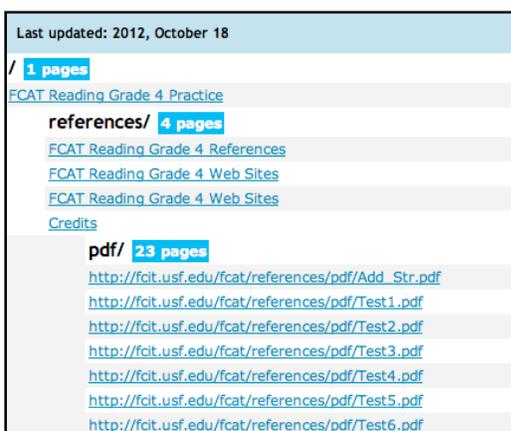


FIGURE 3.11. First page of automatically generated site map.

Using this outline, I generated a graphic representation of the structure using concept mapping software (Inspiration). This visual model allowed an empirically-grounded description of the nature of the structure, e.g., non-linear, hierarchical, mixed, etc. (McDonald & Stevenson,

1998; Mohageg, 1992). For example, a non-linear topology would show cross-referential links to and from many nodes, which might result in a map like figure 3.12. A website that is organized with a hierarchical division of content of different types or for different purposes may result in a map like figure 3.13. This analysis allowed me to see the general topography of the website and helped guide the remainder of the content analysis.

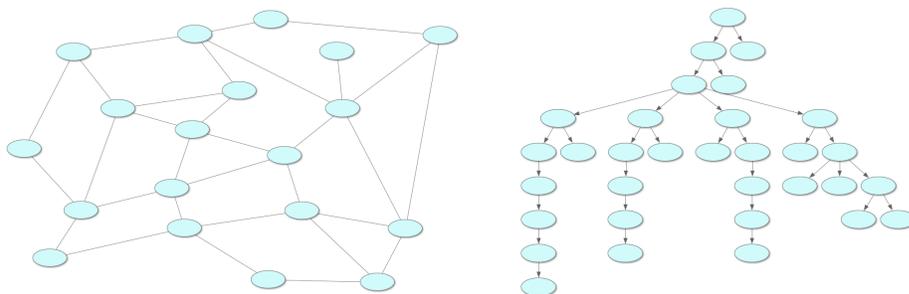


FIGURE 3.12. Non-linear structure. **FIGURE 3.13.** Hierarchical structure.

Next, I attempted to download the entire contents of the website to my computer hard drive using Internet archive software (<http://www.ansemond.com>). Creating a local copy of the website was necessary since website content changes frequently. In addition, I documented each page on the website using a series of screenshots while navigating the website. These static images of the web pages as they were displayed was used to ensure the local archive versions of the websites contain all stylesheet information and media content from the original web-based version of each website. Out of consideration for the dynamic nature of websites, data collection for all websites in the initial sample occurred on the same day and were considered as a point in time. Website content changes frequently and the data collected represented the selected websites at one point in time. This “snapshot” method is both a feature and a limitation of the research design for this study.

I applied the *Web Page Data Evaluation Checklist* to all pages on all five selected websites following procedures established during the instrument development phase and

described in the data-coding manual. After applying the *Web Page Data Evaluation Checklist*, I generated descriptive statistics for all five websites. Descriptive statistics included measures of central tendency and distribution of web pages by subgroups of subjects, functions, and element types.

Next, I wrote a rich narrative description of each website. The narrative began with an accounting of the data recorded on the *Web Page Data Evaluation Checklist* and inferences drawn from the generation of site maps. The narrative then proceeded to inferences based on the recorded data and conclude with an overall summary of each website.

Next, I conducted a statistical comparison between the five websites based on the *Web Page Data Evaluation Checklist*. I created a cross tabular distribution of web page data and performed a series of chi square tests of association (Glass & Hopkins, 1996) to explore possible relationships between different subject, function, and element type subgroups and inclusion of pages on web sites. In other words, I described the statistical likelihood of subjects, functions, and element types to be associated with web sites within the sample.

Finally, I wrote a narrative comparison of the websites. The narrative comparison described the statistical accounting of the differences between the subjects, functions, and element types and noted similarities and differences between the websites. I engaged in a process of data triangulation by comparing inferences that are supported by each type of data and looking for concordant and contradictory inferences. Inferences regarding the genre of the websites that were supported by multiple data sources may be regarded as more trustworthy.

Evidence of validity. This section addresses the steps that I have taken to insure the credibility and trustworthiness of the page level analysis, phase two of the study. The validity procedures for phases one and three are addressed in those sections.

First, it was important that I established that the five websites selected for analysis were examples of children's websites. I cross-referenced several different published sources of information, including Nielsen Company, Alexa Web Information, American Library Association, Kidsites.com, Common Sense Media, Kido'z, and Digital Trends, to construct the list of possible websites. This process of selection based on reporting from multiple sources provides evidence of validity as it constitutes structural corroboration based on consensus (Eisner, 1997).

Second, the critical processes of researcher reflexivity (Creswell, 2007; Creswell & Miller, 2000) allow readers to understand a researcher's positions and take them into account when considering conclusions. I have self-disclosed my "assumptions, beliefs, and biases" (Creswell & Miller, 2000, p. 127) in the *Role of the Researcher* and *Researcher's Relationship to the Study* sections of this chapter. Readers may consider social and cultural forces that may shape my interpretations. I maintained a researcher reflective journal (Janesick, 2004; Patton, 2002) in which I monitored, reported, and reflected upon my interpretations of the data and the analysis process throughout the study.

Third, I documented research decisions by establishing an audit trail. The audit trail detailed the data collection and analysis procedures and decisions in minute detail. Establishing an audit trail involved "journaling and memoing, keeping a research log of all activities, developing a data collection chronology, and recording data analysis procedures clearly" (Creswell & Miller, 2000, p. 128). The audit trail helped establish that inferences are logical and that findings are grounded in the data. The audit trail also allowed me to document other actions that I took to increase validity throughout the study in response to emergent events.

Finally, during data analysis, I compared inferences supported by different forms of data in order to triangulate findings. This process included both data triangulation (seeking the same type of data from multiple sources) and method triangulation (seeking the same inferences from multiple methods). Data triangulation in this study was achieved by analyzing the same variables from multiple websites. Method triangulation was achieved in this study by analyzing many different kinds of data for each website. Triangulation (Creswell & Miller, 2000; Denzin, 1978) increases trustworthiness of findings by establishing a broader evidentiary base for claims.

Phase 3: Element Level Analysis

For the third phase of the study, I selected an information-rich case from the initial five sites analyzed in phase 2. The website was selected based on purposeful sampling (Patton, 2002), specifically, a website that was likely to reveal a great deal about issues of importance to addressing my research questions. The specific criteria for selection of the information-rich case was based on the first phase of analysis and was documented in detail to maintain an audit trail. Additional websites from the initial five websites were considered using the same rationale.

Data collection. I used the data collected during Phase 2 to conduct the element level analysis. From the previous analysis, I had a list of all of the web pages on the selected website.

Data analysis. I completed the *Element Data Evaluation Checklist* for all elements on selected pages from this website. The checklist involved analysis based on several different tools and frameworks, including Coh-Metrix (Graesser et al., 2004), web page cluster analysis (Baldry & Thibault, 2006), multimodal analysis of video content (Baldry & Thibault, 2006), and rhetorical analysis (Bazerman, 1988).

Next, I generated descriptive statistics for the targeted website based on the *Element Data Evaluation Checklist*. Descriptive statistics included a distribution of the percentage of the substantive, rhetorical, and formal elements found on selected pages within the targeted website.

Using the descriptive statistics as a starting point, I wrote a rich narrative description of the website including contextualized descriptions of the elements. The narrative was a necessary step to prevent over-reliance on quantitative summaries of the website content, in particular because I was establishing a methodology. While the strength of the checklist is in reinforcing a consistent procedure and a uniformity of analysis, the strength of the narrative complemented this by allowing the researcher to describe the gestalt of a web page. It was a holistic approach that acknowledged the possibility that some salient features of the site may not be reflected in the checklists. The narrative then proceeded to inferences about substantive, rhetorical, and formal aspects of the website grounded in the previous data analysis.

Having two different analyses of this website – Page Level from Phase 2 and Element Level from Phase 3 – afforded me the opportunity to compare the two. I considered the relative strengths of the two different levels of analysis.

At that point, I reviewed the analysis of the targeted website and concluded the analysis after determining that further analysis was no longer introducing categories, themes, or conclusions (Janesick, 2004; Suter, 2012). I determined that I did not need to conduct an element level analysis of a second website.

Evidence of validity. This section addresses the steps that I took to insure the credibility and trustworthiness of the element level analysis, phase three of the study. The validity procedures for phases one and two were addressed in those sections.

As with the validity of the page level analysis, the critical processes of researcher reflexivity (Creswell, 2007; Creswell & Miller, 2000) allow readers to understand a researcher's positions and take them into account when considering conclusions. I have self-disclosed my "assumptions, beliefs, and biases" (Creswell & Miller, 2000, p. 127) in the *Role of the Researcher* and *Researcher's Relationship to the Study* sections of this chapter. Readers may consider social and cultural forces that may shaped my interpretations and perspectives. I maintained a researcher reflective journal (Janesick, 2004; Patton, 2002) in which I monitored, reported, and reflected upon my interpretations of the data and the analysis process throughout the study.

In addition, I documented research decisions by establishing an audit trail. The audit trail described the data collection and analysis procedures and decisions in minute detail. Establishing an audit trail involved "journaling and memoing, keeping a research log of all activities, developing a data collection chronology, and recording data analysis procedures clearly" (Creswell & Miller, 2000, p. 128). The audit trail helped establish that inferences are logical and that findings are grounded in the data. The audit trail also allowed me to document other actions that I took to increase validity throughout the study in response to emergent events.

Finally, during data analysis, I compared inferences supported by different forms of data in order to triangulate findings. This process included both data triangulation (seeking the same type of data from multiple sources) and method triangulation (seeking the same inferences from multiple methods). Data triangulation in this study was achieved by analyzing the same variables from multiple websites. Method triangulation was achieved in this study by analyzing many different kinds of data for each website. Triangulation (Creswell & Miller, 2000; Denzin, 1978) increased trustworthiness of findings by establishing a broader evidentiary base for claims.

Interpretation of Results

I began interpretation by reviewing and rereading all data analyses and descriptive statistics. This was an important step to adequately ground my interpretation because of the large volume of raw data and large volume of summary analyses of that data. I used color coding throughout the process of data collection and analysis to differentiate between substantive, rhetorical, and formal elements. Each of the three aspects received equal importance because all three are essential and over-reliance on one aspect can result in a view of genre that is overly mechanistic or insufficiently grounded in rhetorical practice (Devitt, 2009; Miller, 1984).

I annotated the data analyses, descriptive statistics, researcher notes, and rich narrative descriptions for evidence of each research question and sub-question. I used the question numbering as an ad hoc code for this purpose. For example, statistics communicating the concentration of different topics were coded (2B) because that piece of information helped me answer research sub-question 2b, “What are the typical topics addressed by selected children’s websites?” Likewise, a section of my process notes that reflected an insight about the nature of children’s website topics would also be coded (2B). Data sources were reorganized by research question so that I could pull across many different data types to seek answers to research questions.

Once all data analyses, descriptive statistics, researcher notes, and rich narrative descriptions had been annotated with appropriate codes, I reorganized the data by code and reread the sections corresponding to each sub-question. I composed an answer to each sub-question based on all available data and analysis. After answering each sub-question, I considered the sub-question answers along with data analyses, descriptive statistics, researcher notes, and

rich narrative descriptions that had been annotated with each question code and composed an answer to each research question.

Role of the Researcher

In qualitative paradigms, the researcher is the primary instrument of data analysis. All conclusions are subjective and context-bound. However in all paradigms, quantitative and qualitative, the researcher plays a central interpretive role. Data never speaks for itself. The researcher determines what questions will be asked and how those questions will be answered. More importantly, the researcher's assumptions and beliefs frame what questions *can* be asked and how those questions *can* be answered. Because of the central role of the researcher in framing the research, disclosure is a necessary part of interpreting the research. When discussing content analysis in particular, Krippendorff (2004) expressed:

Consider common findings of political biases, racial prejudices, and the silencing of minorities on television as such issues. Although counts of evident incidences of such phenomena can give the impression of objectivity, they make sense only in the context of accepting certain social norms, such as the value of giving equal voice to both sides of a controversy, neutrality of reporting, or affirmative representations. Implying such norms hides the context that analysts need to specify. Unless analysts spell out whose norms are applied, whose attitudes are being inferred, who is exposed to which mass media, and, most important, where the supposed phenomena could be observed, their findings cannot be validated.

(p. 28)

Researcher's Relationship to the Study

I grew up less than two miles from the university in an economically distressed area that has been referred to as “suitcase city” because of a highly transient population. When I grew up there in the 1970s, it was a working class suburb of major metropolitan area. I was the youngest of five in a solidly working class family. From a young age, I had a strong interest in television and movies, which later translated to a love of theater of all kinds. I love reading and from a very young age I found myself drawn to understanding the structure of texts and how texts related to each other. When I was eleven, I loved mysteries, including the Choose-Your-Own-Adventure series. Books in that series included decision points at which the reader chose which path the characters would take and followed instructions in the text to proceed along a certain path. For example, a prompt might say, “If you think the team should enter the door on the left, turn to page 76. If you think the team should proceed down the staircase, turn to page 112.” I loved reading and rereading those books to discover all of the possible combinations of outcomes. I also loved creating complete maps of the text that showed every possible path to each outcome. I now see those texts as multilinear, as postmodern, and as forerunners to hypertext novels. I now see my desire to map them as an early and basic desire to understand the structure and possibilities of text.

Prior to entering the field of education, I worked as a master control technician at a community television station. I was trained in the technical operation of television equipment and was responsible for channels control operations (e.g. playback of television programming, quality control of programming, scheduling, liaising with community producers, generating schedules, etc.) I also supported and facilitated community producers in the production of their shows and produced my own shows. Community producers range from professional

videographers, editors, technicians, and hosts to talented amateurs with commitment and time. Public access television encourages both innovation and mimicry of mature genre forms. Some producers attempted (with varying results) to recreate mainstream television genres such as variety shows, news magazines, call-in community affairs talk shows, televised sermons, and talent shows. The atmosphere at the television station was one of experimentation with genre and informed innocence with the medium of television.

I also spent almost ten years working for a company that produced corporate events. I travelled around the U.S., to Europe and the Caribbean, producing parties for corporate conventions, leadership retreats, sales meetings, and trade shows. The time I spent there influenced how I think about approximations of culture, allusion, cultural reference, appropriation of symbols, and multimodal symbolization. My job was to provide a bridge between the discourse of corporations and the discourses of caterers, performers, scenic artists, musicians, choreographers, and designers.

I became a classroom teacher because I was seeking a sense of fulfillment in my work. While I enjoyed everything I had done before, I wanted to help make the world a better place. I believed and still believe that education is sacred work. Helping others improve their lives is a worthy way to spend one's life. My background with video became a part of my teaching, as did my background with event planning. I had developed an innate appreciation for communication in many modes and I encouraged it in my students.

At the same time, I began working for the Florida Center for Instructional Technology. For FCIT, I shot and edited video for web based collections for classroom use. I began working on my Masters degree and eventually FCIT offered me the opportunity to pursue my degree full time while working at the university. Leaving the classroom was an extremely difficult decision.

Working at FCIT afforded me the opportunity to work with pre-service teachers, including many projects working with children on multimodal literacy and technology and working with university faculty, school leadership, and district personnel. As part of my work at FCIT, I helped found and continue to run a digital video summer camp. I have watched and helped children make movies and understand the elements of film genre for the past ten years. In ways that are very similar to public access television, the summer camp is a place for talented and interested amateurs to consider problems of communication with different multimodal discourses, to mimic mature forms, and to innovate in a safe, supported space.

My work at FCIT has included the creation and redesign of many websites and databases for education. While I have some level of technical expertise in some areas, my primary role is to liaise with people in various roles. On FCIT project teams, we spend a lot of time considering design elements and application functions on websites and how they will be interpreted by teachers and administrators in school settings. When considering a design element on a new website, I work with web designers who consider how other similar website design teams, or rhetors, have solved similar communication problems, or exigencies. We consider current typical generic practices in the field, including expectations of audience and how those expectations are interpreted by rhetors. My job is to provide a bridge between the discourses of education and educational research and the discourses of web designers and database programmers.

My doctoral preparation in literacy studies has centered on new literacies and how reading and writing have changed and are changing in response to technological and societal changes in the past fifty years. Our work has been grounded in two distinct contexts: K-5 classrooms and pre-service teacher education.

My approach to genre, in addition to all of the genre researchers cited in this study, is influenced by my past experiences. I identify as a feminist and a postmodernist, and I tend to adopt a critical perspective. I value education and believe that educators have a responsibility to help people understand how to change their lives by navigating complex structures of rhetoric and power. From varying levels of experience and expertise, I have helped create television shows, movies, websites, and stage productions and I have been a fan of all of these forms. I approach problems of communication from a literacy perspective with a grounding in childhood education and classroom practices. These are the lenses through which I view the problem of genres of children's websites.

Chapter Four: Findings

"Does the category of genre remain important, useful, necessary; does it become more or less important in the era of multimodal communication? The answer is that the category of genre is essential in all attempts to understand text, whatever its modal constitution." Kress, 2003.

The purpose of this study is to establish a comprehensive methodology for analyzing children's website content, based on both linguistic and rhetorical data, by employing defensible criteria to evaluate both qualitative and quantitative data. By applying genre theory through a mixed methods content analysis, this study addresses gaps between research into online genres and research into children's reading on the Internet. This study is developed around the supposition that understanding the texts that children choose to read online, not exclusively informational or academic texts, is necessary in order to fully understand reading in online environments designed for children. I use genre theory as a prism for examining form, substance, and rhetorical action within children's websites. The research questions for this study are:

- 1) What methodology allows for the comprehensive study of genre of children's websites?
 - a) What instruments can facilitate the systematic evaluation of children's websites?
 - b) What methods of analysis can be applied to understand the genre of children's websites?
- 2) What is the nature of the formal elements of selected children's websites?

- a) What are the formal textual elements associated with selected children's websites?
 - b) What are the formal visual elements associated with selected children's websites?
 - c) What text-image relationships are found within selected children's websites?
- 3) What is the nature of the substance of selected children's websites?
- a) What are the organizational and thematic units of within selected children's websites?
 - b) What are the topics addressed by selected children's websites?
 - c) How are the author and reader positioned by the text?
- 4) What is the nature of the rhetorical action of selected children's websites?
- a) What stated or implied communicative purposes are served by selected children's websites?
 - b) Within what context or community do selected children's websites operate?
 - c) Who is the intended or purported audience of selected children's websites?
 - d) What attitudes and prior knowledge does the author presume the reader to have?
 - e) How is authorship of selected children's websites identified?
 - f) How are implicit and explicit external references used to position the website by the author?

- g) What evidence points toward antecedent genres that fulfilled the rhetorical purposes of the genres represented within selected children's websites prior to the existence of children's websites?
- 5) What evidence supports or refutes the identification of a genre or genres within the body of children's websites examined within this study?

To answer these questions, I conducted a content analysis of five purposefully-sampled websites that are recognized as having children as their primary audience. For this purpose, I developed evaluation instruments, the Web Page Data Evaluation Checklist (WPDEC) and the Element Data Evaluation Checklist (EDEC), which incorporate many different analysis tools and techniques.

In this chapter, I will briefly describe the selection of the websites included in the analysis, and then address each research question and sub-question, citing specific examples from the data I have collected and analyzed.

Websites Selected

I selected five websites based on recommended lists published by the American Library Association (2014), KidSites.com (2014), Common Sense Media (2014), Digital Trends (2013), and based on market data from media ratings services Nielsen and Alexa. Ratings from Nielsen and Alexa provide an indication of the popularity of the websites with their target audience. Each of the four published lists meet the following criteria: a) a relatively current list of websites for children, b) a list compiled by a national organization with a stated interest in addressing the needs of children and families, and c) an organization that is not promoting their own website with content for children. I cross-referenced each of these lists with ratings information and chose these five sites: National Geographic Kids (kids.nationalgeographic.com) was the only

website (out of more than 500 websites) common to all four curated lists. PBS Kids (pbskids.org) is included on three of the four curated lists and ranks third in children's website rankings by traffic. FunBrain (funbrain.com) is included on three of the four curated lists and has an average traffic ranking of 14th. Seussville (seussville.com) was common to three of the four curated lists. Finally, Nickelodeon (nick.com) is consistently ranked as the most visited site children's website by far with an estimated 7,000,000 unique visitors each month. I chose to include it based solely on the strength of its popularity.

PBS Kids. PBS Kids (pbskids.org) is a website published by the Public Broadcasting Service (PBS), an American broadcast television network. PBS is a private, non-profit organization owned and operated by its 348 member stations across the United States (PBS, 2014). PBS is funded by its member stations and by the Corporation for Public Broadcasting, a private corporation created by and funded by the federal government. PBS has published content for children online since at least 1996. The PBSKids.org website was launched in 1999 (Internet Archive, 2014). The site's current design was launched in October 2013. PBS Kids, including all property sites comprises 13,944 individual pages³.

Nickelodeon. Nickelodeon (www.nick.com/) is the online home of the Nickelodeon cable television network and is published by Viacom International Incorporated, an American media corporation (Viacom, 2014). Nick.com was launched in 1996 (Eng, 1997) and a completely redesigned site was launched in July, 2014 (Steel, 2014). The redesign of the site is still in progress, a fact that affords an interesting opportunity to study the old and new designs at the same time. Nickelodeon is composed of 5,990 individual web pages.

³ The count of web pages published on a given website fluctuates as new pages are published and other pages are removed. The frequency and regularity of changes has not been studied. Studying the frequency and nature of changes and exploring factors that drive these changes would add to our understanding of online environments.

Seussville. Seussville (<http://www.seussville.com/>) is “the official home of Dr. Seuss on the Web.” The website is published by Random House Children’s Books in partnership with Dr. Seuss Enterprises (Random House, 2014). “Dr. Seuss” was the pen name of Theodor Geisel, an American author/illustrator of more than forty children’s books. Random House, a division of Penguin Random House, is one of the largest publishers in the world and is owned by German media company Bertelsmann and education publisher Pearson. Seussville was first published in 1997 and the current design was launched in 2010 (Internet Archive, 2014). Seussville comprises 747 pages.

FunBrain. FunBrain (www.funbrain.com) is published by the Family Education Network (FEN), which is owned by Pearson, an American education publishing and assessment company (Internet Archive, 2000). FunBrain was first published in 1997 by software company PMPublishing. FunBrain was acquired by FEN in March, 2000, which was itself acquired by Pearson later that year. The current site design was launched in 2011 (FunBrain, 2011). The site includes 661 pages.

National Geographic Kids. National Geographic Kids (<http://kids.nationalgeographic.com/>) is published by the National Geographic Society, an American nonprofit institution with a scientific education mission, founded in 1888. NGS publishes the magazines, *National Geographic* and *National Geographic Kids*. Children’s content has been published online by National Geographic since 1998 and the site had a major redesign in May, 2014 (Internet Archive, 2014). The site includes 458 pages.

Question 1: Instruments and Methods of Study

The instruments necessary for website content analysis include tools for parsing and indexing content, text analysis tools, color identification tools, and tools for managing data.

This study is a semiotic analysis, not a description of the technology used to create websites. Talking about website authoring can very quickly become a discussion of various file formats and programming techniques. This is not without cause, since technological dimensions and changes among them can have important implications for literacy and the study of literacy. However, while the constantly-changing array of technologies used to author websites are interesting and deserving of more attention, this study is focused on the human-perceivable elements as rendered on children's websites. In this study, I concern myself with the symbol, not with the formation of the symbol.

Identification and parsing of website content. The first task that must be accomplished in a website content analysis is the systematic parsing of content into smaller units. Websites are complex, multimodal environments. The first function of the instruments I have created, the Web Page Data Evaluation Checklist (WPDEC) and the Element Data Evaluation Checklist (EDEC), is to act as mechanisms to assist in identifying and breaking website content into a number of units.

A cursory examination of two pages from a children's website quickly reveals why a careful accounting for content is necessary. The homepage for Nickelodeon (www.nick.com) includes a link to a game page titled "Nickelodeon Kingdoms" (<http://www.nick.com/games/nickelodeon-kingdoms.html>). The SpongeBob Squarepants property page on the Nickelodeon site (<http://www.nick.com/spongebob-squarepants/>) also includes a link to "Nickelodeon Kingdoms." At least fifteen other Nickelodeon properties and the Nickelodeon Games directory also link to the "Nickelodeon Kingdoms" game page. Like most websites, Nickelodeon publishes no single exhaustive list of all pages and content. Instead content flows to different pages based on the choices of the user and the programming of the

website authors. This is the essence of web-based multilinearity. Content can be located and experienced in many different ways. Creating a list that is exhaustive involves a careful, systematic approach and the application of instruments like the WBDEC and EDEC.

Initially, I believed that pages could be catalogued by a largely automated process involving web crawler (also called “spider”) programs that generate sitemaps (Mirtaheri et al., 2013). While automated sitemap generation is an integral part of the process, it is insufficient to arrive at a complete catalog of website pages. Although some sites can be documented via web crawlers, this is not the case with sites like Nickelodeon for two reasons. First, sites like Nickelodeon are dynamically generated from a database of content. Rather than a static set of HTML documents housed on a publicly addressable server, pages on database-driven sites are generated upon the website user request (Rumianek, 2013). An analogy in a traditional book would be if you could see page 15 listed in the table of contents, but page 15 didn’t actually exist until you turned to the page. A database-driven website uses a set of rules and instructions to assemble the pages requested by users as they are requested. Since most pages don’t exist before requested through a browser, some crawler technology will not find them. This content is a part of what is referred to as “the deep web” (Bergman, 2001), because it exists inside databases and is not readable by search engines. Second, most website crawler software responds to any special instructions given by the website being crawled. By including a file named “robots.txt,” web programmers provide instructions for spiders, identifying which parts of the website the programmers would like to have indexed, which parts they would like ignored, and other behaviors such as frequency of indexing (Webmaster Tools, 2014). While these instructions are only guidelines, most spiders are programmed to obey them. There are many other strategies that web programmers can use to control the behavior of “bad” robots. Whether due to database

structure or to limitations placed on spiders by programmers, web crawling does not identify most content on websites like Nickelodeon. Using a spider is a useful first step, but the results of a crawl must be verified with direct observation of the live website. The WPDEC supports the verification process using conditional formatting to help identify and remove duplicate pages. An iterative process of harvesting pages and removing duplicates continues until no new pages are identified. This process, although labor-intensive, is necessary to provide a full picture of website content for database-driven websites.

Text analysis tools. As discussed extensively in chapter 3, the Coh-Metrix text analysis tool was applied to determine dimension of text cohesion and complexity. The Coh-Metrix variables that proved most informative were word count, Flesch-Kincaid Grade Level, syntactic simplicity, word concreteness, referential cohesion, and deep cohesion.

In addition, I used the WordNet database (Fellbaum, 1998; G. A. Miller, Beckwith, Fellbaum, Gross, & Miller, 1990) to produce consistent descriptions of concepts present on webpages. WordNet, a lexical database published by Princeton University, is one of the tools accessed by Coh-Metrix to aid in Coh-Metrix text analysis. WordNet contains 117,000 synsets (unordered groupings of synonymous concepts) and organizes them as part of a semantic, hierarchical network. For example, searching the word “babysitter” in WordNet produces a definition, synonymous terms “baby-sitter” and “sitter,” a list of sister terms, and a conceptual hierarchy that is eleven levels deep.

Color identification. Human perception of color varies greatly based on the age, medium, light source, and other environmental factors (Kandel, Schwartz, & Jessell, 2000). Human responses to color are both cognitive and affective and can be influenced by a variety of factors, including culture (Kress & Van Leeuwen, 2002). Therefore, interpretation of color

necessitates the identification of an interpretive frame. The color ‘blue’ has no intrinsic meaning; it acquires meaning by virtue of its relative position within a system of meaning (Kress & Van Leeuwen, 2002). Although color has the ability to convey meaning (Gage, 2001), color may or may not meet all criteria to be considered a semiotic system in its own right (Kress & Van Leeuwen, 2002). Regardless, Kress & Van Leeuwen argue convincingly that color is an important meaning-making element in a multimodal environment:

Language, image and music have been conceived of (and have in various ‘purist’ practices often operated) as relatively independent semiotic modes... This is not the case with color... Then again, maybe colour is a characteristic mode for the age of multimodality. It can combine freely with many other modes, in architecture, typography, product design, document design, etc., but not exist on its own. It can survive only in a multimodal environment. (pp. 350-351)

Regardless of the interpretive framework applied, it is necessary for a researcher studying websites to employ a consistent method for identifying and recording colors. Color is a visual property perceived by emitted or reflected light. Color is rendered differently depending on a variety of factors, including the software used to create the visual element, the software used to render the visual element, the hardware used, the light source, and the color model applied. A color model is an abstract system of describing color based on decomposing color into components (Fairchild, 2013). There are several different systems (e.g. RGB, CMYK, HSB, hexadecimal) and no one system is more correct than another system, although each system has different affordances (Fairchild, 2013). In a practical sense, this means that the Nickelodeon orange that displays on my computer screen, with my particular combination of hardware and software, may not be identical to the Nickelodeon orange displayed on someone else’s computer

screen, much less printed material or other real world objects. Internally, Nickelodeon identifies their specific orange within a commercial color system (Pantone) and directs their creative staff to use the CMYK color model when it is not possible to use a Pantone color (AIGA, 2003).

For the purposes of this study, color was identified using the format: [color name]-[hue]-[hexadecimal color code]. Colors were identified within the hexadecimal color space using graphics software, then translated into a color name and hue by employing the Color Name and Hue tool published by Colbindor (<http://www.color-blindness.com/color-name-hue/>). The online tool accesses a database of color names created by cross-referencing several color name lists and color name dictionaries (Mehta, 2014). By providing a description of a color within the RGB, HSB, or hexadecimal color space, the database provides a description of the color within each of the other color spaces, in addition to a unique name.

Managing data. The WPDEC and EDEC aid in the management of website data in three ways: controlled data entry, guided and supported application of codes, and aggregation of data from disparate media types. These instruments incorporate a number of frameworks for evaluation that are essential in understanding the work of children's websites, taking into account their multimodal, multilinear, and unstable characteristics.

The WPDEC and EDEC are applications created using Microsoft Excel and Visual Basic for Applications (VBA). Custom data entry screens guide the user through documenting and describing each aspect of a website and aggregating those variables for analysis.

Question 2: Formal Characteristics of Selected Websites

One of the purposes of this study is to look for evidence of one or more genres within the websites studied. Although rhetorical action is central to the notion of genre (Miller, 1984),

formal characteristics may be easier to spot because of their surface nature. Among the five websites studied, the formal characteristics revealed some immediate similarities and differences.

The total page count ranged from a low of 458 pages on National Geographic Kids to a high of nearly 14,000 pages on PBS Kids (see Table 4.1). Nickelodeon included 3,380 short video clips and 201 full episodes of network shows, which means that 60% of Nickelodeon pages had video as their primary mode. Nick is a site built around video. Figure 4.1 provides a graphical comparison of some of the data reported in table 4.1.

TABLE 4.1

Multimodality as Demonstrated by Formal Characteristics of Selected Websites.

	PBS Kids	Nick	Seuss- ville	Fun Brain	Nat Geo Kids
Total pages	13,944	5,990	747	661	458
<i>Primary element type</i>					
Video (Image)	41%	60%	2%	-	21%
Photo or illustration (Image)	6%	9%	32%	72%	10%
Game (Cluster)	28%	7%	1%	15%	8%
Text Passage (Text)	15%	20%	23%	9%	60%
Text Label (Text)	2%	0%	39%		
Internal Link (Action)	4%	4%	1%	4%	1%
External Link (Action)	4%	-	2%	-	-
Avg. reading lvl. for passages >50 words	7.1	7.2	10.6	6.0	9.3
Standard deviation for reading level	2.8	1.6	2.3	2.0	2.1

Two other sites displayed that kind of modal dominance. The largest content category by volume on the National Geographic Kids website was short informational articles (262, or 57% of the total site content). Each article included text (an average of 245 words at 9th grade level on Flesch-Kincaid), between one and four photographs, a data table, and a diagram or map. The

images included in these informational articles had a close relation to the text (Marsh, 2002). The site also hosted 95 short video clips, 48 printable coloring pages, and 35 online games.

By volume, FunBrain’s content was dominated by a format that may best be described as “web comics”. The site includes 378 pages of web comic content. A typical example of this content can be found in a series called, “Galactic Hot Dogs.” The series combines narrative text and comics-style illustrated panels. The series includes 26 “episodes,” most consisting of six pages of content. “Galactic Hot Dogs” was one of eight serialized titles available. FunBrain includes no videos and 106 games.

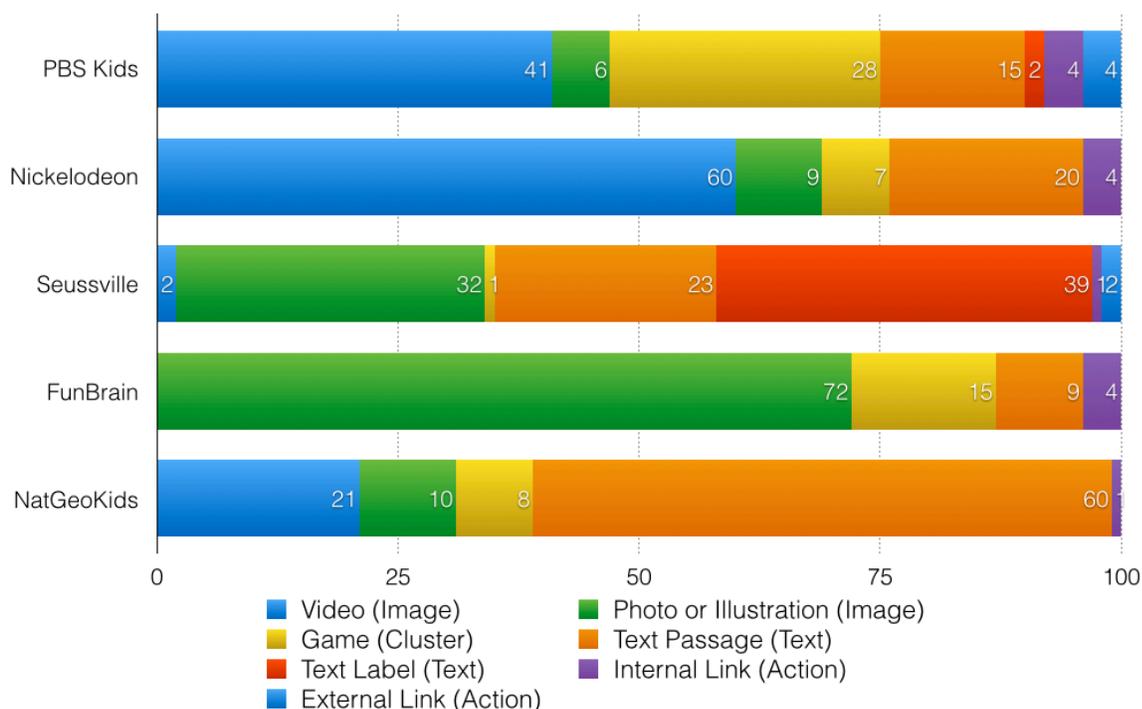


FIGURE 4.1. Percentage of pages by primary element type on five sites.

Although many pages showed a clear dominance of one mode of communication over others, multimodality was the rule. Pages that were primarily text also included logos, colors, and internal and external links. Pages dominated by images included text labels and text passages. The presence of multiple modes was documented on all pages, on all sites. The websites examined, although varied, were richly, thoroughly multimodal.

From a formal perspective, PBS Kids and Nickelodeon resembled each other. Both included a large number of short video clips and full episodes from televised series. Both included games themed around the TV series. However, the apparent similarity in reading levels (7.1 and 7.2) did not reflect the actual literacy differences on the sites. Many of the properties on PBS Kids, such as “Super Why” (<http://pbskids.org/superwhy/>), “Dinosaur Train” (<http://pbskids.org/dinosaurtrain/>), and “Sesame Street” (<http://pbskids.org/sesame/>) did not contain reading passages at all. Most text passages found on PBS Kids are either written for parents and teachers or concentrated on a few properties, such as “Wilson and Ditch” (<http://pbskids.org/wilsonandditch/>) and “Wild Kratts” (<http://pbskids.org/wildkratts/>).



FIGURE 4.2. On PBS Kids, an extended informational text reading passage. Aids to text comprehension include videos, maps, pictures, and diagrams. Reprinted from “Wild Kratts Grey Squirrel”. (2014). PBS Kids. Retrieved from <http://pbskids.org/wildkratts/creaturepedia/>.

An example of an extended text passage available on the PBS Kids website can be found in the “Creaturepedia” section of the “Wild Kratts” site. “Creaturepedia” provides informational text about a variety of animals. The entry for “gray squirrel” (see figure 4.2) includes 246 words under five headings. The text has a Flesch-Kincaid Grade Level score of 5.4, high syntactic simplicity (70%), word concreteness (98%), and referential cohesion (68%). These scores

indicate that the text employs simple sentence structures and easily-visualized vocabulary.

Referential cohesion is a measure of the “presence of overlapping ideas and concepts in a text” (Coh-Metrix, 2014). Text comprehension is supported by two short videos, a world map showing gray squirrel habitat, a photograph, and an illustration.



FIGURE 4.3. On PBS Kids, a brief and decontextualized scientist biography. Reprinted from “PBS Kids Sci Girls”. (2014). PBS Kids. Retrieved from <http://pbskids.org/scigirls/index/>.

Another example of a text passage can be found on the “SciGirls” site (<http://pbskids.org/scigirls/index/>). The home page for SciGirls, a property focused on increasing the interest of girls in science, math, and engineering, is based on a visual metaphor of a scrapbook montage. The page (see figure 4.3) includes photographs, illustrations, animations, video, text labels, and several different types of clusters. Clicking a photo of Jane Goodall reveals a brief bio:

Jane Goodall. Born: 1934 in London, England. When this pioneering primatologist began her career, British officials refused to let a young woman venture alone into the African jungle. Problem solved: she took her mom! Dr. Jane’s work with chimpanzees is now lauded worldwide, as is her current environmental focus.

This 49-word biography has a Flesch-Kincaid Grade Level score of 8.4, somewhat complex sentence structure (syntactic simplicity=38%), and a word concreteness score of 62%. The passage has a low level of cohesion (3% referential cohesion, 10% deep cohesion), which may be due in part to the shortness of the passage. Comprehension of the text may be hindered by other formal characteristics of the presentation. The text was not accompanied by any elements that reinforce the content presented. There were no links to other information about Jane Goodall, primatology, Africa, environmental science, or conservation. There were no photographs or illustrations of primates or “Dr. Jane” doing her work and no maps to help identify the location of her work. More problematic, while the text was displayed, several overlapping audio clips played without user control: a sixty second music loop; a buzzing sound from an animated cell phone at the bottom of the page played every 7 seconds; and a voice recording repeated every 18 seconds that said, “Move your mouse around to check out the home page.” The formal arrangement of this page, from the lack of contextual support for the text to the chaotic visual and auditory environment, seemed to counter comprehension of this text passage.

Throughout PBS Kids, spoken text labels activated by mouse-over were present on twenty-six of the forty-one properties. This feature allows a user who cannot read alphabetic text to hear the text label spoken. Because the spoken text responds to the user’s pointer location regardless of other sounds, this feature can cause overlapping audio. For example, if the user moved the pointer quickly across a series text labels, each text label was spoken simultaneously, resulting in a cacophonous cascade of navigation help.

Question 3: Substance of Selected Websites

Nickelodeon, PBS Kids, and Seussville were organized in a somewhat similar fashion: a central corporate hub with multiple properties⁴, (shows, characters, or brands). Nickelodeon had two content organization strategies. The primary strategy had two levels. First, content was organized around by property, with most properties accessible through the “green slime” menu available at the top of every page. On each property page, the second level primary navigation by media type (games, videos, episodes, message boards, pictures, property description, and character biographies) appeared at the top right side of the screen (see figure 4.4). Secondary navigation for the entire site was organized by media type (games, videos, episodes, radio). From the top left of every page, the media type navigation menu provides access to content from all properties. Despite the extremely consistent parallel construction of the Nickelodeon property pages, great variety existed in the amount of subject matter related to each property, as described in figure 4.5.

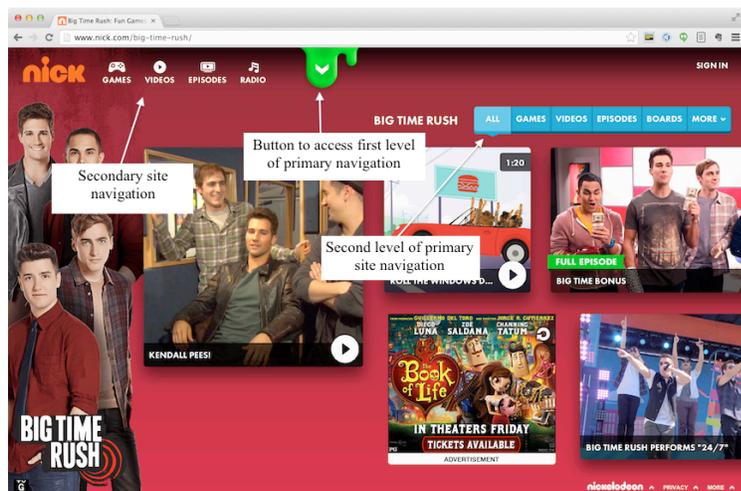


FIGURE 4.4. Typical arrangement of a property home page on nick.com. Reprinted from “Big Time Rush”. (2014). Nickelodeon. Retrieved from <http://www.nick.com/big-time-rush/>.

⁴ In this context, the term “property,” referring to the concept of intellectual or creative property, is useful to encompass all aspects of each separate program. For example, “Sid the Science Kid” is the name of the TV show, website, and a character featured in games, videos, and activities. “Property” is inclusive of all creative content associated with “Sid the Science Kid.”

(<http://pbskids.org/help/>), TV schedules (<http://pbskids.org/tvschedules/stationfinder.html>),
 Writer's Contest (<http://pbskids.org/writerscontest/>), pre-school menu page
 (<http://pbskids.org/hooper/>), and the Cartoon Studio activity (<http://pbskids.org/go/studio/>).

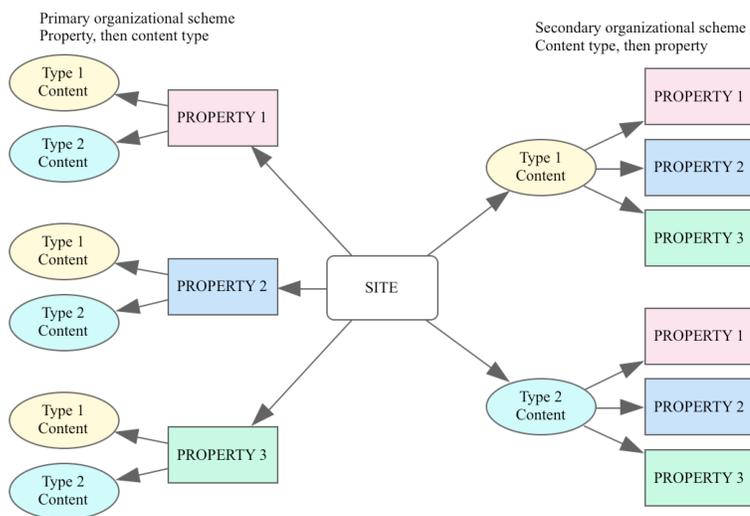


FIGURE 4.6. Organizational pattern found on three of five sites.

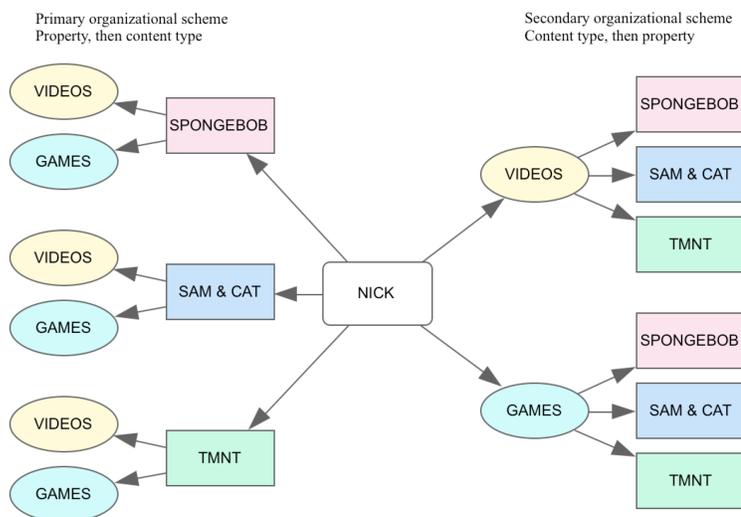


FIGURE 4.7. Organizational pattern illustrated with examples from Nickelodeon.

The organizational patterns of Nickelodeon, PBS Kids, and to a lesser extent, Seussville, followed the same scheme depicted in figures 4.6 and 4.7. These diagrams demonstrate a

possible instantiation of multilinearity on children's websites by documenting multimedia traversals (Lemke, 2005) that are available to readers.

The most common subjects on National Geographic Kids were animals, geography, conservation, science, and scientists. The most common subjects on PBS Kids were learning, friendship, creativity, feelings, and enjoyment of learning. The most common subjects on Seussville were creativity, social skills, and tolerance of difference. The most common subjects on Nickelodeon were teamwork, tolerance of difference, trust, personal relationships, and fighting for ones beliefs. The most common subjects on FunBrain were fun, humor, and facing fears. Although many properties on PBS Kids had a stated goal of improving reading skills, almost no text passages exist at a primary grade level.

Searching, browsing, and sharing. Search features have been ubiquitous on websites for years (Saeed, 2011), so it is interesting to note that Nickelodeon's redesigned site has no search feature. Since this was a complete redesign of an industry-leading site (Steel, 2014), one can assume that removing search was not an oversight and it was not born of a disregard for user's needs. Without a search feature, the user must browse for content on a site that includes nearly 6,000 pages. On Nickelodeon, it is very easy to find a video featuring characters from a given property. It is very difficult to find one specific video because the content on a given page changes constantly. Perhaps this change is meant to encourage users to spend more time browsing, discovering new content, being exposed to advertisements. The change may instead be related to changes in user habits. The PBS Kids website included a search feature, but it does not appear on the home page and ironically it was difficult to locate. National Geographic Kids featured a search box at the top of its home page. FunBrain and Seussville provide indexes and directories for content, but no search features. The lack of search may be a common feature of

web design for children, given that usability studies have found that children use search features less than adults. Of the websites examined, Seussville offered the ability to share content via various social media services, while National Geographic Kids, FunBrain, Nickelodeon, and PBS Kids did not. For the three sites with significant amounts of video content (Nickelodeon, PBS Kids, and National Geographic Kids), only Nickelodeon's site generated a unique URL for each video that could be copied or bookmarked by the user. Video content on National Geographic Kids and PBS Kids displayed within a video player with a common URL.

Question 4: Rhetorical Action of Selected Websites.

Table 4.2 summarizes several rhetorical dimensions of the five websites studied and is followed by a narrative description of these findings.

All of the sites except Nickelodeon had an implied or stated education purpose. PBS Kids had a clearly-stated and comprehensive mission statement published on their website (<http://pbskids.org/help/faq.html#mission>):

PBS KIDS is committed to making a positive impact on the lives of children through curriculum-based entertainment with positive role models and content designed to nurture a child's total well-being. With a 360-degree approach towards learning and reaching children, PBS KIDS leverages the full spectrum of media and technology to build knowledge, critical thinking, imagination and curiosity. PBS KIDS encourages children to interact as respectful citizens in a diverse society. By involving parents, teachers, caregivers and communities as learning partners, PBS KIDS helps to empower children for success in school and in life. PBS' bottom line is measured by how much it contributes to the welfare of America's children.

TABLE 4.2
Rhetorical Characteristics of Selected Websites.

	PBS Kids	Nick	Seussville	FunBrain	Nat Geo Kids
Stated Purpose	Build knowledge, critical thinking, imagination, curiosity in children	None	Support creativity and discovery of Dr. Seuss's work	Providing a safe environment that bridges learning and entertainment	None
Context/Community	Public media, Television	Commercial television	Commercial book publishing	Commercial education publishing	Non-profit education, magazine, television network
Stated Audience	Children, parents, teachers, caregivers	None	Children, parents, teachers	Parents, librarians, teachers	None
Implied Audience	None	Children and adolescents	Adults who purchase books	Children	Children, parents, teachers
Prior knowledge	PBS Kids Characters	Nick Characters	Dr. Seuss, Characters	None	None
Treatment of Authorship	Authors not identified	Authors not identified	Author-focused website	Book and comic authors identified	Full creative team listing; scientist bios
Use of external references	US Dept. of Educ., university researchers, teachers, licensed properties, sister sites	Licensed properties, Sister properties, Corporate owner	Publishing house, Corporate owner	Sister sites, Corporate owner, safety and quality endorsements	TV channels, magazine
Advertising	Apps, books, and related products	Yes	Books and related products	Yes	Yes

PBS's children's media and family and educator resources include PBS KIDS television series, pbskids.org, PBS Parents, PBS Teachers, PBS KIDS Raising Readers and literacy events across the country. PBS is a nonprofit media enterprise owned and operated by the nation's 356 public television stations.

Throughout their materials, PBS Kids aligned itself with parents and teachers in preparing children for school, employing phrases like "ready to learn" and providing extensive resources and suggested activities for parents and teachers. PBS Kids' sister site PBS Parents was described this way: "Parenting resources and tips on raising children, planning birthday parties & kids activities" (PBS, 2014). Providing parents with tips on raising children goes beyond the role of simply providing television programming. External references to federal education initiatives and university research bolstered a rhetorical stance of expertise, responsibility, and professionalism.

FunBrain's "About" page (<http://www.funbrain.com/aboutus.html>) emphasized messages of safety, trust, and confidence, specifically and repeatedly naming teachers, librarians, and parents as the intended recipients of these messages. FunBrain also provided grade level and subject area indices to all games.

National Geographic Kids did not state an audience or an education mission, although their home page included the text, "Play games, watch videos, learn about animals, and places, and get fun facts on the National Geographic Kids website" (National Geographic Kids, 2014). National Geographic was alone among the websites studied in identifying the entire creative and technical team involved in producing the website. In addition, the brief scientist biographies

included in written articles and videos are very similar in content and style to brief author biographies and character biographies from other sites.

Some rhetorical elements on Seussville indicated children as the intended audience, but others identified parents or other adults. For example, the games and coloring pages were clearly intended for children. The videos seemed to be aimed at adults. The videos present on the site appear to be for an adult audience, employing language like, “The Lorax Pop Up! This bestselling ecological warning is now available in an elaborate pop-up book, published in conjunction with the release of The Lorax feature film on March 2, 2012 – Ted Geisel's birthday.” On Seussville, no page was more than a few clicks away from buying a book (see figure 4.8). The extensive character menu served as a mechanism to locate the 61 activities or the 293 books available for purchase.

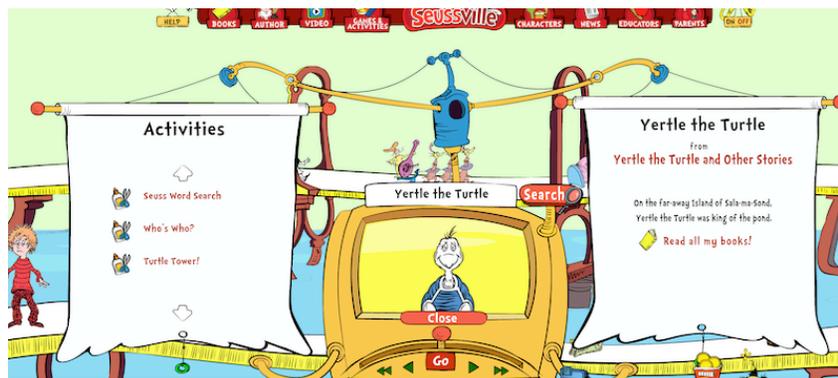


FIGURE 4.8: Seussville character menu displaying links to activities and books to purchase. Reprinted from “Characters Dr. Seuss Seussville”. (2014). Seussville. Retrieved from <http://www.seussville.com/?home#/characters>.

Most commercial activity on PBS Kids was limited to the SHOP and to the APPS section. However, unlike the SHOP link, the APPS link did not alert the user to a transition in content. The APPS & MORE section included paid content and a limited amount of free content. The paid content was sold via external links to the Apple App Store, Amazon, the Google Play Store, the Barnes & Noble Nook Store, the Windows Store, and the Windows Phone Store,

depending on the app, album, or e-book format. In addition, some content of the Parent's headband section may have been considered commercial in nature. Headband content provided brief text labels and hyperlinks intended for parents to access additional information and support materials. The hyperlinked pages include paid apps.

PBS Kids did not engage in cross-property promotion (including characters from more than one property in promotional materials) or cross-marketing of programs (including content segments from one program in the programming stream of another program). As discussed earlier, PBS Kids was a portal site connecting to at least forty-one property websites that appear to be produced and managed more or less independently. The size, depth of content, and frequency with which property websites are updated varies greatly. Based on information provided for parents and teachers, each property website targeted a narrow age range between 2 and 10 years. Authors and other creative and technical staff were not identified on the PBS Kids website, however, the names of the production companies responsible for each property were listed with contact information (<http://pbskids.org/help/write.html>).



FIGURE 4.9. An example of cross-property promotion. Reprinted from “Nick’s All-Time Best Dance Moves”. (2014). Nickelodeon. Retrieved from <http://www.nick.com/pictures/nick-digital-nicks-best-dance-moves-pictures.html>.

Cross-promotion. Nickelodeon regularly engaged in cross-property promotion and cross-marketing of programs. An example of cross-property promotion was identified as “Nick’s All-Time Best Dance Moves,” see figure 4.9. The photo set includes images of dance moves from multiple Nickelodeon series. The photo set was used in the main content stream on the Nickelodeon home page and also appeared in the property feeds for the properties represented in the pictures. Figure 4.10 depicts an example of cross-marketing. A content item for the property “Big Time Rush” appeared in the content stream on the property home page for “iCarly.” This is significant because the content stream appears to be dynamically-generated from a database of content (Rumianek, 2013). Presumably a database query for current “iCarly” content would result in exactly and solely “iCarly” content. The fact that content pertaining to other properties appears in a given content stream – in a way, a “broken” result from the database – suggests a

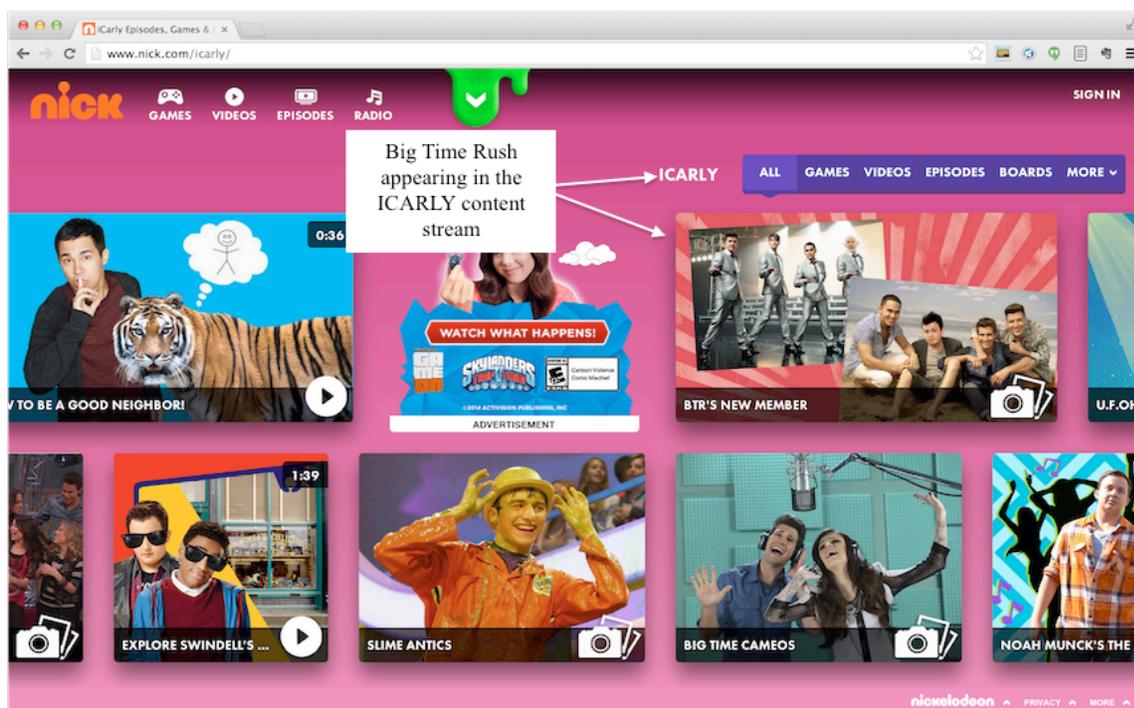


FIGURE 4.10. Cross-marketing of properties: Big Time Rush promoted in the iCarly feed. Reprinted from “iCarly Episodes, Games & Pictures Nickelodeon”. (2014). Nickelodeon. Retrieved from <http://www.nick.com/icarly/>.

strategy of cross-promotion, not unlike a television commercial for one television show running during commercial breaks for another television show on the same network. However, in this case, the network has a content stream, a channel, for each property. In addition to a Nickelodeon “channel” that mixes content from each property, Nick has fifty channels running property-specific programming (and cross-promotions) entirely on-demand. With more than 3,000 pieces of video content alone at the time of this study, and with advertising interspersed, it is untenable to think of a website like nick.com as the “online home” of a television network. Websites like this demonstrate the evolution of old media companies into new media companies and give us a glimpse of what is possible.

Comments and message boards. Some individual property sites on PBS Kids allowed users to post comments, art, curate digital sticker books, collect screen captures, or create some form of customized user home page. National Geographic Kids encouraged users to join their “My Shot” project to post and share photographs in a moderated forum. Seussville and FunBrain both allowed users to create simple, personally non-identifiable accounts so that users could save progress on games. Message boards are electronic bulletin boards that on which website users can create conversation threads associated with various properties. Nickelodeon had 710 pages of message board content, ostensibly created by website users. The majority of these posts were expressing opinions about shows or characters. Just over 30% of the posts contained some form of fan fiction. In addition, users on Nickelodeon can post comments on virtually any individual piece of content. All five sites had privacy policies in place to restrict the amount of personal information users shared online. Table 4.3 summarizes options available for online participation and highlights some differences in the kind of information solicited by the websites.

TABLE 4.3
Participation options on selected websites.

	PBS Kids	Nick	Seuss-ville	Fun Brain	Nat Geo Kids
Create a displayed user name	Yes	Yes	No	No	Yes
Asks for child's date of birth	No	No	No	Grade	Yes
Asks for email address	No	No	No	No	Parent
Asks for child's gender	No	No	No	Yes	Yes
Connection to social media	No	No	Yes	No	No
Comments/message boards	Yes	Yes	No	No	Yes

Question 5: Evidence of Genre.

As previously stated, I see genre as a “stabilized-for-now or stabilized-enough sites of social and rhetorical action” (Schryer, 1993, p. 204). Although genre has formal and substantive aspects, rhetorical action is the key to understanding a genre. However, by nature, formal and substantive aspects are easier to document than rhetorical action when a study takes place, as this one does, in the absence of rhetors. This study was based on the artifacts, the tracings, of genre, in which form and substance are more clearly visible. The evidence collected in this study supported the assertion that these five websites do not participate in identical genres.

PBS Kids and Nickelodeon share several characteristics. Both were websites produced for children by companies that also produce television networks. Both were primarily organized around a finite set of creative properties. Both included large numbers of short video clips and more limited numbers of full-length episodes that originally aired on television. Both provided methods for children to participate online. Both included games themed to their properties.

The evidence suggests that PBS Kids is more de-centralized than Nickelodeon, functioning more as a portal to a family of websites than as a single website, although it retains a strong central identity. Most of Nickelodeon's properties are geared toward older children than

most of PBS Kids' properties. While PBS Kids has a clear and ideologically-articulated mission to support education, Nickelodeon does not state a mission. PBS Kids placed a strong emphasis on educating and enlisting parents and other caregivers to improve education of children.

Nickelodeon addresses parent concerns specifically about children's safety and privacy while on nick.com.

Three of the sites, PBS Kids, Nickelodeon, and Seussville, were organized around narrative, character-based creative properties. An argument could be made that FunBrain was also organized around character-based creative properties because of the large proportion of pages dedicated to electronic books and comics. However, on the former three sites all site content was tethered to those character-based creative properties. In addition to and separate from its books, FunBrain also hosts a large proportion of games, none of which are tied to its narrative properties. Seussville differs from PBS Kids and Nickelodeon in that Seussville does not have a large proportion of content meant to be consumed on the site. The largest proportion of pages on Seussville was book detail pages (39%), essentially catalog entries, followed by printable activity sheets (23%).

FunBrain, Seussville, and National Geographic Kids were all connected to the print-based publishing, although FunBrain's connection came later in its life. All five sites are now owned by "old media" companies.

The evidence collected in this content analysis was sufficient to make tentative claims about some possible genres in which these children's websites participate. PBS Kids and Nickelodeon participate in separate, but closely related, genres. Tentatively, we could say that these are two different varieties of a genre that I will call "Pop Culture Carousel."

I derive the word “carousel” from a web design term for the scrolling menu system found at the tops of the pages on both of these sites, and from the metaphorical amusement ride it references. The designers responsible for these sites likely employ carousels for navigation because they organize their sites around a number of properties large enough that they can’t easily or effectively be displayed in a static row. “Pop culture” is based on the idea that the content on each of those sites participates in, or creates, or contributes to, or borrows from, popular culture. This is significant to the definition of the genre because of the potential for rhetorical action based on involvement with pop culture.

PBS Kids is a Pop Culture Carousel with younger children as a target audience and an expressed education mission. Nickelodeon is a Pop Culture Carousel with older children as a target audience and an implied entertainment mission.

Tentatively, a Pop Culture Carousel is a website genre with children and/or adolescents as one of its primary intended audiences, organized around a large number of narrative properties, probably produced by different creative teams. Each narrative property may differ greatly from other properties. Despite the differences among the constituent properties, a Pop Culture Carousel has a consistent central identity that holds the group together and prevents it from becoming a group of loosely affiliated separate websites.

The data suggest that the three remaining websites do not participate in the Pop Culture Carousel genre. Further, primarily based on the substantive and formal data, I would suggest those three sites, FunBrain, Seussville, and National Geographic Kids, do not share a common genre.

Further study of many more children’s websites and direct studies of website creators and of children reading these and other websites is necessary before making any further claims

regarding genre. Many possible examples of the proposed genre need to be examined in order to determine the essential and optional characteristics of the genre (Miller & Shepherd, 2004).

Summary

This content analysis sought to identify methods useful in studying children's websites and identifying genre in those spaces. I employed many different instruments and techniques to examine multilinear, multimodal, dynamic website content. I refined a replicable process for analyzing websites. I used purposeful sampling to identify five websites for study and applied and further refined my process.

I believed this study was needed because children read, write, listen, watch, and play on children's websites every day. The environments they explore are far too diverse to be described by the single term "children's website." We would never think of applying the same kind of blanket term to print-based media, regarding anything between two covers as simply a book without differentiating between an activity book, a chapter book, a picture book, or even a cookbook. Each of those terms is more useful than simply "book," but still not specific enough to be truly useful. Identifying specific genres is extremely important studying literacy in online spaces.

The technology of web design is constantly changing. Three of the five websites included in this study were completely redesigned within the past year. Two were completely redesigned during the period of data collection and analysis. One set of tools will not work for all websites. Rather, a researcher requires a flexible set of tools that can be employed strategically when confronting different challenges in accessing and documenting website content.

I produced rich descriptions of the websites studied, employing indexical, formal, substantive, and rhetorical data. Based on these data, I offered a tentative identification of a

children's website genre. Pop Culture Carousel references pop culture, a site of significant rhetorical action, and an endless looping procession of interesting ideas held together by a strong central concept. Further research will be necessary to know whether or not this new genre is a useful concept.

Chapter Five: Conclusions

“I think the increasing importance of visual modes of communication and multimodal modes of communication where modes are working with each other needs to sort of tear us away from a devotion to the text and the word” Carolyn Miller.

Summary

Children spend an increasing amount of time on the Internet reading, writing, listening, viewing, and playing (Snyder & Dillow, 2011). Reading, already a complex cognitive task, is even more complex in online environments (Coiro, 2003; D. J. Leu & Kinzer, 2000). Online texts involve multiple possible paths through a text, multiple means for navigation, an ever-changing corpus, and a functionally infinite amount of text (Duke, 2012). In online reading environments, it is more difficult for readers to ascertain the author of a given work and more difficult to separate advertising from content (Coiro & Dobler, 2007). Additionally, online texts have their own unique set of genres that are richly multimodal (Miller & Shepherd, 2004; Ratliff, 2003). Multimodal refers to the co-deployment of multiple modes in communication. Modes include written text, image, audio, video, color, gesture, and speech (Kress, 2003a, 2009).

Although many online genres have been studied (Dillon & Gushrowski, 2000; Herring, Scheidt, Wright, & Bonus, 2005; Miller & Shepherd, 2004; Puschmann, 2009; Ratliff, 2003; Ihlström & Henfridsson, 2005; So, 2005; Wise, Bolls, & Schaefer, 2008), no prior study has examined an online genre geared toward children. Researchers who have investigated children reading online have confined their work to informational texts read for academic purposes (Cho, 2011; Coiro, 2007; Coiro & Dobler, 2007; Henry, 2007; Donald J. Leu et al., 2005; Zhang &

Duke, 2008). Although this is important work that directly informs classroom practice, it excludes the majority of what children read online in the spaces they choose to visit. This study is a precursor to further study in an area that is largely unexamined.

The purpose of this study was to establish a comprehensive methodology for analyzing children's website content. By applying genre theory through a mixed methods content analysis, this study addresses gaps between research into online genres and research into children's reading on the Internet. This research will support new pedagogical and analytic approaches to children's online reading behavior and experience.

I began with a convenience sample fifty web pages and fifty web page elements. I used them to establish procedures, investigate techniques, and create tools, such as the Web Page Data Evaluation Checklist (WPDEC) and the Element Data Evaluation Checklist (EDEC) to assist in web page content analysis. I then collected a purposeful sample of five popular children's websites and applied the tools and techniques I established.

My analysis revealed patterns of similarity and difference between and among the five selected sites. Based on formal, substantive, and rhetorical data, I have posited a previously unidentified website genre. Further research is necessary to verify the usefulness of this description.

Conclusions

Establishing meaningful genres. Results of the analysis strongly suggest that the term "children's websites" does not refer to a genre. Each of the five websites studied could be called a "children's website," but they differ greatly from each other in terms of form and substance. Evidence suggests that they also differ in terms of the social and rhetorical actions accomplished

by their use. These differences provide a firm basis from which to argue that more than one genre is present in the sample.

Further, the results of this analysis suggest that grouping websites into meaningful genres based on formal, substantive, and rhetorical data is possible and would inform further research into children's reading and writing development in online spaces. In addition, establishing meaningful genres can inform classroom instruction by making better use of exemplar texts.

Documenting multilinearity and multimodality. Although the concepts of multilinearity and multimodality have been described and associated with Internet texts, no previous study has documented them as extensively as this content analysis. Genres change over time. By documenting the instantiations of multilinearity and multimodality over time, it will be possible to learn how these important dimensions shift in new literacy contexts. The study of reading in online spaces is not truly possible until we completely describe online spaces. Complete descriptions are necessary in order to know what to study and how to study it.

The current study provided a basis from which to study instability of children's websites and a methodology that supports such inquiry. Three websites included in this study were completely redesigned during the preceding 18 months. One redesign was launched between the creation of the instruments and the beginning of data collection. The shift from a synchronic perspective to diachronic perspective (Saussure, 1983) will illuminate the characteristic of instability on individual websites, but more importantly will allow descriptions of the instability of genres of children's websites. How frequently do children's websites change? Which genres change most frequently and in what ways? What are the intended and unintended consequences of changes on children's websites? How does the volume of available content change over time? The current study provides a necessary platform from which to address questions such as these.

Given that all inquiry has an ideological and theoretical perspective regardless of whether or not it is articulated, the procedures described in this study are not expressly and intentionally ideological. Instead, a theoretical lens can be applied to the analysis of data collected using these procedures. A researcher could for example conduct a Marxist analysis, or examine data from a constructivist perspective, or take a historical perspective.

Rhetorically-based classification. According to Carolyn Miller, when approaching genre identification "...we need to be sensitive to the need not to impose pre-fabricated theoretical frameworks on them, without at the same time jettisoning explanatory frameworks that have been useful in the past" (transcribed by author from Association for the Rhetoric of Science and Technology, 2013). When the study of a genre forces data to conform to prior understandings, we risk reinforcing the "tiresome and useless taxonomies" of which historian of rhetoric Thomas Conley warned (as quoted in Miller, 1984, p. 155). The system of classification must be open and rhetorically-based. In the current study, the list of variable values for substance and for rhetorical action were derived from the text base and allowed for continual modification to account for the actual content on the websites.

Recommendations

Buried treasure. New information is added to the Internet at staggering rates. Resources that can be very valuable for classrooms can be lost in the deluge. Locating a single site that provides a useful resource for a given teaching context is not as powerful as understanding the ways in which that single site is similar to an entire class of websites. A single website may change, a resource may disappear, but a genre will change more slowly because genres are necessarily more stable than individual texts (Miller, 1984). By identifying genres of websites, we can better understand how to locate and make use of resources that exist and new resources

that become available. Many teachers already make intuitive use of genres to differentiate between different types of websites and the resources they are likely to provide. A more complete and nuanced understanding of online genres will aid in discovery of resources.

For example, when teaching creative writing, a “pop culture carousel” site is likely to be a rich source of examples of character, conflict, setting, and other aspects of storytelling. Moreover, the examples are parsed into clips of a few minutes in length which makes them easier to integrate into a larger lesson. Video clips from pop culture carousel sites also involve contexts that are meaningful and engaging to school-age children.

Knowing what changes; understanding why. When nick.com launched its redesigned site in July 2014, user comments, which had been featured on the home page, disappeared. In fact, user comments are displayed in a different way, but not featured in a “Twitter” style feed on the home page in the current design. In the previous design, even casual visitors to the site could read comments from other users, presumably children⁵, before they engaged with any content provided by the site authors. Comments are still available, but they are not displayed in an aggregated feed and they must be accessed by first clicking on individual items of content.

There is a rhetorical dimension to that change. In one sense, children now have less direct access to an audience for their comments. Whereas the previous design dedicated a significant amount of home page screen space to the current online activity of other users, the current design emphasizes the site author’s priorities for attention. However, the site authors created the previous design and controlled the content that was displayed, so the messages, even if typed by children, were expressing the message and perspective of the authors. Perhaps the change was in response to concerns about privacy or perhaps it was related to controlling focus. While

⁵ The relationship between the actual identities and the adopted personas of users who comment is necessarily unknown from an examination of the text, in the interest of protecting the privacy of children. Studies of children creating online personas provide a window into this process.

speculation is possible based on analyzing the content, direct research with website users and authors would be necessary in order to fully understand the changes. Regardless, the changes cannot be studied if the changes are not noticed and documented. This change and others like it are examples of the instability of online texts.

Although projects like the Internet Archive and the International Internet Preservation Consortium (IIPC) are valuable, the technological limitations of web archiving are such that much useful information reveals itself to be essentially ephemeral when websites change. Information about the context in which information is displayed is particularly vulnerable. In the example above, the text of user comments may be captured by web archiving technology and preserved, but the rhetorical choices that the site author makes about how to display that text in relation to other content may be lost. For this reason, the documentation and study of websites used by children should be an ongoing project of literacy researchers.

SpongeBob is lost in time. SpongeBob is arguably the most famous character on the Nickelodeon site. Based on the frequency with which he appears in cross-promotional contexts, SpongeBob occupies a place of central importance to the Nickelodeon brand. By reading the character bio on the site, you can discover many things about SpongeBob (all written in the present tense), but nothing about his creation. SpongeBob, like virtually all properties on the Nickelodeon site, has no creator and no history. No authors, illustrators, directors, producers, or any other creative or technical staff is listed on the Nickelodeon site. I suspect this is true of many similar sites. Within the websites examined in this study, only National Geographic Kids provided a comprehensive list of creative and technical staff⁶.

⁶ In children's literature, science information text should be written by content experts and include references to research. National Geographic Kids' list of credits may exist for the same reasons. Likewise, children's literature includes a tradition of narrative texts authored by syndicate "ghost" writers. A prominent example is the Stratemeyer syndicate, publisher of the *Nancy Drew*, *Hardy Boys*, and *Bosby Twins* book series, among others.

Books generally include the author's name on the cover. A reader sees the name every time he or she picks up the book. A TV show lists the creative team at the end and sometimes at the beginning of an episode, ostensibly available to every viewer, every time an episode is viewed. In fact, the only place Nickelodeon creative and technical staff are identified is in the closing credits at the end of full episode of a television show. As such, the text is embedded in the video and not searchable on the site and not indexed by search engines.

Since creators are listed in episode credits, it is possible although very difficult to know who created the character. There is no information about when he was created. An episode could be 6 months old or 6 years old; there is no information. Of course, complete details of every episode of every show can be found on other websites, such as Wikipedia.

What are the implications of having creative content on websites has no 'author'? It's as if these texts spring forth in final form, author-less. One powerful aspect of listing the author and/or illustrator on the cover of a book is that the reader is confronted with the concept that the text was authored, that the character was illustrated. The extension is that the role of writer or illustrator is available to the reader. Child readers in school can conduct author studies, even going so far as to correspond with the author and learn directly about the creative process. The lack of authors on some websites leaves no author position open for the children who read there.

The lack of dates, growth, and history for a character like SpongeBob may have to do with keeping the character current and relevant to a continually renewed audience. The owners of a brand have a market-driven interest in keeping the brand from seeming old.

I grew up watching Looney Toons cartoons and I remember being surprised later in life when I discovered that some of my favorite episodes were decades old when I first experienced them. As a five year old, I was a dedicated fan of the television series *Gilligan's Island*, also

created before I was born. These creative properties were relevant and current to me as a child. The character of SpongeBob exists in a timeless condition, possibly for the same reasons that the character of Gilligan did. This notion of timelessness may be linked to the absent author. When an author is identified, the discourse moves into the past tense. The text would change from “SpongeBob is...” to “SpongeBob was created...” The acknowledgement of creation ties the character to a timeline, acknowledges that there was a time before the character.

Website content related to the SpongeBob character differs from content experienced via other media in the treatment of time and authorship. Further research into the ways in which notions of authorship change in online spaces is warranted. The methodology established in the current study supports rhetorical analysis of website content.

Additionally, lack of authorship may be related to the concept of intertextuality. Prior to the Renaissance, the concept of individual authorship and the origination of ideas did not carry the same importance as they do today (Chandler, 2014; Goldschmidt, 1943). Perhaps as in Medieval texts, the creativity or individual expression of a children’s website author or artist is subordinate to the pre-existing concepts and conventions of the genre. In this sense, intertextuality may define the text to a greater extent than the agency of any individual creator. Regardless, further study of the role of authorship on children’s websites is warranted.

Eye-tracking studies. Comprehensive, multimodal content analysis may be of particular importance to researchers applying eye-tracking data to literacy questions. Because of technological advances in recent years, researchers are able to collect much more accurate data about the position of a reader's or viewer’s gaze. In the field of literacy, this has led to calls for revisions to some longstanding models of reading and writing processes (Anson & Schwegler, 2012). While large scale eye-tracking studies of reading have yet to be conducted, the approach

has the potential to add greatly to our understanding of reading and writing. However, tracking the position of the eyes is only one side of the equation. Eye-tracking studies need to distinguish between task-relevant, task-redundant, and extraneous areas within a visual display (Gegenfurtner, Lehtinen, & Säljö, 2010). The validity of their findings depends upon these distinctions. The methodology of the current study provides a systematic and reliable way to determine significance and meaning within a complex visual display that includes alphabetic text, images, and content.

Implications for instructional design. Many of the webpages examined in this study dense with overlapping concepts represented in multiple modes. These websites are extremely information-rich environments. The information is highly culturally-contextualized and richly multi-modal. The environment is complex and dynamic. Children choose to spend time on these spaces and apparently read and write in these spaces. By studying how children learn in authentic web spaces that they choose to visit, we may learn new approaches to designing learning spaces for children. Assuming that children can become accustomed to reading and viewing in virtual spaces as dense and rich as these websites (an assumption that should be fully investigated), more research is needed into difficulties that may be presented by on the comparative dearth of stimulation in most academic environments.

References

- Afflerbach, P. (1990). The influence of prior knowledge and text genre on readers' prediction strategies. *Journal of Literacy Research*, 22(2), 131-148. doi: 10.1080/10862969009547700
- AIGA (2003). *The orange book: Nickelodeon employee branding book*. Retrieved on 10/6/2014 from http://designarchives.aiga.org/#/entries/%2Bid%3A651/_/detail/relevance/asc/0/7/651/the-orange-book-nickelodeon-employee-branding-book/2.
- Altman, R. (1984). A semantic/syntactic approach to film genre. *Cinema Journal*(1225093), 6-18. doi: papers://9B5639B8-E52F-4416-81F5-893D7028B012/Paper/p8
- Altman, R. (1999). *Film/genre*. London: BFI Publishing.
- Alvy, L. M., & Calvert, S. L. (2008). Food marketing on popular children's web sites: A content analysis. *Journal of the American Dietetic Association*, 108(4), 710-713. doi: 10.1016/j.jada.2008.01.006
- American Library Association. (2014). ALA's List of Top Rated "Great Websites for Kids". Retrieved on 9/8/2014 from <http://gws.ala.org/>.
- Ant Ozok, A., & Salvendy, G. (2003). The effect of language inconsistency on performance and satisfaction in using the web: Results from three experiments. *Behaviour & Information Technology*, 22(3), 155-163. doi: 10.1080/0144929031000107081
- Aristotle, Buckley, T. W. A., & Hobbes, T. (1900). *Aristotle's treatise on rhetoric*. London: George Bell & sons.

- Association for the Rhetoric of Science and Technology. (2013). *ARST oral history project: An interview with Dr. Carolyn Miller*. Retrieved 4/5/2014 from https://www.youtube.com/watch?v=Pz_rN8g8WN0.
- Bakhtin, M. M., Holquist, M., & Emerson, C. (1986). *Speech genres and other late essays* (V. W. McGee, Trans.). Austin, TX: University of Texas Press.
- Baldry, A., & Thibault, P. J. (2006). *Multimodal transcription and text analysis*. London, UK: Equinox Publishing, Ltd.
- Bang, M. (2000). *Picture this: how pictures work*. San Francisco, CA: Chronicle Books.
- Barab, S. A., Bowdish, B. E., & Lawless, K. A. (1997). Hypermedia navigation: Profiles of hypermedia users. *Educational Technology Research and Development*, 45(12870469798480505584), 23-42. doi: [papers://9B5639B8-E52F-4416-81F5-893D7028B012/Paper/p1155](https://doi.org/10.1007/BF027028B012/Paper/p1155)
- Barab, S. A., Young, M. F., & Wang, J. (1999). The effects of navigational and generative activities in hypertext learning on problem solving and comprehension. *International Journal of Instructional Media*, 26(3), 283-301.
- Bazerman, C. (1980). A relationship between reading and writing: The conversational model. *College English*, 41(6), 656-661. doi: <http://www.jstor.org/stable/375913>
- Bazerman, C. (1988). *Shaping written knowledge: The genre and activity of the experimental article in science*. Madison, Wisconsin: The University of Wisconsin Press.
- Bazerman, C. (1997). The life of genre, the life in the classroom. In W. Bishop & H. Ostrum (Eds.), *Genre and writing* (pp. 19-26). Portsmouth, NH: Boynton/Cook.

- Bazerman, C. (2004). Speech acts, genres, and activity systems: How texts organize activity and people. In C. Bazerman & P. Prior (Eds.), *What writing does and how it does it: An introduction to analysis of text and textual practice* (pp. 309-339). Mahwah, NJ: Lawrence Erlbaum Associates.
- Bazerman, C., Bonini, A., & Figueiredo, D. b. (Eds.). (2009). *Genre in a changing world*. Fort Collins, Colorado: The WAC Clearinghouse.
- Bergman, M. K. (August 2001). "The deep web: Surfacing hidden value". *The Journal of Electronic Publishing* 7 (1). doi:10.3998/3336451.0007.104.
- Bezemer, J., & Kress, G. (2008). Writing in multimodal texts: A social semiotic account of designs for learning. *Written Communication*, 25(2), 166-195. doi: 10.1177/0741088307313177
- Biber, D. (1988). *Variation accross speech and writing*. Cambridge, UK: Cambridge University Press.
- Biber, D. (1995). *Dimensions of register variation: A cross-linguistic comparison*. Cambridge, UK: Cambridge University Press.
- Bilal, D. (2000). Children's use of the yahooligans! Web search engine: I. Cognitive, physical, and affective behaviors on fact-based search tasks. *Journal of the American Society for Information Science*, 51(5069329002400775922), 646-665. doi: papers://9B5639B8-E52F-4416-81F5-893D7028B012/Paper/p1157
- Bilal, D. (2001). Children's use of the yahooligans! Web search engine: Ii. Cognitive and physical behaviors on research tasks. *Journal of the American Society for Information Science*, 52, 118-136.
- Bitzer, L. F. (1968). The rhetorical situation. *Philosophy and Rhetoric*, 1, 1-14.

- Bordwell, D. (1989). *Making meaning: Inference and rhetoric in the interpretation of cinema*. Cambridge, MA: Harvard University Press.
- Bordwell, D. (2008). *Poetics of Cinema*. Berkeley: Routledge.
- Bourdieu, P. (1991). *Language and symbolic power: The economy of linguistic changes* (G. Raymond & M. Adamson, Trans.). Cambridge, MA: Harvard University Press.
- Busch, C., De Maret, P. S., Flynn, T., Kellum, R., Le, S., Meyers, B., Saunders, M., White, R., & Palmquist, M. (2012). *Writing@CSU: Content analysis*. Retrieved from <http://writing.colostate.edu/guides/guide.cfm?guideid=61>
- Caballero-Luque, A., Aragonés-Beltrán, P., García-Melón, M., & Dema-Pérez, C. (2010). Analysis of the alignment of company goals to web content using anp. *International Journal of Information Technology & Decision Making*, *09*(03), 419-436. doi: 10.1142/s0219622010003889
- Calisir, F., & Gurel, Z. (2003). Influence of text structure and prior knowledge of the learner on reading comprehension, browsing and perceived control. *Computers in Human Behavior*, *19*(2), 135-145.
- Callow, J. (2008). Show me: Principles for assessing students' visual literacy. *The Reading Teacher*, *61*(8), 616-626. doi: 10.1598/rt.61.8.3
- Carter, M., Ferzli, M., & Wiebe, E. (2004). Teaching genre to English first-language adults: A study of the laboratory report. *Research in the Teaching of English*, *38*(4), 395-419.
- Chafe, W. L. (1982). Integration and involvement in speaking, writing, and oral literature. In D. Tannen (Ed.), *Spoken and written language: Exploring orality and literacy* (pp. 35-53). Norwood, NJ: Ablex.

- Chandler-Olcott, K., & Mahar, D. (2003). " Tech-savviness" meets multiliteracies: Exploring adolescent girls' technology-mediated literacy practices. *Reading Research Quarterly*, 38(7861595479185802522), 356-385. doi: <http://www.jstor.org/stable/4151823>
- Chapman, M. L. (1995). The sociocognitive construction of written genres in first grade. *Research in the Teaching of English*, 29(2), 164-192.
- Cho, B.-Y. (2011). *Adolescents' constructively responsive reading strategy use in a critical internet reading task*. Dissertation.
- Clay, M. M. (1991). *Reading: The patterning of complex behaviour* (2nd ed.). Auckland, NZ: Heinemann.
- Common Sense Media. (2014). *Editor's picks: Best websites for kids*. Retrieved on 9/8/2014 from <https://www.commonsensemedia.org/website-lists>.
- Coiro, J. (2003). Exploring literacy on the Internet: Reading comprehension on the internet: Expanding our understanding of reading comprehension to encompass new literacies. *The Reading Teacher*, 56(5), 458-464.
- Coiro, J. (2007). *Exploring changes to reading comprehension on the Internet: Paradoxes and possibilities for diverse adolescent readers*. (Ph.D. Dissertation), University of Connecticut.
- Coiro, J. (2011). Predicting reading comprehension on the Internet: Contributions of offline reading skills, online reading skills, and prior knowledge. *Journal of Literacy Research*, 43(4), 352-392. doi: 10.1177/1086296x11421979
- Coiro, J., & Dobler, E. (2007). Exploring the online reading comprehension strategies used by sixth-grade skilled readers to search for and locate information on the Internet. *Reading Research Quarterly*, 42(2), 214-257. doi: 10.1598/rrq.42.2.2

- Cope, B., & Kalantzis, M. (1993). Introduction: How a genre approach to literacy can transform the way writing is taught. In B. Cope & M. Kalantzis (Eds.), *The powers of literacy: A genre approach to teaching writing*. Pittsburgh, PA: University of Pittsburgh Press.
- Cope, B., & Kalantzis, M. (Eds.). (2000). *Multiliteracies: Literacy learning and the design of social futures*. New York: Routledge.
- Creswell, J. W. (2007). *Qualitative inquiry and research method: Choosing among five approaches* (2nd ed.). Thousand Oaks, CA: Sage.
- Creswell, J. W., & Miller, D. L. (2000). Determining validity in qualitative inquiry. *Theory Into Practice*, 39(3), 124-130.
- Culham, R. (2003). *6 + 1 traits of writing: A complete guide*. New York, NY: Scholastic Professional Books.
- Culp, J., Bell, R. A., & Cassady, D. (2010). Characteristics of food industry web sites and "advergames" targeting children. *Journal of Nutrition Education and Behavior*, 42(3), 197-201. doi: 10.1016/j.jneb.2009.07.008
- Dancy, R. M. (2004). *Plato's introduction of forms*. Cambridge, UK: Cambridge University Press, 2004.
- de Sola Pool, I. (1959). *Trends in content analysis*. Urbana, IL: University of Illinois Press.
- Denzin, N. K. (1978). *The research act: A theoretical introduction to sociological methods*. New York, NY: McGraw Hill.
- Devitt, A. J. (1996). Genre, genres, and the teaching of genre. *College Composition and Communication*, 47(4), 605-615.
- Devitt, A. J. (2000). Integrating rhetorical and literary theories of genre. *College English*, 62(6), 696-718.

- Devitt, A. J. (2009). Re-fusing form in genre study. In J. Giltrow & D. Stein (Eds.), *Genres in the internet: Issues in the theory of genre*. Philadelphia, PA: John Benjamins Publishing Company.
- Dillman, D. (2000). *Constructing the questionnaire: Mail and internet surveys*. New York, NY: John Wiley & Sons.
- Dillon, A., & Gushrowski, B. A. (2000). Genres and the web: Is the personal home page the first uniquely digital genre? *Journal of the American Society for Information Science*, 51(2), 202-205. doi: 10.1002/(SICI)1097-4571(2000)51:2<202::AID-ASI11>3.0.CO;2-R
- Digital Trends. (2013). Best websites for kids. Retrieved on 9/8/2014 from <http://www.digitaltrends.com/how-to/ten-best-websites-for-kids/>.
- Dixon, J. (1967). *Growth through English*. Reading, England: National Association for the Teaching of English.
- Dresang, E. T. (1999). *Radical change: Books for youth in a digital age*. New York: H.W. Wilson Co.
- Emigh, W., & Herring, S. C. (2005). *Collaborative authoring on the web: A genre analysis of online encyclopedias*. Proceedings of the Thirty-Eighth Hawai'i International Conference on System Sciences. Retrieved from http://www.ufrgs.br/limc/escritacoletiva/pdf/online_encyclopedias.pdf.
- Eng, P.M. (1997). Cybergiants see the future—and it's Jack and Jill. *Business Week*. Retrieved on 10/6/2014 from <http://www.businessweek.com/1997/15/b352274.htm>.
- Fairchild, M. D. (2013). *Color appearance models / Mark D. Fairchild*. Chichester, West Sussex, UK: Wiley, 2013.

- Flowerdew, L. (2000). Using a genre-based framework to teach organizational structure in academic writing. *ELT Journal*(14150577533898724690). doi: papers://9B5639B8-E52F-4416-81F5-893D7028B012/Paper/p1040
- Fry, P. H. (2012). *Theory of literature*. New Haven, CT: Yale University Press.
- Funbrain. (2011). Press release: Funbrain launches newly redesigned website with fun math & reading games for kids. Retrieved on 9/01/2014 from <http://www.funbrain.com/press/20110810-Funbrain-Redesign-Launch.html>.
- Gage, J. (2001). *Colour and meaning: Art, science and symbolism*. Thames & Hudson.
- George, A. L. (2009). Propaganda analysis: A case study from World War II. In K. Krippendorff & M. A. Bock (Eds.), *The content analysis reader* (pp. 21-27). Thousand Oaks, CA: Sage.
- Gerodimos, R. (2008). Mobilising young citizens in the UK: A content analysis of youth and issue websites. *Information, Communication & Society*, 11(7), 964-988. doi: 10.1080/13691180802109014
- Gerson, L. (2013). Plotinus. In E. N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy* (Summer 2013 ed.).
- Giesbers, J., Verdonck-de Leeuw, I. M., van Zuuren, F. J., Kleverlaan, N., & van der Linden, M. H. (2010). Coping with parental cancer: Web-based peer support in children. [Research Support, Non-U.S. Gov't]. *Psychooncology*, 19(8), 887-892. doi: 10.1002/pon.1636
- Glass, G. V., & Hopkins, K. D. (1996). *Statistical methods in education and psychology* (3rd ed.). Englewood Cliffs, N.J.: Prentice Hall.
- Goldschmidt, E. P. (1943). *Medieval texts and their first appearance in print*. London, UK: Oxford University Press.

- Gow, R. W., Lydecker, J. A., Lamanna, J. D., & Mazzeo, S. E. (2012). Representations of celebrities' weight and shape during pregnancy and postpartum: A content analysis of three entertainment magazine websites. *Body Image, 9*(1), 172-175. doi: 10.1016/j.bodyim.2011.07.003
- Graesser, A. C., McNamara, D. S., & Louwerse, M. M. (2003). What do readers need to learn in order to process coherence relations in narrative and expository text. In A. P. Sweet & C. E. Snow (Eds.), *Rethinking reading comprehension* (pp. 82-98). New York: Guilford Publications.
- Graesser, A. C., McNamara, D. S., Louwerse, M. M., & Cai, Z. (2004). Coh-metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, and Computers, 36*, 193-202.
- Habib, M. A. R. (2008). *A history of literary theory and criticism from Plato to the present*. Malden, MA: Blackwell Publishing.
- Halliday, M. A. K. (1985). *An introduction to functional grammar*. Baltimore, MD: Edward Arnold.
- Halliday, M. A. K. (1989). *Spoken and written language* (2nd ed.). Oxford, UK: Oxford University Press.
- Haneefa, M., & Nellikka, S. (2010). Content analysis of online English newspapers in India. *Journal of Library & Information Technology, 30*(4), 17-24.
- Hartman, D. K. (1992). Intertextuality and reading: The text, the reader, the author, and the context. *Linguistics and Education, 4*(14997280346166516363), 295-311. doi: papers://9B5639B8-E52F-4416-81F5-893D7028B012/Paper/p1151

- Henry, A. E., & Story, M. (2009). Food and beverage brands that market to children and adolescents on the Internet: A content analysis of branded web sites. *Journal of Nutrition Education and Behavior, 41*(5), 353-359. doi: 10.1016/j.jneb.2008.08.004.
- Henry, L. A. (2007). *Exploring new literacies pedagogy and online reading comprehension among middle school students and teachers: Issues of social equity or social exclusion?* (Doctoral dissertation). Retrieved from <http://digitalcommons.uconn.edu/dissertations/AAI3282520/>.
- Herring, S. C., Scheidt, L. A., Wright, E., & Bonus, S. (2005). Weblogs as a bridging genre. *Information Technology & People, 18*(2), 142-171. doi: 10.1108/09593840510601513
- Heylighen, F., & Dewaele, J.-M. (1999). *Formality of language: Definition, measurement, and behavioral determinants*. Internal Report submitted to Center "Leo Apostel", Free University of Brussels. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.30.6280>.
- Hicks, D. (1997). Working through discourse genres in school. *Research in the Teaching of English, 31*, 459-485.
- Hofman, R., & van Oostendorp, H. (1999). Cognitive effects of a structural overview in a hypertext. *British Journal of Educational Technology, 30*(2), 129-140.
- Horace, & Wilkins, A. S. (1964). *Ars poetica*. New York: Macmillan; St. Martin's Press.
- Hyon, S. (1996). Genre in three traditions: Implications for ESL. *TESOL Quarterly, 30*(Journal Article), 693-722.
- Hypertext. (2014). *Encyclopaedia Britannica*. Retrieved from <http://www.britannica.com/EBchecked/topic/279726/hypertext>

- Ihlström, C., & Henfridsson, O. (2005). Online newspapers in Scandinavia: A longitudinal study of genre change and interdependency. *Information Technology & People, 18*(2), 172-192. doi: 10.1108/09593840510601522
- Internet Archive. (2000). FunBrain – About Us – March 1, 2000. Retrieved on 10/3/2014 from <https://web.archive.org/web/20000301013727/http://www.funbrain.com/aboutus.html>.
- Internet Archive. (2014). Internet archive wayback machine. Information retrieved from <https://web.archive.org/>.
- Jamieson, K. M. (1975). Antecedent genre as rhetorical constraint. *Quarterly Journal of Speech, 61*, 406-415.
- Janesick, V. J. (2004). *"Stretching" exercises for qualitative researchers* (2nd ed.). Thousand Oaks, CA: Sage Publications.
- Juarascio, A. S., Shoaib, A., & Timko, C. A. (2010). Pro-eating disorder communities on social networking sites: A content analysis. *Eating Disorders, 18*(5), 393-407. doi: 10.1080/10640266.2010.511918
- Kamberelis, G. (1999). Genre development and learning: Children writing stories, science reports, and poems. *Research in the Teaching of English, 33*(4), 403-460. doi: papers://9B5639B8-E52F-4416-81F5-893D7028B012/Paper/p397
- Kamberelis, G., & Bovino, T. D. (1999). Cultural artifacts as scaffolds for genre development. *Reading Research Quarterly, 34*(2), 138-170.
- Kandel, E.R., Schwartz, J.H. and Jessell, T.M. (2000). *Principles of Neural Science*, 4th Ed. New York, NY: McGraw-Hill.

- Kelly, B. P., Bochynska, K., Kornman, K., & Chapman, K. (2008). Internet food marketing on popular children's websites and food product websites in Australia. *Public Health Nutrition, 11*(11), 1180-1187.
- KidSites. (2014). Your guide to the best kid sites on the web. Retrieved on 9/8/2014 from <http://www.kidsites.com>.
- Kinneavy, J. (1971). *A theory of discourse*. Englewood Cliffs, N.J.: Prentice Hall.
- Kramarski, B., & Feldman, Y. (2000). Internet in the classroom: Effects on reading comprehension, motivation and metacognitive awareness. *Educational Media International, 37*(9244071320269056344), 149-155. doi: papers://9B5639B8-E52F-4416-81F5-893D7028B012/Paper/p1175
- Kress, G. R. (1999). Genre and the changing contexts for English language arts. *Language Arts, 76*(6), 461-469.
- Kress, G. R. (2003a). *Multimodality, representation and new media*. Paper presented at the Expert Forum for Knowledge Presentation: Resources for Communication, Forum for discussion, planning, and collaboration, Chicago, IL.
- Kress, G. R. (2003b). *Reading images: Multimodality, representation and new media*. Paper presented at the Expert Forum for Knowledge Presentation: Resources for Communication, Forum for discussion, planning, and collaboration, Chicago, IL.
- Kress, G. R. (2009). *Multimodality: A social semiotic approach to contemporary communication*. New York, NY: Routledge.
- Krippendorff, K. (2004). *Content analysis: An introduction to its methodology* (2nd ed.). Thousand Oaks, CA: Sage.

- Kristeva, J. (1984). *Revolution in poetic language* (M. Waller, Trans.). New York, NY: Columbia University Press.
- Langford, B. (2005). *Film Genre: Hollywood and Beyond*. Edinburgh, UK: Edinburgh University Press.
- Lankshear, C., & Knobel, M. (2007). A new literacies sampler. 263. doi: papers://9B5639B8-E52F-4416-81F5-893D7028B012/Paper/p943
- Lawless, K. A., Mills, R., & Brown, S. W. (2002). Children's hypertext navigation strategies. *Journal of Research on Technology in Education*, 34(3), 274-284.
- Learn about robots.txt files - Webmaster Tools Help. (2014). Retrieved on September 30, 2014 from <https://support.google.com/webmasters/answer/6062608?hl=en>.
- Lemke, J. L. (2005). Multimedia genres and traversals. *Folia Linguistica*, 39(1-2), 45-56.
- Leu, D.J., Mcverry, J. G., O'Byrne, W. I., Zawilinski, L., Castek, J., & Hartman, D. (2009). The new literacies of online reading comprehension and the irony of No Child Left Behind: Students who require our assistance the most, actually receive it the least. In L. M. Morrow, R. Rueda & D. Lapp (Eds.), *Handbook of research on literacy and diversity* (pp. 173-194). New York, NY: Guilford.
- Leu, D.J., Castek, J., Hartman, D., Coiro, J., Henry, L. A., Kulikowich, J. M., & Lyver, S. (2005). *Evaluating the development of scientific knowledge and new forms of reading comprehension during online learning*: North Central Regional Educational Laboratory/Learning Point Associates.
- Leu, D. J., & Kinzer, C. (2000). The convergence of literacy instruction with networked technologies for information and communication. *Reading Research Quarterly*, 35(1), 108-127.

- Longinus. (1st c./1964). *'Longinus' on the sublime* (D. A. Russell, Trans.). Oxford, England: Clarendon Press.
- Luke, A. (1996). Genres of power: Literacy education and the production of capital. In R. Hasan & G. Williams (Eds.), *Literacy in society*. New York, NY: Longman.
- Marenbon, J. (1988). *Early medieval philosophy (480-1150)*. New York, NY: Routledge.
- Marenbon, J. (2013). Anicius manlius severinus boethius. In E. N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy* (Summer 2013 ed.).
- Marsh, E. E. (2002). *Rhetorical relationships between images and text in web pages*. (Ph.D. Dissertation), University of Maryland, College Park, MD.
- McDonald, S., & Stevenson, R. J. (1998). Effects of text structure and prior knowledge of the learner on navigation in hypertext. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 40(1), 18-27. doi: 10.1518/001872098779480541
- McNamara, D. S., Louwerse, M. M., McCarthy, P. M., & Graesser, A. C. (2010). Coh-metrix: Capturing linguistic features of cohesion. *Discourse Processes*, 47(4), 292-330. doi: 10.1080/01638530902959943
- McPherran, M. L. (2010). *Plato's republic: A critical guide*. Cambridge, UK: Cambridge University Press.
- Mehta, C. (2014). Name that color. Retrieved 7/15/14 from <http://chir.ag/projects/name-that-color/#6195ED>.
- Miller, C. R. (1984). Genre as social action. *Quarterly Journal of Speech*, 70, 151-167.
- Miller, C. R., & Shepherd, D. (2004). Blogging as social action: A genre analysis of the weblog. *Into the Blogosphere: Rhetoric, Community, and Culture of Weblogs*, from http://blog.lib.umn.edu/blogosphere/blogging_as_social_action.html

- Mohageg, M. (1992). The influence of hypertext linking structures on the efficiency of information retrieval. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 34(3), 351-367. doi: 10.1177/001872089203400308
- Mokhtari, K., Kymes, A., & Edwards, P. (2008). Assessing the new literacies of online reading comprehension: An informative interview with W. Ian O'Byrne, Lisa Zawilinski, J. Greg Mcverry, and Donald J. Leu at the University of Connecticut. [10.1598/RT.62.4.9]. *The Reading Teacher*, 62(4), 354-357. doi: papers://9B5639B8-E52F-4416-81F5-893D7028B012/Paper/p1162
- Moore, E. S. (2006). *It's child's play: Advergaming and the online marketing of food to children*. Menlo Park, CA: Kaiser Family Foundation.
- Muller, H. J. (1967). *The uses of English*. New York, NY: Holt, Rinehart and Winston.
- National Geographic Kids. (2014). Kids' games, animals, photos, stories, and more -- National Geographic Kids. Retrieved on 10/15/14 from <http://kids.nationalgeographic.com/>.
- Nikolajeva, M., & Scott, C. (2001). *How picturebooks work*. New York, NY: Garland.
- O'Byrne, W. I., & McVerry, J. G. (2009). Measuring the dispositions of online reading comprehension: A preliminary validation study. *The 57th National Reading Conference yearbook* (pp. 362-375).
- Opoku, R. A., Pitt, L. F., & Abratt, R. (2007). Positioning in cyberspace: Evaluating bestselling authors' online communicated brand personalities using computer-aided content analysis. *South African Journal of Business Management*, 38(4), 21-32.
- Pappas, C. C. (1991). Young children's strategies in learning the "book language" of information books. *Discourse Processes*, 14(2), 203-225.

- Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage.
- PBS. (2014). PBS parents: Your resource for parenting tips & parenting advice. Retrieved on 10/2/2014 from <http://www.pbs.org/parents/>.
- Portela, M. (2007). New textualities. *European Journal of English Studies*, 11(2), 121-132. doi: 10.1080/13825570701452649
- “Public Broadcasting Service”. (2014). In *Encyclopaedia Britannica*. Retrieved on 10/1/2014 from <http://www.britannica.com/EBchecked/topic/482306/Public-Broadcasting-Service>.
- Puschmann, C. (2009). Lies at Walmart: Style and the subversion of genre in the life at Wal-mart blog. In J. Giltrow & D. Stein (Eds.), *Genres in the Internet: Issues in the theory of genre*. Philadelphia, PA: John Benjamins Publishing Company.
- Random House. (2014). About us: Dr. Seuss frequently asked questions. Retrieved on 10/2/2014 from http://www.seussville.com/faq/about_us.php.
- Ratliff, C. (2003, September 12). Genre theory, genre analysis, and blog as genre [Web log post]. Retrieved from <http://kairosnews.org/genre-theory-genre-analysis-and-blog-as-genre>.
- Rouet, J. F., Levonen, J. J., Dillon, A., & Spiro, R. J. (1996). *Hypertext and cognition*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Rumianek, M. (2013). Archiving and recovering database-driven websites. *D-Lib Magazine*, 19(1/2), 6.
- Saeed, S. (2011). Showcase of beautiful search box UI designs. Hongkiat.com. Retrieved 10-15-14 from <http://www.hongkiat.com/blog/beautiful-website-search-box-ui-design/>.
- Saiki, D. (2010). Interacting online: A content analysis of museum education websites. *Journal of Learning Design*, 4(1), 52-62.

- Salmeron, L., Kintsch, W., & Canas, J. (2006). Reading strategies and prior knowledge in learning from hypertext. *Memory and Cognition*, 34(14745524428223208226related:ItsWo1egoswJ), 1157-1171. doi: papers://9B5639B8-E52F-4416-81F5-893D7028B012/Paper/p1249
- Saussure, F. (1959). *Course in general linguistics*. (Charles Bally and Albert Reidlinger, Eds). (Wade Baskin, Trans.). New York, NY: Philosophical Library.
- Schryer, C. F. (1993). Records as genre. *Written Communication*, 10, 200-234.
- Shaw, J. F. (1886). *Writings of Augustine: On Christian doctrine*. New York, NY: Christian Literature Publishing.
- Shechter, M., & Schechter, S. R. (1987). *Children's acquisition of literary genre: Science fiction versus fantasy*. Paper presented at the 12th Annual Meeting of the Boston University Conference on Language Development, Boston, MA.
- Shin, W., & Huh, J. (2009). Multinational corporate website strategies and influencing factors: A comparison of us and Korean corporate websites. *Journal of Marketing Communications*, 15(5), 287-310. doi: 10.1080/13527260802481207
- Snow, C. E. (2002). *Reading for understanding toward an R & D program in reading comprehension*. Santa Monica, CA: Rand.
- Snyder, T.D., and Dillow, S.A. (2011). *Digest of Education Statistics 2010* (NCES 2011-015). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.
- So, B. P. C. (2005). From analysis to pedagogic applications: Using newspaper genres to write school genres. *Journal of English for Academic Purposes*, 4(1), 67-82. doi: papers://9B5639B8-E52F-4416-81F5-893D7028B012/Paper/p369

- Spivak, G. C. (1990). *The post-colonial critic*. New York, NY: Routledge.
- Steel, E. (2014). Nickelodeon's Digital Generation. *New York Times*. Retrieved 10/2/2014 from http://www.nytimes.com/2014/08/04/business/media/nickelodeons-digital-generation.html?_r=0.
- Sublette, J. R. (1973). The Dartmouth conference: Its reports and results. *College English*, 35(3), 348-357.
- Suter, W. N. (2012). *Introduction to educational research: A critical thinking approach* (2nd ed.). Thousand Oaks, CA: Sage Publications.
- Swales, J. M. (2009). Worlds of genre—metaphors of genre. In C. Bazerman, A. Bonini & D. B. Figueiredo (Eds.), *Genre in a changing world* (pp. 3-16). Fort Collins, CO: The WAC Clearinghouse.
- Swales, J. M. (2011). Coda: Reflections on the future of genre and l2 writing. *Journal of Second Language Writing*, 20(1), 83-85. doi: 10.1016/j.jslw.2010.12.005
- Telford, K. A. (1961). *Aristotle's poetics : Translation and analysis* (Gateway ed.). Chicago: Henry Regnery.
- Tsiotsou, R. H., Rigopoulou, I. D., & Kehagias, J. D. (2010). Tracing customer orientation and marketing capabilities through retailers' websites: A strategic approach to Internet marketing. *Journal of Targeting, Measurement and Analysis for Marketing*, 18(2), 79-94. doi: 10.1057/jt.2010.5
- Tyrkkö, J. (2007). Making sense of digital textuality. *European Journal of English Studies*, 11(2), 147-161. doi: 10.1080/13825570701452722
- Viacom. (2014). Viacom brands. Retrieved on 10/10/2014 from <http://origin.www.viacom.com/brands/>.

- Villanueva-Alfonso, M. L., Luzón-Marco, M. J., & Ruiz-Madrid, M. N. (2008). Understanding digital genres as semiotic artefacts: Meaning and cognition beyond standardised genres. *Computers and Composition Online*. Retrieved from <http://www.bgsu.edu/departments/english/cconline/DigitalGenres/>
- Weber, K., Story, M., & Harnack, L. (2006). Internet food marketing strategies aimed at children and adolescents: A content analysis of food and beverage brand web sites. *Journal of the American Dietetic Association, 106*(9), 1463-1466. doi: 10.1016/j.jada.2006.06.014
- Wierzbicka, A. (1990). The meaning of color terms: Semantics, culture, and cognition. *Cognitive Linguistics, 1*(1), 99-150.
- Wise, K., Bolls, P., & Schaefer, S. (2008). Choosing and reading online news: How available choice affects cognitive processing. [10.1080/08838150701820858]. *Journal of Broadcasting & Electronic Media, 52*(1), 69-85. doi: papers://9B5639B8-E52F-4416-81F5-893D7028B012/Paper/p1250
- Wohn, D. Y. (2011). Gender and race representation in casual games. *Sex Roles, 65*(3-4), 198-207. doi: 10.1007/s11199-011-0007-4
- Wollman-Bonilla. (2000). Teaching science writing to first graders: Genre learning and recontextualization. *Research in the Teaching of English, 35*, 35-65.
- Xi, Y., Zhuang, Y., Huang, W., She, C., & Zhang, Z. (2007). The quality assessment and content analysis of corporate websites in china: An empirical study. *International Journal of Information Technology & Decision Making, 6*(2), 389-405.

- Yu, J., King, K. W., & Yoon, H. J. (2010). How much are health websites influenced by culture? Content analysis of online diet programs in the United States, the United Kingdom, and Korea. *Journal of Promotion Management, 16*(3), 331-359. doi: 10.1080/10496490903294703
- Zammit, K. (2011). Moves in hypertext: The resource of negotiation as a means to describe the way students navigate a pathway through hypertext. *Linguistics and Education, 22*(2), 168-181. doi: 10.1016/j.linged.2011.02.005
- Zhang, S., & Duke, N. (2008). Strategies for Internet reading with different reading purposes: A descriptive study of twelve good Internet readers. *Journal of Literacy Research, 40*(1), 128-162. doi: 10.1080/10862960802070491

Appendix A: Coh-Metrix Output Fields

Indices in the Coh-Metrix output.

No	Description	Measure	Full description
1	Title	Title	Title
2	Genre	Genre	Genre
3	Source	Source	Source
4	JobCode	JobCode	JobCode
5	LSASpace	LSASpace	LSASpace
6	Date	Date	Date
7	Adjacent anaphor reference	CREFP1u	Anaphor reference, adjacent, unweighted
8	Anaphor reference	CREFPau	Anaphor reference, all distances, unweighted
9	Adjacent argument overlap	CREFA1u	Argument Overlap, adjacent, unweighted
10	Argument overlap	CREFAau	Argument Overlap, all distances, unweighted
11	Adjacent stem overlap	CREFS1u	Stem Overlap, adjacent, unweighted
12	Stem overlap	CREFSau	Stem Overlap, all distances, unweighted
13	Content word overlap	CREFC1u	Proportion of content words that overlap between adjacent sentences
14	LSA sentence adjacent	LSAassa	LSA, Sentence to Sentence, adjacent, mean
15	LSA sentence all	LSApsa	LSA, sentences, all combinations, mean
16	LSA paragraph	LSAppa	LSA, Paragraph to Paragraph, mean
17	Personal pronouns	DENPRPi	Personal pronoun incidence score
18	Pronoun ratio	DENSPR2	Ratio of pronouns to noun phrases
19	Type-token ratio	TYPTOKc	Type-token ratio for all content words
20	Causal content	CAUSVP	Incidence of causal verbs, links, and particles
21	Causal cohesion	CAUSC	Ratio of causal particles to causal verbs (cp divided by cv+1)

22	Intentional content	INTEi	Incidence of intentional actions, events, and particles.
23	Intentional cohesion	INTEC	Ratio of intentional particles to intentional content
24	Syntactic structure similarity adjacent	STRUTa	Sentence syntax similarity, adjacent
25	Syntactic structure similarity all-1	STRUTt	Sentence syntax similarity, all, across paragraphs
26	Syntactic structure similarity all 2	STRUTp	Sentence syntax similarity, sentence all, within paragraphs
27	Temporal cohesion	TEMPta	Mean of tense and aspect repetition scores
28	Spatial cohesion	SPATC	Mean of location and motion ratio scores.
29	All connectives	CONi	Incidence of all connectives
30	Conditional operators	DENCOND _i	Number of conditional expressions, incidence score
31	Pos. additive connectives	CONADpi	Incidence of positive additive connectives
32	Pos. temporal connectives	CONTPpi	Incidence of positive temporal connectives
33	Pos. causal connectives	CONCSp _i	Incidence of positive causal connectives
34	Pos. logical connectives	CONLGpi	Incidence of positive logical connectives
35	Neg. additive connectives	CONADni	Incidence of negative additive connectives
36	Neg. temporal connectives	CONTPni	Incidence of negative temporal connectives
37	Neg. causal connectives	CONCSni	Incidence of negative causal connectives
38	Neg. logical connectives	CONLGni	Incidence of negative logical connectives
39	Logic operators	DENLOGi	Logical operator incidence score (and + if + or + cond + neg)
40	Raw freq. content words	FRQCRacw	Celex, raw, mean for content words (0-1,000,000)
41	Log freq. content words	FRQCLacw	Celex, logarithm, mean for content words (0-6)
42	Min. raw freq. content words	FRQCRmcs	Celex, raw, minimum in sentence for content words (0-1,000,000)
43	Log min. freq. content words	FRQCLmcs	Celex, logarithm, minimum in sentence for content words (0-6)
44	Concreteness content words	WORDCacw	Concreteness, mean for content words
45	Min. concreteness content words	WORDCmcs	Concreteness, minimum in sentence for content words

46	Noun hypernym	HYNOUNa w	Mean hypernym values of nouns
47	Verb hypernym	HYVERBa w	Mean hypernym values of verbs
48	Negations	DENNEGi	Number of negations, incidence score
49	NP incidence	DENSNP	Noun Phrase Incidence Score (per thousand words)
50	Modifiers per NP	SYNNP	Mean number of modifiers per noun-phrase
51	Higher level constituents	SYNHw	Mean number of higher level constituents per word
52	Words before main verb	SYNLE	Mean number of words before the main verb of main clause in sentences
53	No. of words	READNW	Number of Words
54	No. of sentences	READNS	Number of Sentences
55	No. of paragraphs	READNP	Number of Paragraphs
56	Syllables per word	READASW	Average Syllables per Word
57	Words per sentence	READASL	Average Words per Sentence
58	Sentences per paragraph	READAPL	Average Sentences per Paragraph
59	Flesch Reading Ease	READFRE	Flesch Reading Ease Score (0-100)
60	Flesch-Kincaid	READFKG L	Flesch-Kincaid Grade Level (0-12)