Reading in a Foreign Language April 2008, Volume 20, No. 1
ISSN 1539-0578 pp. 92–122

Vocabulary recycling in children's authentic reading materials: A corpus-based investigation of narrow reading

Dee Gardner
Brigham Young University
United States

Abstract

Fourteen collections of children's reading materials were used to investigate the claim that collections of authentic texts with a common theme, or written by one author, afford readers with more repeated exposures to new words than unrelated materials. The collections, distinguished by relative thematic tightness, authorship (1 vs. 4 authors), and register (narrative vs. expository), were analyzed to determine how often, and under what conditions, specialized vocabulary recycles within the materials. Findings indicated that thematic relationships impacted specialized vocabulary recycling within expository collections (primarily content words), whereas authorship impacted recycling within narrative collections (primarily names of characters, places, etc.). Theme-based expository collections also contained much higher percentages of theme-related words than their theme-based narrative counterparts. The findings were used to give nuance to the vocabulary-recycling claims of narrow reading and to more general theories and practices involving wide and extensive reading.

Keywords: narrow reading, vocabulary, themes, registers, authorship

Over the past 30 years, a large body of literature has touted reading as the major source of students' vocabulary development (e.g., Cunningham & Stanovich, 2003; Krashen, 1989, 1993a, 1993b; Nagy & Anderson, 1984; Nagy & Herman, 1985, 1987). This claim has also received some empirical support from studies that have found small, incremental gains in word knowledge through contextual exposure during reading (reviewed in Swanborn & de Glopper, 1999), as well as studies that have correlated amount of print exposure with large vocabulary differences among school-aged children (reviewed in Cunningham & Stanovich, 2003). As a result, *wide reading* (reading large amounts of "authentic" material) and its more robust conceptualization, *extensive reading*, have been advocated for expanding the vocabularies of various learners in first-language (L1), second-language (L2), and foreign-language instructional settings (Cunningham & Stanovich, 2003; Day & Bamford, 1998, 2002; Graves, 2006; Krashen, 1989, 1993a, 1993b).

At the heart of this issue is the assumption that readers will encounter new (unfamiliar) words multiple times in multiple and varied contexts during extensive reading experiences, eventually

resulting in the "incidental acquisition" of those words (Nagy, 1997; Nagy, Anderson, & Herman, 1987; Nagy & Herman, 1987; Shu, Anderson, & Zhang, 1995). Proponents have also put forth this hypothesis as being the best explanation of how young L1 learners acquire the bulk of their large vocabularies through the 12th grade, with estimates ranging somewhere between 40,000 (Nagy & Herman, 1987) and 80,000 words (Anderson, 1996; Anderson & Nagy, 1992), depending on what is counted as a word.

As appealing as this hypothesis has been in reading research and pedagogy, there remains a relative dearth of research studies that have carefully considered the vocabulary input of children's authentic reading materials to determine how well, and under what conditions, they do recycle vocabulary, particularly those words that are not from the relatively small pool of high-frequency forms found in most texts (*the*, *of*, *and*, *a*, *take*, *get*, *mother*, *play*, etc.). The much larger group of more specialized vocabulary items—to which the words of this study belong—constitutes the bulk of the English word stock (Nation, 1990), thus effectively representing the large-scale vocabulary (e.g., 40,000 to 80,000 words) that can potentially be acquired during the school years and beyond.

With this background in mind, the aim of the current study is to extend the earlier work of Gardner (2004), in which he analyzed the vocabulary input of a 1.5 million-word extensive reading corpus, consisting of seven children's narrative collections (four texts each) and seven grade-equivalent expository collections (four texts each). One of his major findings was that the words children are exposed to during narrative reading are vastly different than those they are exposed to during expository reading, particularly at the more specialized, content-rich levels of vocabulary (i.e., beyond the high-frequency words of the language) where 17,921 of 23,857 word types (72.5%) were either found in narrative texts only or grade-equivalent expository texts only (i.e., zero overlap). Additionally, this lack of sharing of critical word types occurred even though many of the narrative and expository collections were related by common themes—one of two conditions proposed by advocates of *narrow reading* for improving vocabulary recycling in reading curricula of English as an L2 or a foreign language (e.g., Cho, Ahn, & Krashen, 2005; Day, 1994; Krashen, 1981, 1985, 2004; Schmitt & Carter, 2000), the other being authorship (i.e., using texts written by the same author). Krashen (1985) has articulated these two conditions as follows:

> If the Input Hypothesis is correct . . . it suggests that narrow input is more efficient for L2 acquisition, that early specialization rather than late specialization is better, that students should be encouraged to read on only one topic at a time, or several books by the same author, in the intermediate stage, and that [L2] students stay on somewhat familiar ground when they first enter the mainstream. . . . In addition, each topic has its own vocabulary, and to some extent its own style; the same can be said for each author. Narrow input provides many exposures to these new items in a comprehensible context and built-in review. (p. 73)

The current study examines this assertion from the standpoint of authentic vocabulary input from the children's reading corpus (Gardner, 2004), considering *theme* (*topic*) and *authorship*, in addition to *register*, as primary variables of interest in order to tease apart nuances of specialized vocabulary recycling in authentic reading collections.

At the outset, the potential benefits of narrow reading are recognized to extend beyond vocabulary recycling only (e.g., exposing L2 readers to consistent stylistic and discourse moves of certain authors). However, because vocabulary recycling is a central tenet of this position, it deserves more careful examination. A clearer understanding of the impact of text relationships on vocabulary recycling will serve as a guide for theories and practices in English language education in general, particularly in the areas of wide and extensive reading and vocabulary development. The findings may also prove informative in L1 settings, where the assumed language benefits of theme-based instruction (e.g., Walmsley, 1994) and the known challenges with content-area, nonfiction reading materials (e.g., Bamford, Kristo, & Lyon, 2002; Vacca & Vacca, 1996) have also received a great deal of attention.

Why a Focus on Authentic Reading Materials?

Before proceeding, it is important to note that the existence of narrow reading and similar approaches (e.g., Dubin, 1986) is largely a result of the linguistic characteristics of authentic reading materials. Such materials, unlike graded readers (e.g., Waring, 2003; Wodinsky & Nation, 1988), basal readers (e.g., Bello, Fajet, Shaver, Toombs, & Schumm, 2003), decodable texts (e.g., Mesmer, 2001), or other linguistically engineered materials, do not intentionally control for the presentation of vocabulary and other language structures. By their nature, authentic reading materials are fairly unpredictable in terms of the language demands they place on readers, as well as the language-learning opportunities they afford. While authentic oral communication is often simplified and repeated in order to achieve the conditions of comprehensible input, the same is not true for most authentic written language, which is made permanent in print, thus removing the author from the reader in terms of both time and space. While modern technology may hold the key to making written text more flexible as a language learning tool (Cobb, 2007; Huang & Liou, 2007), and while such technology has also introduced e-mailing, on-line chatting, and text-messaging with their real-time, two-way communication capabilities, these modes of written communication are vastly different from the linguistically frozen materials of printed school English (novels, trade books, textbooks, etc.). By extension, narrow reading is simply one attempt to deal with this challenge of authentic written input by suggesting that collections of authentic texts written on similar topics or by one author will improve the chances that essential linguistic redundancy will actually take place, or in other words, that readers, especially L2 readers, will be exposed to necessary levels of repetitive, comprehensible input as they move from one text to the next.

> Essentially, vocabulary learning from extensive reading is very fragile. If the small amount of learning of a word is not soon reinforced by another meeting, then that learning will be lost. It is thus critically important in an extensive reading program that learners have the opportunity to keep meeting words they have met before. (Nation, 1997, p. 15)

A clearer understanding of how relationships between authentic reading materials might affect such crucial vocabulary recycling is at the heart of the current study.

Why a Focus on Specialized Vocabulary?

The work of Paul Nation and his colleagues has been instrumental in showing the distributions of vocabulary in authentic written and spoken materials. Table 1 is a repurposing of Nation's (2001) analysis of the distribution of vocabulary in the American Heritage Intermediate (AHI) corpus (Carroll, Davies, & Richman, 1971), which consists of 5 million running words taken from a random selection of third- through ninth-grade texts. This corpus is particularly important to the current study because it was the primary source for the landmark claims associated with the incidental hypothesis (Nagy & Anderson, 1984) and the call for wide reading in reading instruction (Nagy, Herman, & Anderson, 1985).

Table 1. *Vocabulary coverage in the American Heritage Intermediate corpus*

Number of word families	Cumulative % of text coverage
10	23.7
100	49
1,000	74.1
2,000	81.3
3,000	85.2
4,000	87.6
5,000	89.4
12,448	95
43,831	99
86,741	100

Note. Adapted from Nation (2001, p. 15).

Table 1 shows clearly that a small subset of high-frequency word families (i.e., base forms plus their inflections and transparent derivations, e.g., *climb*, *climbs*, *climbing*, *climbed*, *climber*, *climbers*) covers most of the running words of the AHI corpus. For instance, the top 100 word families cover nearly half (49%) of the running words, and the top 1,000 word families cover nearly three-fourths (74.1%) of the running words. Examples of these high-frequency words include function words (*the*, *of*, *and*, *a*, *to*, *in*, etc.) and high-frequency content words (*take*, *get*, *said*, *people*, *find*, *water*, *words*, *know*, etc.), many of which can be found in authentic children's texts.

However, the remaining 85,741 word families in the AHI corpus (86,741 minus 1,000) cover only slightly more than one-fourth (25.9%) of the running words. This means that they repeat much less frequently in general than the 1,000 high-frequency word families. In most cases, however, these less frequent word families characterize a particular text or content area. They are also the words that children are less likely to know, and for which the long-term vocabulary learning benefits of extensive reading are most likely to be realized (*nourishment*, *saturated*, *tomb*, *mineral*, *topographic*, *prohibition*, *tomahawk*, etc.). Determining how often, and under what conditions, these words actually repeat in collections of authentic reading materials is the focus of this study.

Linguistic Studies of Vocabulary Recycling in Narrow-Reading Materials

Most of the linguistic studies that consider the impact of text-level variables such as theme or authorship on vocabulary recycling have focused on adult-level materials. The findings are nonetheless important to the current study. For instance, Hwang and Nation (1989) performed an analysis of the vocabulary load in running stories from newspapers versus the vocabulary load in unrelated stories, concluding that

> [A] higher proportion of word families outside the 2,000 [most frequent] words will recur in stories from the same series, thus reading running stories reduces the vocabulary load to a greater extent than reading unrelated stories … [and] running stories provide *more repetitions of more words outside the first 2,000 words* [italics added] than unrelated stories, and thus provide *more favorable conditions for learning vocabulary* [italics added] than unrelated stories. (p. 332)

The authors also suggested that their findings have implications for other texts besides newspapers, especially in settings of English as a foreign language, where several disparate topics often comprised textbooks.

Sutarsyah, Nation, and Kennedy (1994) also found substantial differences in the distribution of vocabulary between a single content text (economics), consisting of approximately 300,000 words, and a corpus of 160 shorter academic texts (from over 15 subject areas), consisting of approximately the same number of words. While the diverse corpus contained a much larger vocabulary base than the single text, the words were mostly of lower frequency. In contrast, "a small number of words that were closely related to the topic of the text occurred with very high frequency in the economics text" (p. 34). Additionally, with the exception of higher general frequency words (from the 2,000 word family list) and a few subtechnical terms common to many disciplines, there was little overlap in vocabulary between the narrower textbook and the broader corpus, leading the researchers to conclude the following:

> Most English courses make use of a series of unrelated texts. This can increase the vocabulary load of the course enormously. If teachers or course designers wish to avoid this, it is worth considering making the course consist of a few themes so that the texts within a theme bear more relationship to each other and thus make use of a smaller vocabulary. (p. 49)

It is important to note that the single text in this particular study was an expository textbook, consisting of a tight theme (macroeconomics) written by one author.

Finally, Schmitt and Carter (2000) compared the vocabulary of a series of nine theme-related newspaper stories (the tragic death of Princess Diana) to the vocabulary of nine unrelated stories from the same newspapers, containing the same number of total running words (7,843). The findings indicated that the theme-related Diana stories contained 156 fewer types (different words) for L2 readers to deal with, and repeated those types more often in general than the unrelated stories. This overall trend was also true when content words and proper nouns were examined, leading the researchers to the general conclusion that narrow reading may facilitate

earlier access to authentic L2 reading materials "by lowering the lexical load required of the learner" (p. 8).

The important point for purposes of the current study is that the theme-related texts were tightly related to each other (i.e., death of Princess Diana), whereas the unrelated stories had no connections beyond the fact that they belonged to the newspaper register in general. It is also crucial to note that only five of the content words actually listed in the study (*crash*, *palace*, *photographers*, *police*, *princess*—all occurring in the Diana stories) would be considered as specialized vocabulary in the current investigation, as the rest would have been identified as general high-frequency forms (e.g., *said*, *car*, *pay*, *school*, *people*, *work*, *time*, *one*, *year*); in other words, they come from the relatively small pool of general high-frequency forms and are therefore likely to occur in many texts, regardless of the relationships between those texts. While there is no question that continued exposure to such high-frequency forms is essential for building general reading fluency and text comprehension, it is equally clear that they do not represent the types of topic- and content-related words upon which a reader can build an extensive vocabulary.

Taxonomy of Textual Relationships

To date, very little has been done to formalize the potential relationships between authentic reading materials in terms of how such relationships might affect language sharing and recycling. While popular book-leveling schemes in elementary education (e.g., Fountas & Pinnell, 1996, 1999, 2005) have provided important guidelines for grouping texts according to general linguistic and print characteristics (percentages of higher-frequency vs. lower-frequency words, numbers of morphologically and conceptually complex words, font size, words per line, etc.), they do not address specific vocabulary redundancy that may occur as a result of thematic, authorship, or similar relationships between those texts (e.g., genre and register). Viewed another way, traditional leveling schemes tend to relate two or more texts based on the linguistic demands they place on young readers (i.e., how well such readers will be able to comprehend those texts), not on the potential redundancy of the textual content. Therefore, a book about plants and a book about outer space could both be rated at the same difficulty level, even though there is likely to be very little overlap in the topic-related words of the two texts (e.g., *blossom* and *root* vs. *star* and *comet*). Likewise, a children's adventure novel and a children's trade book about magnets could both be rated as at the same level, depending on their general linguistic and print characteristics.

Figure 1 depicts a proposed taxonomy for classifying relationships between authentic texts that could directly impact specialized (topic-related) vocabulary recycling within such materials. The taxonomy is in essence a classification scheme that could be used to more accurately predict the chances that *blossom* and *root* or, alternatively, *star* and *comet*, will appear in Text 1, Text 2, and so forth. Three primary text relationships are considered in the taxonomy: themes, authorship, and registers. In the case of themes, the primary considerations are twofold: (a) the general presence or absence of thematic relationships between texts and (b) the relative tightness of a given theme. For instance, *mummy* is a tighter theme than *mystery* in this study and might therefore be expected to recycle specialized vocabulary more efficiently. In general, *Gold Rush*

is a tighter theme than *Westward Movement*, which is a tighter theme than *American History*; *bees* is a tighter theme than *insects*, and so forth.

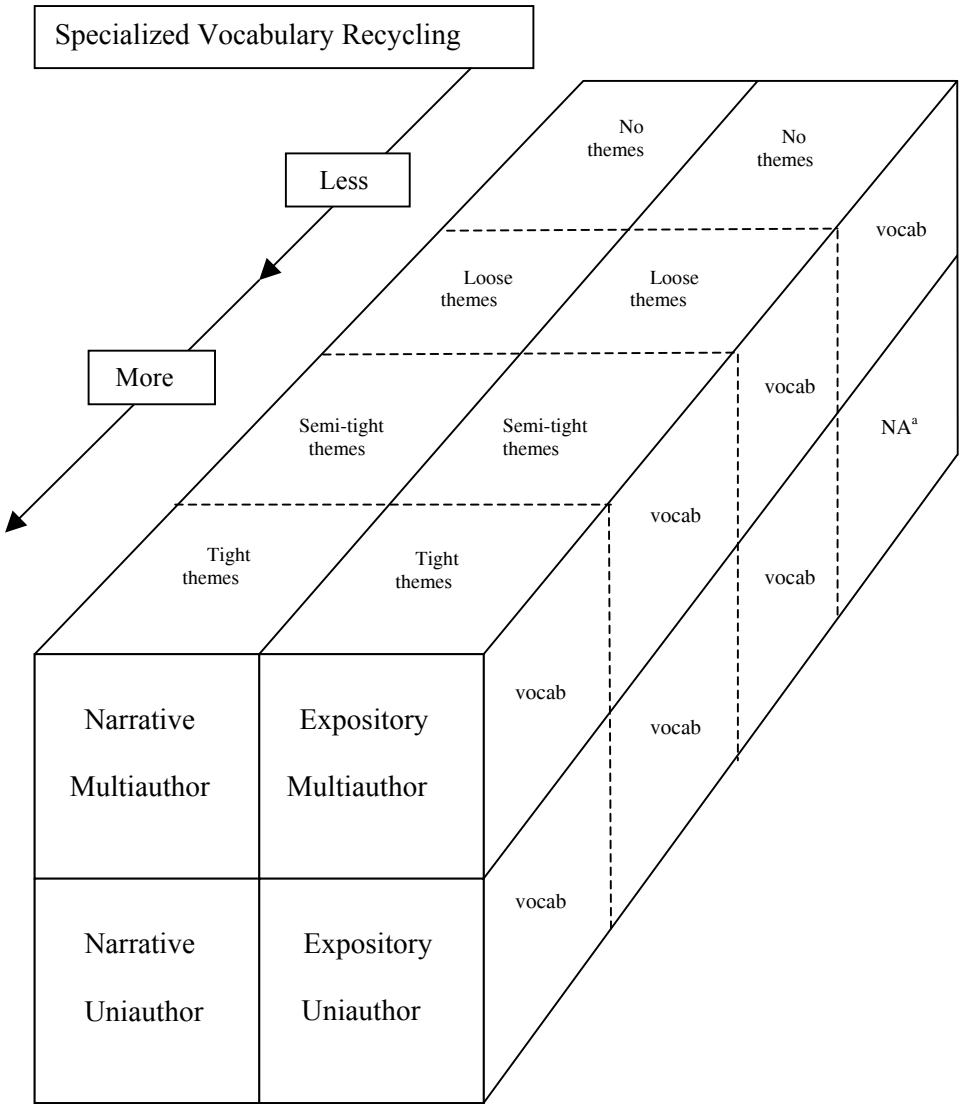

Figure 1. Proposed taxonomy of textual relationships for specialized vocabulary recycling in collections of authentic reading materials.

[a]There are no collection possibilities for the two cells on this particular row (Expository Uniauthor and Narrative Uniauthor) because *uniauthor* creates a potential relationship between the texts, even though they are not related by a content theme.

With regard to authorship issues, it has been broadly accepted that text collections written by one author (uniauthor) are more efficient in recycling vocabulary than text collections written by multiple authors (multiauthor). Finally, regarding register issues, the primary consideration has been the differences between the culturally- and socially-oriented vocabulary of narrative fiction (storybooks) and the informationally-oriented vocabulary of expository nonfiction.

The possible combinations of these variables are depicted in Figure 1 and will subsequently be referred to as the *taxonomy of textual relationships*.

A fourth dimension, content-area, may also have a bearing on vocabulary recycling in text collections. For instance, science-based materials may exhibit different vocabulary characteristics than history-based materials. However, because of the practical constraints of using an existing corpus, this dimension will only be addressed loosely in the current study by analyzing possible differences between the history-based collections under Westward Movement and the science-based collections under Mummy and Mystery.

It is clear that relatively little is known about vocabulary recycling as a function of text relationships, especially with regard to authentic children's reading materials. The current study will more carefully examine this issue by analyzing the specialized vocabulary of several collections of children's texts (Gardner, 2004) written at approximately the fifth- to sixth-grade level. The following question will be used to focus the analyses:

> To what extent do specialized words recycle within various collections of authentic children's reading materials that are related by (a) theme (Mystery, Westward Movement, Mummy), (b) authorship (texts written by different authors vs. texts written by one author), (c) register (narrative fiction vs. expository nonfiction), and (d) the various combinations of (a–c) above?

Method and Procedure

Constructs of Word, Vocabulary, and Type

The terms *word*, *vocabulary*, and *type* are used broadly and interchangeably in this study, and all three are defined conservatively as "unique spellings." While it is realized that some children may be able to make connections during reading between the morphologically related words of English (e.g., *climb*, *climbs*, *climbing*, *climbed*, *climber*), there is growing evidence of disparities in this ability based on children's individual reading skills (Carlisle, 2000; Mahony, Singson, & Mann, 2000; Singson, Mahony, & Mann, 2000) and the amount of direct instruction they receive in raising their morphological awareness (Carlo et al., 2004; Cunningham, 1998; Stahl & Shiel, 1992). Furthermore, differences in children's awareness of morphological relationships have been isolated as one of several significant variables predicting early vocabulary acquisition (McBride-Chang, Wagner, Muse, Chow, & Shu, 2005). The fact that many of the studies cited above deal with native English-speaking children or bilinguals suggests that this morphological-awareness problem may be even more pronounced for nonnative children trying to negotiate the complex morphological system of English. In fact, Schmitt and Zimmerman (2002) found that even adult learners of English (university students) struggle to make many morphological connections without explicit help, particularly when derivation is involved.

It should also be noted that the definition of word, vocabulary, and type used in this study does not account for multiword items (phrasal verbs, idioms, etc.) or variant meaning for the same

word forms (homonymy and polysemy). However, the more specialized nature of the words in this study suggests that there will be fewer chances for form-meaning error than if high-frequency words were being analyzed (Ravin & Leacock, 2000).

Children's Thematic Corpus

The actual corpus of children's extensive reading materials comes from Gardner's (2004) earlier study. Hereafter the corpus will be referred to as the Children's Thematic Corpus. With the aid of an experienced fifth-grade teacher and a children's librarian, Gardner established four collections of four texts each for each of three popular themes used in upper elementary education (fifth and sixth grades): Mummy (tight, science-based theme), Westward Movement (semitight, history-based theme), and Mystery (loose, science-based theme). This collaboration resulted in the 12 text collections outlined in Appendix A. A total of 48 texts were used to establish the four collections in each of the three themes: 3 × 4 × 4 (Themes × Collections × Texts).

Of the 48 texts in the 12 collections, 27 are from documented (published) thematic units (see the theme unit source key in Appendix A), and 21 were chosen with the expert assistance of the fifth-grade teacher and children's librarian, based on (a) subjective grade-level readability assessments or readability scores printed on the back covers of several books, (b) thematic fit, and (c) popularity of texts. A narrative and an expository control collection (no thematic or authorship relationships between the texts) were also established with the assistance of the teacher and children's librarian (see Appendix A). The control narrative collection consisted of four popular *Newbery Medal* books from four different genres of fiction (science, mystery, adventure, and romance), and the control expository collection consisted of four grade-equivalent informational books from four different content-areas (earth science, political science, life science, and geography-culture). From the perspective of the current investigation, the two control collections could alternatively be viewed as examples of wide reading, whereas the thematic collections would be more appropriately labeled as narrow reading.

Preliminary Procedure for Analyzing Vocabulary

Scanning. Each of the 56 texts (48 thematic and 8 control) was scanned into the computer using Omnipage text scanning software. Words not able to be scanned because of font and background problems were entered into the computer by keyboard. Each electronic document was then carefully edited to correct the relatively few scanning errors that occurred.

Equalization of word counts. For comparative purposes in the current study, each of the electronic texts was reduced to equal chunks of running words as follows: the first 5,000 running words of each text, beginning with the first word on page one.

This was done for two reasons: to account for differences in text length, especially between the lengthy narrative texts and the relatively short expository texts at the same grade level, and to allow comparisons of vocabulary repetition within a consistent number of running words that a child could encounter in a normal reading experience.

Identification of specialized vocabulary. The texts in each of the 14 collections (12 thematic and

2 control) were run, by collection, through the Range vocabulary program (Heatley, Nation, & Coxhead, 2002) and sorted into lists of High-Frequency Words and Other Words. The predetermined high-frequency list consisted of words from the first 1,000 word families of the General Service List (GSL; West, 1953), which accompanies the *Range* program, and which, unlike the second 1,000 GSL word families, have been found to be fairly stable over time (Nation & Hwang, 1995). An additional 108 function words and numerical terms that were not found in the first 1,000 GSL list were also added (e.g., *ahead*, *amid*, *billion*, *eighths*, *during*). The Other Words (i.e., not in the high-frequency list) were subsequently identified as being *specialized* if they appeared in at least three texts of a four-text collection.

Hereafter these words are referred to as *specialized words*, *specialized vocabulary*, or *specialized types* interchangeably. These are the words of interest in the current study, because they tend to characterize the content of the various collections of extensive reading materials (e.g., *mummy*, *pyramids*, *museum*, *archeologist*, *buffalo*, *prairie*, *investigation*). It is crucial to reiterate that these are shared, specialized words, occurring in several different texts of a collection instead of one text only (e.g., Hirsh & Nation, 1992). They are thus more representative of the types of words that children could encounter in an extensive reading program that uses themes, authors, and registers to organize instruction. They also fulfill the well established assumption for successful incidental word acquisition, namely, that children will encounter new words multiple times in multiple and varied contexts within a reasonable time frame.

Data Analysis

Once the specialized words were identified, three measures of vocabulary recycling were selected for comparison purposes:

> 1. Total number of specialized types (number of different words occurring in several texts). For instance, the words *mummy* and *prairie* would be counted as one type each, even though they might repeat 100 times and 4 times respectively.
>
> 2. Total number of specialized tokens (raw frequency counts). For instance, the word *mummy* would have a token count of 100 in the example above, and the word *prairie* would have a token count of 4.
>
> 3. Total number of specialized types that repeat at least six times (6+). For instance, the word *mummy* would be counted as one 6+ type in the scenario above (i.e., it repeats at least six times), whereas *prairie* would not be counted (i.e., it does not occur at least six times). The 6+ figure is a fairly conservative estimate of the number of incidental encounters that is generally necessary for new vocabulary to be acquired during extensive reading (see Zahar, Cobb, & Spada, 2001, for review.)

Each of these three measures addresses a different aspect of vocabulary recycling in authentic extensive reading collections. The first gives a general sense of how many different specialized words are drawn together by the relationships between texts or collections. The second provides an indication of how often these different words repeat in general, and the third provides information about specific specialized words that reach repetition levels conducive to incidental

vocabulary acquisition while reading (i.e., 6+ times).

Results and Discussion

Table 2 lists all 14 collections according to their average rank for the three vocabulary recycling measures (types, tokens, and 6+ types) based on the first 5,000 running words in each text (i.e., 20,000 per collection). It is clear that large differences exist in the number of specialized types, tokens, and 6+ types in the various collections. Looking first at the extremes, the Mummy Expository Uniauthor collection has 136 more specialized types than the non-thematic Control Expository Multiauthor collection (164 minus 28), 1,717 more specialized tokens (1,925 minus 208), and 85 more 6+ types (97 minus 12). These recycling differences become even more staggering when one considers that all of this happens in roughly 80 pages of text (20,000 running words). Additionally, the Mummy Expository Multiauthor collection is only slightly behind the Mummy Expository Uniauthor collection in this regard. This suggests that presence of a tight theme (Mummy) may have an important impact on vocabulary recycling.

Table 2. *Rank order of collections considering specialized types, tokens, and 6+ types appearing in at least three of four texts in each collection* (*based on 5,000 tokens per text, i.e., 20,000 per collection*)

Theme	Register	Authorship	Type total	Rank	Token total	Rank	6+ Type total	Rank	Average rank[a]
Mummy	Expository	Uniauthor	164	1	1,925	2	97	1	1.3
Mummy	Expository	Multiauthor	141	2	1,935	1	95	2	1.7
Westward	Narrative	Uniauthor	108	3	1,467	4	60	3	3.3
Westward	Expository	Multiauthor	93	4.5	1,004	5	54	4	4.5
Mystery	Narrative	Uniauthor	90	6	1,643	3	47	6.5	5.2
Westward	Expository	Uniauthor	80	7	720	7	48	5	6.3
Mummy	Narrative	Uniauthor	93	4.5	708	8	42	8	6.8
Mystery	Expository	Multiauthor	72	8.5	751	6	47	6.5	7.0
Westward	Narrative	Multiauthor	72	8.5	616	9	38	9	8.8
Mystery	Expository	Uniauthor	54	12	473	10	33	10	10.7
Control	Narrative	Multiauthor	58	10	360	11	21	11.5	10.8
Mystery	Narrative	Multiauthor	56	11	334	13	21	11.5	11.8
Mummy	Narrative	Multiauthor	52	13	346	12	20	13	12.7
Control	Expository	Multiauthor	28	14	208	14	12	14	14.0

Note. Westward = Westward Movement.
[a]Rank order based on average of three ranks for types, tokens, and 6+ types.

However, while Mummy Narrative Uniauthor appears to recycle specialized words much more efficiently than Control Narrative Multiauthor (no theme), the same is not true of its multiauthor counterpart (Mummy Narrative Multiauthor), which actually has fewer specialized types, tokens, and 6+ types than the narrative control. In fact, two of the theme-based narrative collections have average ranks that fall below the narrative control collection. The same is not true for the expository collections, which all exhibit more vocabulary recycling than the expository control collection. This suggests that theme may have a more important impact on vocabulary recycling in expository collections than in narrative collections. Also noteworthy in these data is the great disparity between Mummy Expository Uniauthor (tight, science-based theme) and Mystery Expository Uniauthor (loose, science-based theme) in terms of specialized types (164 vs. 54),

tokens (1,925 vs. 473), and 6+ types (97 vs. 33). Roughly the same disparity exists between Mummy Expository Multiauthor and Mystery Expository Multiauthor. The same is not true of their narrative equivalents. In fact, the order is exactly reversed, with the Mystery narratives (loose theme) showing slightly more specialized word recycling than the Mummy narratives. Interestingly, the Westward Movement collections (semitight, history-based theme) appear to congregate more toward the middle of the rankings as a whole. Again, these findings add support to the ongoing conclusion that thematic relationships between texts have their greatest impact on vocabulary recycling in expository, rather than narrative, text collections.

Table 3 reflects the major register distinction between expository and narrative texts as this distinction appears to be the primary variable of interest. The table reflects two clear vocabulary differences between the expository and narrative collections:

1. The Mummy expository collections (tight theme) occupy the top two rankings among expository texts; yet their Mummy narrative counterparts are ranked much lower among narratives, with the Mummy Narrative Multiauthor collection actually coming in last among narratives, falling below the Control Narrative Multiauthor collection (no theme) and the Mystery Narrative Multiauthor collection (loose theme).

2. The three highest ranked narrative collections, Westward Narrative Uniauthor (3.3), Mystery Narrative Uniauthor (5.2), and Mummy Narrative Uniauthor (6.8) are all single-author collections. This same pattern does not hold true for the expository collections. While the Mummy Expository Uniauthor collection (1.3) is slightly higher than the Mummy Expository Multiauthor collection (1.7), the exact opposite is true for the Westward Movement expository texts (multiauthor, 4.5; uniauthor, 6.3) and the Mystery expository texts (multiauthor, 7.0; uniauthor, 10.7). This suggests that single authorship has its greatest impact on vocabulary recycling within narrative text collections.

Table 3. *Rank order by expository and narrative collections*

Theme	Authorship	Average rank
	Expository collections	
Mummy	Uniauthor	1.3
Mummy	Multiauthor	1.7
Westward	Multiauthor	4.5
Westward	Uniauthor	6.3
Mystery	Multiauthor	7.0
Mystery	Uniauthor	10.7
Control	Multiauthor	14.0
	Narrative collections	
Westward	Uniauthor	3.3
Mystery	Uniauthor	5.2
Mummy	Uniauthor	6.8
Westward	Multiauthor	8.8
Control	Multiauthor	10.8
Mystery	Multiauthor	11.8
Mummy	Multiauthor	12.7

In summary, the major differences in vocabulary recycling seem to be at the register level, that is, between narrative and expository texts. With regard to the narrative collections, neither presence of theme (see Narrative Control vs. Narrative Mystery Multiauthor and Narrative Mummy Multiauthor) nor tightness of theme (see Mummy vs. Mystery) appears to make any real difference in terms of vocabulary recycling. Number of authors, on the other hand, appears to make a big difference, with all three uniauthor collections recycling specialized vocabulary better than their multiauthor counterparts.

With regard to expository collections, the situation is completely reversed. Not only does presence of theme make a difference (i.e., the Control Expository Collection is ranked last, and by a good margin), but the relative tightness of the themes matches the taxonomy perfectly (see Figure 1), with the two thematically tight Mummy collections ranked the highest, followed by the two thematically semitight Westward Movement collections, and, lastly, the two thematically loose Mystery collections. Number of authors, on the other hand, appears to produce inconsistent, even random differences in specialized vocabulary recycling within the expository collections.

Table 4 provides the top 10 most frequent specialized words appearing in all 14 collections. (See Appendix B for a list of all specialized words by collection.) The table is arranged horizontally by theme and vertically by register and authorship. An examination of these words is very informative. First, the presence of theme can be clearly identified in the theme-based expository collections (e.g., Mummy = *mummy*, *Egypt*, *tombs*, *pyramids*, *preserved;* Westward Movement = *trail*, *cattle*, *wagon*, *fort*, *Indians;* Mystery = *bones*, *evidence*, *clues*, *buried*, *horror*). This is in stark contrast with the more semantically random words of the expository control collection (e.g., *feet*, *America*, *area*, *huge*, *ice*).

However, this same contrast is not so readily apparent between the theme-based narrative collections and the narrative control collection. With the exception of two words in the Narrative Mummy Multiauthor collection (*Egypt* and *Egyptian*), it is difficult to find any theme-related differences between the two Mummy narrative collections and the narrative control collection. There is also no discernable, theme-related difference between the specialized vocabulary of the two Mystery narrative collections and the narrative control collection. However, several words in the Westward Movement narratives do exhibit some of the same thematic characteristics as their expository counterparts (e.g., *wagon*, *prairie*, *Indians*). Such similarities may be a result of the content area that they are drawn from (history vs. science)—a topic that should be more carefully considered in future research.

Also apparent in the table is the role of character names in the narrative uniauthor collections, along with their large token counts (e.g., *Anthony*, 150; *Laura*, 198; *Pa*, 192; *Kayo*, 338; *Rosie*, 327). Even the words *bone* and *breath* in the Mystery Narrative Uniauthor collection refer primarily to the name of a dog (*Bone Breath*), and the word *club* in the same collection refers primarily to a children's club for animals (*Care Club*). Obviously, the repetition of names (people, places, groups, organizations, etc.) is one of the vocabulary advantages of reading fictional stories written by one author (e.g., a series), and it likely explains why the narrative uniauthor collections are ranked higher than their narrative multiauthor counterparts in terms of vocabulary recycling (see Tables 2 and 3). Interestingly, only one of the single-authored

expository collections seems to be impacted by names—the Westward Expository Uniauthor collection. Here, however, one might argue that there are important characteristic differences between names of historical significance (e.g., *John*, *James*, *George*) and names in fictional narratives.

Table 4. *Top 10 specialized types (with token counts) appearing in at least 3 of 4 texts in each collection (based on 5,000 tokens per text, i.e., 20,000 per collection)*

MUMMY_EXP_MA		MUMMY_EXP_UA		MUMMY_NAR_MA		MUMMY_NAR_UA	
MUMMY	166	EGYPT	189	HALL	46	ANTHONY	150
MUMMIES	161	PYRAMID	124	HAIR	18	LOT	19
EGYPTIANS	77	PYRAMIDS	80	EGYPT	15	DESK	17
EGYPT	56	EGYPTIANS	72	LOT	13	FINALLY	14
EGYPTIAN	50	BC	65	EGYPTIAN	10	LIKED	12
TOMBS	46	EGYPTIAN	51	ANGRY	10	SUDDENLY	12
PYRAMID	46	TOMBS	42	JOB	10	TEA	12
BURIED	45	DYNASTY	41	SUDDENLY	9	JOB	11
TOMB	43	NILE	41	FINALLY	9	BIT	10
PRESERVED	40	TOMB	34	KITCHEN	9	CORNER	10
WESTWARD_EXP_MA		**WESTWARD_EXP_UA**		**WESTWARD_NAR_MA**		**WESTWARD_NAR_UA**	
TRAIL	69	AMERICAN	46	WAGON	69	LAURA	198
CATTLE	52	CALIFORNIA	43	PRAIRIE	33	PA	192
WAGON	41	AMERICANS	34	GRASS	29	MA	113
WAGONS	38	AMERICA	29	WAGONS	27	MARY	94
FORT	35	JOHN	25	MA	22	WAGON	87
AMERICAN	33	JAMES	22	INDIANS	20	JACK	44
MISSOURI	30	INDIAN	21	HAIR	15	BROWN	23
INDIANS	28	SMITH	20	SLOWLY	15	PATH	23
TERRITORY	24	GEORGE	17	FEET	13	TALL	23
SAN	23	MISSISSIPPI	17	SUDDENLY	13	FEET	19
MYSTERY_EXP_MA		**MYSTERY_EXP_UA**		**MYSTERY_NAR_MA**		**MYSTERY_NAR_UA**	
BONES	67	ISLAND	59	SHOP	21	KAYO	338
JOHN	66	FEET	33	FAT	18	ROSIE	327
BONE	51	FINALLY	28	JEANS	16	CAT	99
EVIDENCE	51	CREATURE	18	FUNNY	11	BREATH	62
SKULL	31	AMERICAN	13	JACKET	11	BONE	60
TEETH	17	BURIED	13	CORNER	10	SAMMY	55
CLUES	15	HORROR	13	HAIR	10	CLUB	49
HAIR	14	MAD	13	LOT	10	SAUNDERS	35
FOOT	14	EVIL	12	SUDDENLY	9	HOMER	33
PHYSICAL	13	HOLE	12	PROBABLY	8	BENTON	31
CONTROL_EXP_MA				**CONTROL_NAR_MA**			
FEET	32			FEET	21		
AMERICA	16			CLIFF	20		
AREAS	14			HAIR	18		
AREA	12			FOOT	15		
PACIFIC	12			KITCHEN	13		
HUGE	10			SUDDENLY	12		
OCEAN	10			RUG	11		
CALIFORNIA	8			NOSE	10		
ALASKA	8			SKIN	10		
ICE	7			CHAIR	10		

Note. EXP = Expository; NAR = Narrative; MA = Multiauthor; UA = Uniauthor.

"Theme" Word Rating Comparison

Observed differences in the thematic nature of the specialized words prompted a subsequent in-

depth analysis of this variable. Two master's-level linguists, not associated with the study, were given an alphabetized list of the specialized words appearing in all 14 collections (759 distinct types, some of which were shared between collections) and were asked to independently rate each word based on the following criteria:

1. The word definitely or possibly belongs to the Mummy theme
2. The word definitely or possibly belongs to the Mystery theme
3. The word definitely or possibly belongs to the Westward Movement theme
4. The word does not seem to belong to any of the three themes

The raters were instructed that a given word could be rated only once per theme, for example, *definitely* or *possibly*, but not both. However, raters could rate a particular word as definitely or possibly belonging to more than one theme (e.g., the word *burial* was rated as possibly Mummy, possibly Mystery, and possibly Westward Movement). Specialized words from the two control collections (Narrative and Expository) were also included in the list to serve as controls for the rating process. It was assumed that most of the specialized words in the control collections would be rated as not belonging to any theme. The raters used dictionaries and internet searching browsers (e.g., *Google*) to gain understanding of specialized words they were not familiar with in the lists (e.g., *Abusir*—a site in ancient Egypt). Initial interrater agreement was 56.7%, with 430 of 759 total words rated exactly the same. The raters subsequently compared their ratings; differences were worked out through discussion and by collapsing the *definitely* and *possibly* distinction into one category—In Theme. Because the purpose of the ratings was to arrive at a consensus for each word, initial ratings were only used as points of reference for discussion purposes.

The words in the alphabetized list along with their ratings were then assigned back to the original themes from which they were drawn, and appropriate descriptive calculations were run. Table 5 displays the results of the thematic ratings for all 12 thematic collections and for the two control collections. The order of the theme-based collections is based on highest-to-lowest percentage of specialized types that were rated as being In Theme. Note that the results of this analysis highlight a divide (see the broken line in Table 5) that coincides with the natural distinction between the two macroregisters. Several important conclusions can be drawn from the figures in Table 5:

1. Percentages of theme-related specialized words are higher in the expository collections than in the narrative collections, with a high of 51.2% in Mummy Expository Uniauthor and a low of 3.6% in Mystery Narrative Multiauthor.

2. The percentages of thematic types among the expository collections follow the taxonomy of textual relationships perfectly—that is, Mummy expositories (tight theme) have the highest percentages of thematic types, followed by Westward Movement expositories (semitight theme), and then Mystery expositories (loose theme). The same is not true of the narrative collections.

3. The number of theme-related specialized words in the narrative collections is abysmal, with the best case being 14 of 72 (Westward Narrative Multiauthor), and the worst case

being 2 of 56 (Mystery Narrative Multiauthor).

Table 5. *Theme-word-rating comparisons by theme-based collections and control collections*

Theme	Register	Authorship	Total Types	Number and percentage (%)					
				In Theme		Not in Theme		In Different Theme	
Theme-Based Collections									
Mummy	Expository	Uniauthor	164	84	51.2	70	42.7	10	6.1
Mummy	Expository	Multiauthor	141	66	46.8	69	48.9	6	4.3
Westward	Expository	Multiauthor	93	40	43.0	49	52.7	4	4.3
Westward	Expository	Uniauthor	80	31	38.8	45	56.3	4	5.0
Mystery	Expository	Uniauthor	54	13	24.1	36	66.7	5	9.3
Mystery	Expository	Multiauthor	72	17	23.6	53	73.6	2	2.8
Westward	Narrative	Multiauthor	72	14	19.4	57	79.2	1	1.4
Westward	Narrative	Uniauthor	108	11	10.2	94	87.0	3	2.8
Mummy	Narrative	Multiauthor	52	4	7.7	47	90.4	1	1.9
Mystery	Narrative	Uniauthor	90	4	4.4	80	88.9	6	6.7
Mummy	Narrative	Uniauthor	93	4	4.3	81	87.1	8	8.6
Mystery	Narrative	Multiauthor	56	2	3.6	53	94.6	1	1.8
Control Collections									
Control	Narrative	Multiauthor	58	NA	NA	55	94.8	3	5.2
Control	Expository	Multiauthor	28	NA	NA	21	75.0	7	25.0

Note. Words in the *In Different Theme* category of the theme-based collections were judged by raters to belong to a theme that they did not actually appear in, whereas words in this same category of the control collections were judged by raters to belong to a theme even though no thematic relationships were assumed in the control collections.

The data for the control collections provide useful validation of the subjective ratings themselves. As was expected, most of the specialized words in the narrative control collection (no theme) were rated as not belonging to any of the themes (55 of 58, i.e., 94.8%). The three narrative control words judged to be theme related were *cliff* (Mystery), *screamed* (Mystery), and *stomach* (Mummy). While only 75% (21 of 28) of the specialized words in the expository control were rated as not belonging to any theme, a visual inspection of the seven that were judged to be thematic suggests why this is the case: *Alaska*, *America*, *Atlantic*, *California*, *Louisiana*, *Oregon*, and *Pacific*. All seven were rated as definitely or possibly belonging to the Westward Movement theme. Overall, it appears that the raters' intuitions were consistent with the expectation that most of the words in the control collections would be rated as not belonging to a theme, thus providing a secondary validation of the rating accuracy itself.

General Summary and Extensions to Pedagogy

The primary purpose of this study was to investigate the vocabulary-recycling claims of narrow reading, namely, that authentic reading materials related by a common theme or written by a single author will recycle content vocabulary more efficiently than unrelated materials, thus

providing English language learners with more exposure to such items for potential acquisition. In this regard, the linguistic findings of the study suggest the following five conclusions:

1. Themes have their greatest impact on specialized vocabulary recycling among authentic informational (expository) materials, with little or no impact among authentic fictional (narrative) materials. While all theme-based expository collections exhibited more specialized vocabulary recycling (and by quite a large margin) than the expository control collection (no thematic relationship), the same was not true with the narrative collections, where two collections actually exhibited less specialized vocabulary recycling than the narrative control.

2. Tighter themes draw together and recycle specialized vocabulary more efficiently among authentic expository materials than looser themes, but relative thematic tightness has no bearing on specialized vocabulary recycling in authentic narrative collections. While the trend in specialized vocabulary recycling among the expositories followed the proposed taxonomy of textual relationships perfectly (i.e., more specialized vocabulary recycling in the tight Mummy theme, followed by the semitight Westward Movement, followed by the loose Mystery, followed by the Control), the same was not true among the narratives, with one loose Mystery collection ranked highest among the seven narrative collections, and one tight Mummy collection actually ranked last among all narratives, even falling below the narrative control collection.

3. The specialized words in the expository collections were more recognizably theme based (i.e., Mummy-related, Westward Movement-related, Mystery-related) than their narrative counterparts. Furthermore, the extent of thematic fit among the specialized vocabularies of the expository collections followed the proposed taxonomy perfectly (i.e., more theme-specific words in the tight Mummy theme, followed by the semitight Westward Movement theme, followed by the loose Mystery theme). No such theme-specific pattern existed among the specialized vocabulary of the narratives, which, with the possible exception of the history-based Westward Movement collections, were essentially *theme-less*.

4. Authentic children's narratives written by the same author have substantially more specialized vocabulary recycling than narratives written by several different authors, but authorship has no observable impact on specialized vocabulary recycling among authentic children's expository materials. While all three of the narrative uniauthor collections exhibited more specialized vocabulary recycling than their four multiauthor counterparts, the same was not true among the expositories, where no such pattern was observed.

5. Advantages in specialized vocabulary recycling among authentic narrative collections written by one author can be largely accounted for by the repetition of names in stories (characters, places, etc.), not thematic content. By far, the highest repetitions of specialized words in the uniauthor narratives were names of characters (*Anthony*, *Laura*, *Pa*, *Ma*, *Mary*, *Kayo*, *Rosie*, etc.).

Taken together, these linguistically based conclusions indicate that narrow reading has some concrete advantages in terms of specialized vocabulary recycling, but that such advantages may be much more complex than previously indicated (e.g., Cho et al., 2005; Day, 1994; Krashen, 1981, 1985; Schmitt & Carter, 2000). For one, the impact of themes and the impact of authorship on specialized vocabulary recycling may not be mutually supportive constructs, appearing instead to be heavily register-sensitive. That is, themes work best for expository collections, and single authorship works best for narrative collections.

Second, the impact of themes and authorship on vocabulary recycling is as much an issue of *what vocabulary* as it is *what register*. English language educators should not assume that because a group of fictional narratives have been classified as *theme based*, they will greatly improve their learners' exposure to theme-based words and concepts, nor should they assume that a fictional series written by a single author (theme-based or not) will provide substantially more specialized vocabulary redundancy than unrelated materials, with the noted exception of character names and places. This should not be taken to mean that there are no additional advantages to narrow fiction reading that might lead to gains in word knowledge (lowering the overall lexical load, increasing motivation to read, etc.). However, one must be somewhat skeptical of the loose claims that such reading leads to increased, repetitive exposures to new or less familiar words.

Conversely, the power of themes to draw together and recycle the specialized vocabulary of expository materials should also be duly noted in many areas of English language education, especially in content-based instruction and English for academic purposes. Indeed, there is irony in the fact that the expository collections, which are not known for being friendly to incidental word learning from context (Anderson, 1996; Coté, Goldman, & Saul, 1998), produce the best conditions for recycling specialized, theme-specific vocabulary, especially when they are related by tighter themes such as Mummy.

Pedagogically, this apparent paradox of improved vocabulary recycling in learner-unfriendly contexts suggests that any vocabulary recycling advantages gained through theme-based expository reading may still need to be augmented by direct vocabulary instruction and more "word consciousness" raising for young L1 and L2 readers (Graves, 2006; Zahar et al., 2001). In other words, the findings regarding narrow (tight-themed) expository reading should not necessarily be translated into more opportunities for incidental vocabulary acquisition. However, the lexical advantages of narrow expository reading should also be recognized. Because such reading tends to draw together and recycle a greater proportion of specialized words that are more easily identifiable as being theme related (e.g., *mummy*, *pyramid*, *tombs*, *Egypt*, *embalming*, *pharaoh*, *mummification*), several potential advantages accrue for classroom instruction: (a) the reading materials will likely have more connections and relevance to other theme-related classroom discussions or projects; (b) the words in such materials are prime candidates for direct vocabulary instruction because of their salience to both the theme in general and the reading materials that support that theme; and (c) there appears to be a greater likelihood of specialized, theme-based vocabulary redundancy between different expository reading materials from the same tight theme, thus allowing teachers to more confidently choose from a variety of grade-equivalent authentic materials that will recycle many of the same crucial vocabulary items.

Finally, the proposed taxonomy of textual relationships (Figure 1) appears to require some revision based on the disparities in specialized vocabulary recycling noted between the narrative and expository text collections (see Figure 2). The separation of narrative fiction and expository nonfiction in the revised taxonomy reflects the register-related nuances of narrow reading: Number of authors affects specialized vocabulary recycling in collections of narrative fiction, but not in collections of expository nonfiction; conversely, thematic relationships affect specialized vocabulary recycling in collections of expository nonfiction, but not in collections of narrative fiction. The revised taxonomy also reflects the general advantages of narrow reading over wide reading in terms of exposing young readers to repetitive encounters with new or less familiar vocabulary, and in reducing the overall lexical load placed on such readers.

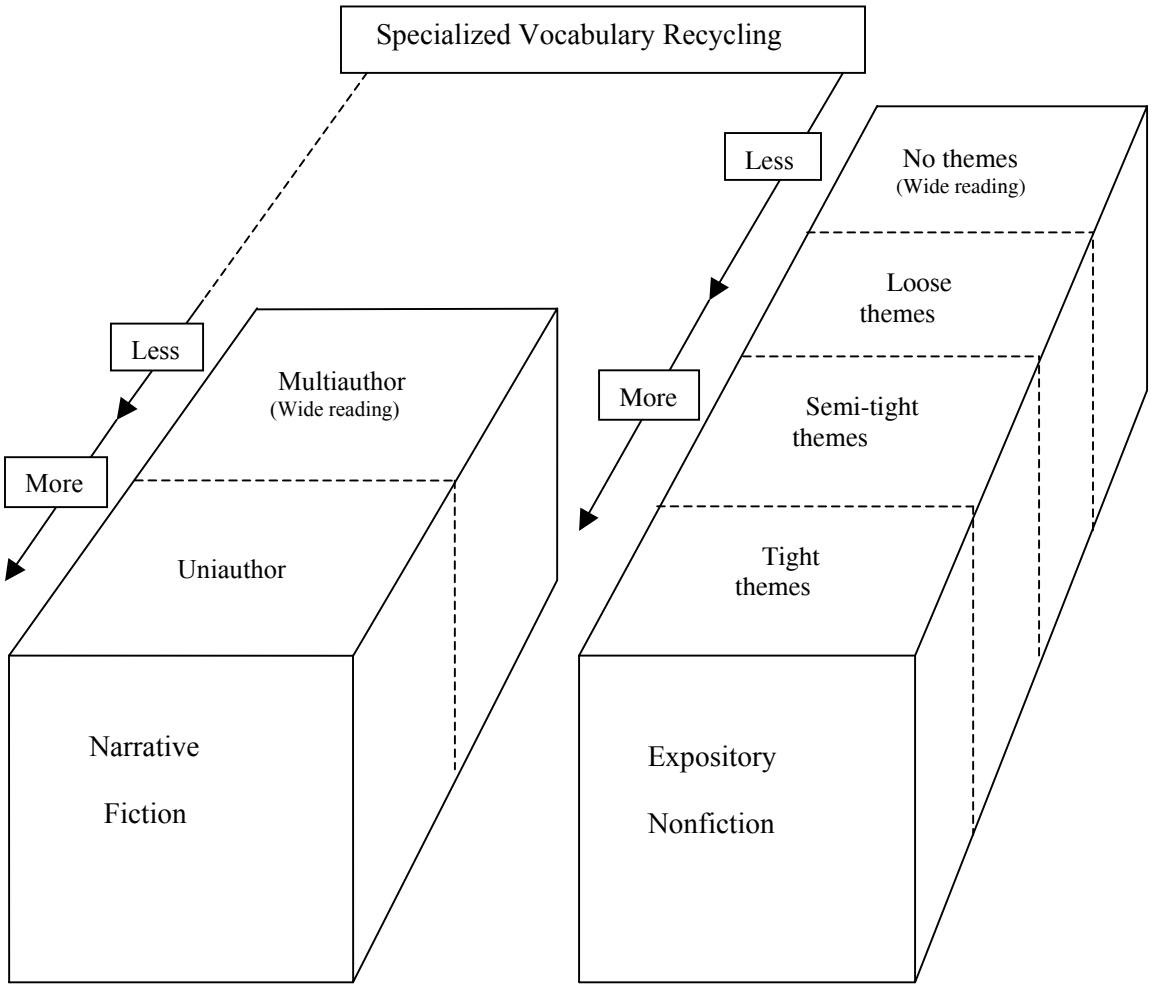

Figure 2. Revised taxonomy of textual relationships for specialized vocabulary recycling in collections of authentic reading materials.

Conclusion

More research is needed to validate the findings of this study. Future research could include different themes, texts, and controls, perhaps using a more randomized selection of materials. In

addition, the definition of *word* used in this study (i.e., unique spellings) is open to challenge on several grounds, including the potential for children to link morphologically related words and the potential presence of multiword items, homonymy, and polysemy. However, there was no visual indication in the data to suggest that any of these variables would have altered the primary findings of the study, which clearly suggest major vocabulary-recycling differences based on the textual relationships between authentic reading materials. For example, Appendix B provides evidence of rich morphological relationships between several of the specialized words in the theme-based expository collections.

There would also appear to be some obvious strengths in the type of corpus-based research conducted in this study, particularly in bringing some degree of accountability to the broad claims of new vocabulary exposure through extensive reading of authentic materials, which have often lacked the nuances necessary to make them truly informative for pedagogical purposes. All too often, people are simply told that young readers must read widely or extensively in class or at home in order to substantially grow their vocabularies, without careful consideration of the types of materials that they could possibly read, the types of words that such reading will expose them to, and the varying levels of essential lexical redundancy that come about as a result of choices regarding which text to read first, second, third, and so forth. In fact, the findings of this corpus-based study suggest that even the more pedagogically focused approach of narrow reading, with its assumed vocabulary-exposure benefits, may have been greatly oversimplified. In short, there is no adequate substitute for real data, in this case, actual words and actual word repetitions in actual authentic reading materials.

It should be emphasized, however, that this study examined specialized, content-rich vocabulary (i.e., the types of words that are representative of long-term, large-scale vocabulary growth during the school years—*mummy*, *embalming*, *laboratory*, *anthropologist*, *frontier*, etc.), and did not consider the recycling of the relatively small set of high-frequency words (2,000 to 3,000 words) that L2 readers must first gain proficiency with in order to achieve basic reading comprehension and to utilize context as a means of building vocabulary knowledge (Laufer, 1989; Nation, 2001). However, given the nature of general high-frequency words (i.e., they appear in many texts) and given that authentic texts are not controlled for the presentation of vocabulary, it is doubtful that narrow reading would provide any appreciable differences in vocabulary recycling at the high-frequency level either. Future research could examine the verity of this assertion.

Finally, it is hoped that the findings of this study and the revised taxonomy of textual relationships will prove useful to English language teachers and curriculum designers who must ultimately make the choices about what their students will be encouraged to read throughout the course of a term, semester, or school year. If the goal is to improve their learners' repetitive exposure to the vocabulary of school, then consideration of the relationships between texts can make a profound difference. In this regard, narrow reading of theme-related expository materials will provide better conditions for such exposure, provided that the theme is sufficiently tight to draw together and recycle theme-related content words in an advantageous manner—in short, not just any theme will do. For instance, a tight expository-based Mummy theme would be better than a loose Mystery theme, but both would be superior to no theme at all (e.g., wide reading). A tight *bee* theme would be better than a loose *insect* theme, but again both would be better than no

theme at all, and so forth. Given the apparent vocabulary-recycling advantages of such thematic decisions, the logical next step is to test whether young L2 readers can actually utilize these benefits to gain vocabulary knowledge through reading or whether the nature of the words in expository materials, and the concepts they often entail, require extratextual support before they can be acquired. A final point with regard to expository materials is that there are no apparent vocabulary-recycling advantages for utilizing text collections written by one author.

In contrast, narrow reading of fictional storybooks written by one author will improve the chances for specialized vocabulary recycling. However, the lexical advantages to the language learner will likely involve issues of general reading fluency (repetition of character names, places, etc.), rather than repeated exposures to theme-based or content-area vocabulary (i.e., the language of school). Indeed, there is no indication in the data that theme-relatedness plays any facilitative role in specialized vocabulary recycling within collections of narrative fiction.

The marked disparities noted in this study between theme-based narrative and expository collections adds to the growing body of evidence suggesting major differences between children's narrative and expository reading materials and reading experiences (reviewed in Grabe, 2002). In this regard, and in line with Gardner's (2004) assertions, the findings of this study suggest that more attention should be paid to the *what* of reading and vocabulary exposure, not merely the *how much*.

References

America's journal. (1996). New York: Scholastic.

Anderson, R. C. (1996). Research foundations to support wide reading. In V. Greaney (Ed.), *Promoting reading in developing countries* (pp. 55–77). Newark, DE: International Reading Association.

Anderson, R. C., & Nagy, W. E. (1992). The vocabulary conundrum. *American Educator*, *16*(4), 14–19, 44–47.

Bamford, R. A., Kristo, J. V., & Lyon, A. (2002). Facing facts: Nonfiction in the primary classroom. *The NERA Journal*, *38*(2), 8–15.

Bello, M., Fajet, W., Shaver, A. N., Toombs, A. K., & Schumm, J. S. (2003). Basal readers and English language learners: A content analysis study. *Reading Research and Instruction*, *42*(2), 1–16.

Carlisle, J. F. (2000). Awareness of the structure and meaning of morphologically complex words: Impact on reading. *Reading and Writing: An Interdisciplinary Journal*, *12*, 169–190.

Carlo, M. S., August, D., McLaughlin, B., Snow, C. E., Dressler, C., Lippman, D. N., Lively, T. J., & White, C. E. (2004). Closing the gap: Addressing the vocabulary needs of English-language learners in bilingual and mainstream classroom. *Reading Research Quarterly*, *39*, 188–215.

Carroll, J. B., Davies, P., & Richman, B. (1971). *The American heritage word frequency book*. New York: Houghton Mifflin.

Cobb, T. (2007). Computing the vocabulary demands of L2 reading. *Language Learning and Technology*, *11*, 38–63.

Cho, K.-S., Ahn, K.-O., & Krashen, S. (2005). The effects of narrow reading of authentic texts on interest and reading ability in English as a foreign language. *Reading Improvement*, *42*, 58–64.

Coté, N., Goldman, S. R., & Saul, E. U. (1998). Students making sense of informational text: Relations between processing and representation. *Discourse Processes*, *25*(1), 1–53.

Cunningham, P. (1998). The multisyllabic word dilemma: Helping students build meaning, spell, and read "big" words. *Reading and Writing Quarterly: Overcoming Learning Difficulties*, *14*, 189–218.

Cunningham, A. E., & Stanovich, K. E. (2003). Reading matters: How reading engagement influences cognition. In J. Flood, D. Lapp, J. R. Squire, & J. M. Jensen (Eds.), *Handbook of research on teaching the English language arts* (2nd ed.). Mahwah, NJ: Erlbaum.

Day, R. R. (1994). Selecting a passage for the EFL reading class. *English Teaching Forum*, *32*(1), 20–23, 38.

Day, R. R., & Bamford, J. (1998). *Extensive reading in the second language classroom*. Cambridge: Cambridge University Press.

Day, R. R., & Bamford, J. (2002). Top ten principles for teaching extensive reading. *Reading in a Foreign Language*, *14*, 136–141.

Dubin, F. (1986). Dealing with texts. In F. Dubin, D. E. Eskey, & W. Grabe (Eds.), *Teaching second language reading for academic purposes* (pp. 127–160). Reading, MA: Addison-Wesley.

Fountas, I. C., & Pinnell, G. S. (1996). *Guided reading: Good first teaching for all children*. Portsmouth, NH: Heinemann.

Fountas, I. C., & Pinnell, G. S. (1999). *Matching books to readers: Using leveled books in guided reading, K-3*. Portsmouth, NH: Heinemann.

Fountas, I. C., & Pinnell, G. S. (2005). *The Fountas and Pinnell leveled book list, K-8, 2006–2008 edition*. Portsmouth, NH: Heinemann.

Gardner, D. (2004). Vocabulary input through extensive reading: A comparison of words found in children's narrative and expository reading materials. *Applied Linguistics*, *25*, 1–37.

Grabe, W. (2002). Narrative and expository macro-genres. In A. Johns (Ed.), *Genre in the classroom: Multiple perspectives* (pp. 249–267). Mahwah, NJ: LEA.

Graves, M. F. (2006). *The vocabulary book: Learning & Instruction*. Williston, VT: Teachers College Press.

Heatley, A., Nation, I. S. P., & Coxhead, A. (2002). RANGE and FREQUENCY programs [Computer software]. Retrieved from http://www.vuw.ac.nz/lals/staff/paul-nation/RANGE32.zip

Hirsch, D., & Nation, I. S. P. (1992). What vocabulary size is needed to read unsimplified texts for pleasure. *Reading in a Foreign Language*, *8*, 689–696.

Huang, H-T., & Liou, H-C. (2007). Vocabulary learning in an automated graded reading program. *Language Learning & Technology*, *11*, 64–82.

Hwang, K., & Nation, I. S. P. (1989). Reducing the vocabulary load and encouraging vocabulary learning through reading newspapers. *Reading in a Foreign Language*, *6*, 323–335.

It's a mystery. (1996). New York: Scholastic.

Krashen, S. (1981). The case for narrow reading. *TESOL Newsletter*, *15*(6), 23.

Krashen, S. (1985). *The input hypothesis: Issues and implications*. New York: Longman.

Krashen, S. (1989). We acquire vocabulary and spelling by reading: Additional evidence for the input hypothesis. *The Modern Language Journal*, *73*, 440–464.

Krashen, S. (1993a). The case for free voluntary reading. *The Canadian Modern Language Review*, *50*(1), 72–82.

Krashen, S. (1993b). *The power of reading: Insights from the research*. Englewood, CO: Libraries Unlimited.

Krashen, S. (2004). The case for narrow reading. *Language Magazine*, *3*(5), 17–19.

Laufer, B. (1989). What percentage of text-lexis is essential for comprehension? In C. Lauren & M. Nordman (Eds.), *Special language: From humans thinking to thinking machines* (pp. 316–323). Clevedon, UK: Multilingual Matters.

Mahony, D., Singson, M., & Mann, V. (2000). Reading ability and sensitivity to morphological relations. *Reading and Writing: An Interdisciplinary Journal*, *12*, 191–218.

McBride-Chang, C., Wagner, R. K., Muse, A., Chow, B. W. Y., & Shu, H. (2005). The role of morphological awareness in children's vocabulary acquisition in English. *Applied Psycholinguistics*, *26*, 415–435.

Mesmer, H. A. E. (2001). Decodable text: A review of what we know. *Reading Research and Instruction*, *40*, 121–142.

Nagy, W. E. (1997). On the role of context in first- and second-language vocabulary learning. In N. Schmitt & M. McCarthy (Eds.), Vocabulary: Description, acquisition and pedagogy (pp. 64–83). Cambridge: Cambridge University Press.

Nagy, W. E., & Anderson, R. C. (1984). How many words are there in printed school English? *Reading Research Quarterly*, *19*, 304–330.

Nagy, W. E., Anderson, R. C., & Herman, P. A. (1987). Learning word meanings from context during normal reading. *American Educational Research Journal*, *24*, 237–270.

Nagy, W. E., & Herman, P. (1985). Incidental vs. instructional approaches to increasing reading vocabulary. *Educational Perspectives*, *23*, 16–21.

Nagy, W. E., & Herman, P. (1987). Breadth and depth of vocabulary knowledge: Implications for acquisition and instruction. In M. McKeown & M. Curtis (Eds.), *The nature of vocabulary acquisition* (pp. 19–35). Hillsdale, NJ: Lawrence Erlbaum.

Nagy, W. E., Herman, P. A., & Anderson, R. C. (1985). Learning words from context. *Reading Research Quarterly*, *20*, 233–253.

Nation, I. S. P. (1990). *Teaching and learning vocabulary*. New York: Newbury House Publishers.

Nation, I. S. P. (1997). The language learning benefits of extensive reading. *The Language Teacher*, *21*(5), 13–16.

Nation, I. S. P. (2001). *Learning vocabulary in another language.* Cambridge: Cambridge University Press.

Nation, I. S. P., & Hwang, K. (1995). Where would general service vocabulary stop and special purposes vocabulary begin? *System*, *23*, 35–41.

Paulson, K. J. (1994). Westward movement: A fifth grade theme. In S. A. Walmsley, *Children exploring their world: Theme teaching in the elementary school* (pp. 217–242). Portsmouth, NH: Heinemann.

Ravin, Y., & Leacock, C. (2000). Polysemy: An overview. In Y. Raven & C. Leacock (Eds.), *Polysemy: Theoretical and computational approaches* (pp. 1–29). Oxford: Oxford University Press.

Schmitt, N., & Carter, R. (2000). The lexical advantages of narrow reading for second language learners. *TESOL Journal*, *9*(1), 4–9.

Schmitt, N., & Zimmerman, C. (2002). Derivative word forms: What do learners know? *TESOL*

Quarterly, *36*, 145–171.
Shu, H., Anderson, R. C., & Zhang, H. (1995). Incidental learning of word meanings while reading: A Chinese and American cross-cultural study. *Reading Research Quarterly*, *30*(1), 76–95.
Singson, M., Mahony, D., & Mann, V. (2000). The relation between reading ability and morphological skills: Evidence from derivational suffixes. *Reading and Writing: An Interdisciplinary Journal*, *12*, 219–252.
Stahl, S. A., & Shiel, T. G. (1992). Teaching meaning vocabulary: Productive approaches for poor readers. *Reading and Writing Quarterly: Overcoming Learning Difficulties*, *8*, 223–241.
Sutarsyah, C., Nation, I. S. P., & Kennedy, G. (1994). How useful is EAP vocabulary for ESP? A corpus based case study. *RELC Journal*, *25*(2), 34–50.
Swanborn, M. S. L., & de Glopper, K. (1999). Incidental word learning while reading: A meta-analysis. *Review of Educational Research*, *69*, 261–285.
Vacca, R. T., & Vacca, J. L. (1996). *Content area reading* (5th ed.). New York: Harper Collins.
Walmsley, S. A. (1994). *Children exploring their world: Theme teaching in the elementary schools*. Portsmouth, NH: Heinemann.
Waring, R. (2003). *The Oxford University Press guide to the why and how of using graded readers*. Tokyo: Oxford University Press.
West, M. (1953). *A general service list of English words*. London: Longman.
Willcox-Schnabl, T. (1994). Egyptian mummies: A sixth grade theme. In S. A. Walmsley, *Children exploring their world: Theme teaching in the elementary school* (pp. 243–270). Portsmouth, NH: Heinemann.
Wodinsky, M., & Nation, I. S. P. (1988). Learning from graded readers. *Reading in a Foreign Language*, *5*, 155–161.
Zahar, R., Cobb, T., & Spada, N. (2001). Acquiring vocabulary through reading: Effects of frequency and contextual richness. *The Canadian Modern Language Review*, *57*, 541–572.

Appendix A

References of Books in Children's Thematic Corpus

Theme Unit Source Key: Source A = (*Egyptian Mummies: A Sixth Grade Theme*, Willcox-Schnabl, 1994); **Source B** = (*Westward Movement: A Fifth Grade Theme*, Paulson, 1994); **Source C** = (*America's Journal*, 1996); **Source D =** (*It's a Mystery*, 1996); **Source E** = (Children's Librarian and Fifth-Grade Teacher)

Books in Thematic Collections

Mummy Narrative Multiauthor

Masterman-Smith, V. (1982). *The great Egyptian heist*. New York: Four Winds Press. **[Source A]**
Peck, R. (1986). *Blossom Culp and the sleep of death*. New York: Bantam Doubleday Dell. **[Source A]**
Snyder, Z. K. (1967). *The Egypt game*. New York: Bantam Doubleday Dell. **[Source A]**
Voight, C. (1991). *The Vandemark mummy*. New York: Fawcett Juniper. **[Source A]**

Mummy Narrative Uniauthor

Bellairs, J. (1978). *The treasure of Alpheus Winterborn*. New York: Harcourt Brace Jovanovich. **[Source E]**
Bellairs, J. (1983). *The mummy, the will, and the crypt*.

New York: Dial Books for Young Readers. **[Source B]**

Bellairs, J. (1984). *The dark secret of Weatherend*. New York: Dial Books for Young Readers. **[Source E]**

Bellairs, J. (1996). *The curse of the blue figurine* (rev. ed.). New York: Puffin Books. **[Source B]**

Mummy Expository Multiauthor

Bendick, J. (1989). *Egyptian tombs*. New York: Franklin Watts. **[Source A]**

Lauber, P. (1985). *Tales mummies tell*. New York: Thomas Y. Crowell. **[Source A]**

Putnam, J. (1993). *Eyewitness books: Mummy*. New York: Alfred A. Knopf. **[Source A]**

Wilcox, C. (1993). *Mummies & their mysteries*. Minneapolis: Carolrhoda Books. **[Source A]**

Mummy Expository Uniauthor

Millard, A. (1982). *Ancient Egypt*. London: Granada. **[Source A]**

Millard, A. (1987). *Great civilizations: Egypt 3118 BC-AD 642*. New York: Franklin Watts. **[Source A]**

Millard, A. (1995). *Mysteries of the pyramids*. Brookfield, Connecticut: Copper Beech Books. **[Source E]**

Millard, A. (1996). *Pyramids*. New York: Kingfisher. **[Source E]**

Westward Movement Narrative Multiauthor

Conrad, P. (1985). *Prairie songs*. New York: Harper & Row, Publishers. **[Source B]**

Fleischman, S. (1988). *By the great horn spoon* (rev. ed.). Boston: Little, Brown and Company. **[Source B]**

Lawlor, L. (1986). *Addie across the prairie*. Niles, IL: Albert Whitman & Company. **[Source B]**

Moeri, L. (1994). *Save Queen of Sheba* (rev. ed.). New York: Puffin Books. **[Source B]**

Westward Movement Narrative Uniauthor

Wilder, L. I. (1971a). *Little house in the big woods* (rev. ed.). New York: Harper & Row. **[Source E]**

Wilder, L. I. (1971b). *Little house on the prairie* (rev. ed.). New York: Harper & Row. **[Source E]**

Wilder, L. I. (1971c). *Farmer boy* (rev. ed.). New York: Harper & Row. **[Source E]**

Wilder, L. I. (1971d). *On the banks of Plum Creek* (rev. ed.). New York: Harper & Row. **[Source E]**

Westward Movement Expository Multiauthor

Blumberg, R. (1989). *The great American gold rush*. New York: Bradbury Press. **[Source B]**

Freedman, R. (1983). *Children of the wild west*. New York: Scholastic. **[Sources B & C]**

Sandler, M. W. (1994). *Cowboys*. New York: Harper Collins. **[Sources B & C]**

Tunis, E. (1961). *Frontier living (chapters 9–19)*. Cleveland: The World Publishing Company. **[Source B]**

Westward Movement Expository Uniauthor

Hakim, J. (1994a). *A history of US (Book 5): Liberty for all (chapters 1–9)*. New York: Oxford University Press. **[Source E]**

Hakim, J. (1994b). *A history of US (Book 5): Liberty for all (chapters 10–18)*. New York: Oxford University Press. **[Source E]**

Hakim, J. (1994c*). A history of US (Book 5): Liberty for all (chapters 19–27)*. New York: Oxford University Press. **[Source E]**

Hakim, J. (1994d). *A history of US (Book 5): Liberty for all (chapters 28–36)*. New York: Oxford University Press. **[Source E]**

Mystery Narrative Multiauthor

Brenner, B. (1972). *Mystery of the plumed serpent* (rev. ed.). New York: Alfred A. Knopf. **[Source D]**

Elmore, P. (1992). *Susannah and the purple mongoose mystery*. New York: Dutton Children's Books. **[Source D]**

Konigsburg, E. L. (1967). *From the mixed-up files of Mrs. Basil E. Frankweiler.* New York: Atheneum. **[Source D]**

Wortis, A. (1991). *Windcatcher*. New York: Avon Books. **[Source D]**

Mystery Narrative Uniauthor

Kehret, P. (1995a). *Frightmares: Cat burglar on the prowl*. New York: Pocket Books. **[Source E]**

Kehret, P. (1995b). *Frightmares: Don't go near Mrs. Tallie*. New York: Pocket Books. **[Source E]**

Kehret, P. (1996a). *Frightmares: Backstage fright*. New York: Pocket Books. **[Source E]**

Kehret, P. (1996b). *Frightmares: Screaming eagles*. New York: Pocket Books. **[Source E]**

Mystery Expository Multiauthor

Beattie, O., & Geiger, J. (1992). *Buried in ice*. New York: Scholastic. **[Source D]**

Bisel, S. C. (1990). *The secrets of Vesuvius*. New York: Scholastic. **[Source D]**

Jackson, D. M. (1996). *The bone detectives*. Boston: Little, Brown and Company. **[Source D]**
Sheely, R. (1993). *Police lab: Using science to solve crimes*. Silver Moon Press. **[Source D]**

Mystery Expository Uniauthor

Simon, S. (1976). *Ghosts*. Philadelphia: J. B. Lippincott. **[Source E]**
Simon, S. (1979). *Creature from lost worlds*. New York: J. B. Lippincott. **[Source E]**
Simon, S. (1981). *Mad scientists, weird doctors, & time travelers in movies, TV, & books*. New York: J. B. Lippincott. **[Source E]**
Simon, S. (1997). *Strange mysteries from around the world* (rev. ed). New York: Morrow Junior Books. **[Source E]**

Books in Nonthematic Collections

Control Narrative Collection

L'Engle, M. (1962). *A wrinkle in time*. New York: Farrar, Straus, and Giroux. **[Source E]**
O'Dell, S. (1960). *Island of the blue dolphins*. New York: Dell Publishing. **[Source E]**
Paterson, K. (1977). *Bridge to Terabithia*. New York: Harper & Row. **[Source E]**
Raskin, E. (1978). *The Westing game*. New York: Viking Penguin. **[Source E]**

Control Expository Collection

Dow, L. (1990). *Whales: A great creature of the world book*. New York: Weldon Owen Pty Limited. **[Source E]**
Heinrichs, A. (1992). *America the beautiful: Montana* (2nd ed.). Chicago: Children's Press. **[Source E]**
Maestro, B., & Maestro, G. (1996). *The voice of the people: American democracy in action*. New York: Lothrop, Less & Shepard Books. **[Source E]**
Ride, S., & Okie, S. (1986). *To space & back*. Orlando, FL: Harcourt Brace. **[Source E]**

Appendix B

Specialized Types (With Token Counts) Appearing in At Least Three of Four Texts per Collection
(Words marked with asterisks were rated as being In Theme)

Mummy Exp MA
(141 specialized types)

MUMMY *	166	GRAVE *	17	BRAIN*	9	PROCESS	6
MUMMIES *	161	EMBALMERS*	16	CARVED*	9	SANDS*	6
EGYPTIANS*	77	SAND*	16	JACKAL*	9	SKULL*	6
EGYPT *	56	PHARAOH *	15	SACRED	9	TOE	6
EGYPTIAN *	50	ANUBIS*	15	SEALED	9	CENTURY *	5
TOMBS *	46	HAIR	15	CUSTOMS	8	BOTTOM	5
PYRAMID *	46	WRAPPINGS*	15	EXTRA	8	FREEZING	5
BURIED *	45	ORGANS*	14	GREEK	8	FUR	5
TOMB*	43	BANDAGES*	14	OUTER	8	HORUS *	5
PRESERVED*	40	KHUFU *	14	LOT	8	STATUE	5
LINEN*	37	PRIESTS *	13	MUMMIFICATION*	8	STUFFED	5
PYRAMIDS *	33	DECORATED	12	STATUES	8	UNDERGROUND	5
WRAPPED*	28	ROBBERS	12	WARM	8	WASHED	5
NILE*	25	STRIPS	12	FINALLY	7	WET	5
EMBALMING*	25	EMBALMED*	12	JEWELRY *	7	PRACTICED	4
CAVES*	24	HUGE	12	LAYER	7	CLIFF	4
MUMMIFIED *	23	PRESERVE *	12	FLINT	7	CURIOUS	4
BURIAL*	22	NATRON *	11	GRAVES*	7	DUG	4
PROBABLY	21	JARS*	11	INTESTINES*	7	EUROPE	4
SKIN	21	PERIOD	11	LUNGS*	7	FERTILE *	4
PHARAOHS*	20	CLOTHING	10	PATRON	7	FINGER	4
OSIRIS*	19	REMOVED	10	AREA	6	GOVERNMENT	4
BACTERIA *	18	BITUMEN *	10	CEREMONY*	6	ITEMS	4
CAVE *	18	COFFINS*	10	ARCTIC	6	KNIFE	4
DECAY *	18	PASSAGE	10	CEREMONIES*	6	NOSE	4
COFFIN*	17	SPELLS	10	CLOTHES	6	PERFORMED	4
RESIN *	17	STOMACH*	10	EXACTLY	6	PRESERVING*	4
		TREASURES*	10	HERODOTUS*	6	QUICKLY	4
		WRAPPING *	10	ICY	6	SHELTERS	4
		CLOTH	9	INTERNAL	6	SIMILAR	4
		FEET	9	LAYERS	6	SKELETON *	4
		FUNERAL*	9	LIVER *	6	STICKY	4

TOOLS	4
YARDS	4
BOTHER	3
CIVILIZATION *	3
EVERYDAY	3
EVIL	3
EXACT	3
FEATHERS	3
FINGERS	3
MOLDED	3
NARROW	3
PACKED	3
POURED	3
SAWDUST	3
SERIES	3
SHARP	3
SPICES*	3
WEIGHED	3

Mummy Exp UA
(164 specialized types)

EGYPT *	189
PYRAMID*	124
PYRAMIDS*	80
EGYPTIANS*	72
BC*	65
EGYPTIAN*	51
TOMBS*	42
DYNASTY*	41
NILE*	41
TOMB*	34
PERIOD	25
NUBIA*	25
BLOCKS*	24
GOODS	22
OSIRIS *	21
HORUS*	20
BURIED *	20
GIZA *	20
THEBES*	19
PRIESTS *	18
DYNASTIES*	16
RAMP	16
CHAMBER*	16
UPPER	16
MUD	15
PAPYRUS*	15
COPPER	14
TEXTS	14
HIEROGLYPHS*	14
III	14
KHUFU *	14
MONUMENTS*	13
MEMPHIS*	13
THRONE *	13
FLOOD	12
STRAIGHT	12
BURIAL *	12
MUMMIES*	12
SCRIBES*	12
LINEN*	11
PRESERVED*	11
SAND *	11
GOVERNMENT	11
ROBBERS	11
BRICKS	10
INUNDATION	10
ISIS *	10
GODDESSES*	9
CARVED*	9
HEIGHT	9
MEDITERRANEAN*	9
PHARAOH *	9
SKILLED	9
STRAIGHT-SIDED	9
ARCHITECT*	8
DECORATED	8
DRAGGED	8
GRAIN	8
PROBABLY	8
ASWAN*	8
BLOCK *	8
PRAYERS	8
PUNT	8
QUARRY *	8
REEDS *	8
ZOSER*	8
COMPLEX	7
DIVINE *	7
HUGE	7
II	7
REED *	7
AFFORD	7
BENT	7
CENTURIES *	7
CENTURY*	7
CROPS	7
GEB*	7
HOLY *	7
INVENTED	7
MAGIC	7
NUT	7
SENUSRET*	7
TREASURES*	7
FINALLY	6
IV	6
WRAPPED *	6
ANUBIS *	6
CEREMONY *	6
CIVILIZATION *	6
DEIR*	6
ETERNAL*	6
FEET	6
FERTILE *	6
RAMPS	6
SHU*	6
SURVIVED	6
TIMBER	6
FLOODED	5
JARS*	5
POTTERY	5
QUARRIES*	5
AREA	5
BRICK	5
COURTIERS*	5
DEMOTIC*	5
EDUCATION	5
EL*	5
ETERNITY*	5
GUARDED	5
IMHOTEP*	5
MASTABAS*	5
PASSAGE	5
PERIODS	5
ROOF	5
SCRIBE*	5
STATUE	5
SKILL	4
VAST	4
ABUSIR *	4
ACCOMPANIED	4
AMULETS*	4
ARCHAIC*	4
BANDAGES *	4
CALENDAR	4
CULTURE	4
DOUBLE	4
FOOT	4
FOREVER	4
HATHOR*	4
IVORY	4
LEBANON	4
MASTABA*	4
NARROW	4
OXEN	4
POURED	4
PROFESSIONAL	4
REGULAR	4
ROMANS *	4
TREATED	4
BASKETS *	3
BEER	3
CANALS	3
CEILING	3
DESCENDED	3
DIG *	3
DITCHES	3
DONKEYS*	3
EMPTY	3
FAMINE *	3
GOVERNORS	3
HIERATIC *	3
HIEROGLYPHIC*	3
HUNI*	3
PEASANT	3
PLANETS	3
PTOLEMY*	3
RAINS	3
RUBBLE	3
SMOOTH	3
STEEP	3
TALLER	3
TREASURE *	3
VICTIM	3
WEAPONS	3

Mummy Nar MA
(52 specialized types)

HALL	46
HAIR	18
EGYPT *	15
LOT	13
EGYPTIAN*	10
ANGRY	10
JOB	10
SUDDENLY	9
FINALLY	9
KITCHEN	9
KID	8
THIN	8
EXACTLY	7
HOLE	7
THICK	7
GRINNED	6
BUSY	6
DIRTY	6
DUMB	6
FEET	6
INTERRUPTED	5
AFTERNOON	5
AREA	5
BOTTOM	5
GRAY	5
LIKED	5
NOSE	5
OKAY	5
QUICK	5
SIGHED	5
STUCK	5
ANNOUNCED	4
BUTTON	4
CLOCK	4
LOCKED	4
NARROW	4
NODDED	4
PALE	4
PLENTY	4
PROBABLY	4
SEARCH	4
SLIPPED	4
STARED	4
ATTENTION	3
BURIED *	3
CARVED *	3
HELLO	3
HEY	3
KNEES	3
MESS	3
NICE	3
PICK	3

Mummy Nar UA
(93 specialized types)

ANTHONY	150
LOT	19
DESK	17
FINALLY	14
LIKED	12
SUDDENLY	12
TEA	12
JOB	11
BIT	10
GRAY	10
STARED	10
TALL	10
CORNER	10
NICE	10
CLOCK	9
FUNNY	9
HALL	9
HIT	9
JOHN	9
KITCHEN	9
MYSTERIOUS*	9
SLOWLY	9
CHAIR	8
TOWER	8
CREAM	8
WORRIED	8
CHAPTER	7
FEET	7

GONNA	7
NOSE	7
ODD	7
SICK	7
LEATHER	7
RADIO	7
GRINNED	6
QUICKLY	6
ANGRY	6
CAP	6
CHESS	6
HAIR	6
SCARED	6
SIGH	6
BATHROOM	5
CHEERFUL	5
COFFEE	5
DOORWAY	5
HMMM	5
MUTTERED	5
NODDED	5
OKAY	5
QUIETLY	5
SCREEN	5
SIGHED	5
SNAPPED	5
STUCK	5
AFRAID	4
BROWN	4
CARVED*	4
FLAKES	4
FUN	4
LIQUOR	4
NARROW	4
PAIR	4
POCKET	4
ROOF	4
ROTTEN*	4
SHUT	4
STARING	4
TONE	4
BLOTTER	3
CLUTCHED	3
DIGGING*	3
DIRTY	3
DISCONTENTED	3
DUG	3
ELDERLY	3
GASPED	3
GLOW	3
HIDING	3
KIDS	3
LONELY	3
NEST	3
RIDICULOUS	3
SINISTER	3
SOMEHOW	3
STAIRS	3
STRAIGHT	3
SWEAR	3
SWEPT	3
TONGUE	3
UPSTAIRS	3
WET	3
YEAH	3

Westward Exp MA
(93 specialized types)

TRAIL*	69
CATTLE*	52
WAGON*	41
WAGONS*	38
FORT *	35
AMERICAN *	33
MISSOURI*	30
INDIANS*	28
TERRITORY*	24
SAN	23
OXEN *	23
INDIAN *	22
PRAIRIE*	22
TEXAS *	22
MEXICO*	21
FRONTIER *	21
RANGE *	16
ST	15
AMERICANS *	14
JOB	14
RANCHES *	14
LOUIS	12
MEXICAN*	12
MISSISSIPPI*	12
IOWA *	12
GRASS*	11
AMERICA *	11
SPANISH	11
DUST	10
KANSAS*	10
TRIP	10
FEET	9
JOHN	9
MULES*	9
RAILROAD*	9
HERDS	8
TOUGH	8
GOVERNMENT *	8
JOURNEY*	8
SADDLE	8
SLAVES*	7
ARKANSAS*	6
CONGRESS	6
SHOPS	6
VAST	6
ACRES*	6
CENTS	6
CRUDE	6
FOOT	6
FRAME	6
KNIVES	6
PATENT	6
PLOW	6
RUSH	6
MAP	5
BOOTS*	5
EQUIPMENT	5
FENCES	5
HORN	5
HORSEBACK*	5
LOADED	5
OCCUPIED	5
PACK	5
PLENTY	5
RIFLES *	5
RIO*	5
SLAVERY*	5
SOIL	5
WYOMING*	5
CORNER	4
ESPECIALLY	4
FERTILE	4
GULF	4
PICK	4
PUEBLO*	4
QUICK	4
WEATHER	4
BUSY	3
CAMP*	3
CIVIL	3
CLAY	3
EL	3
EXTRA	3
MALE	3
PICKED	3
QUICKLY	3
RARE	3
RAW	3
SITE	3
STRAIGHT	3
TIN	3
TWENTY-FOUR	3
UNSETTLED*	3

Westward Exp UA
(80 specialized types)

AMERICAN *	46
CALIFORNIA*	43
AMERICANS*	34
AMERICA*	29
JOHN	25
JAMES	22
INDIAN*	21
SMITH	20
GEORGE	17
MISSISSIPPI*	17
SLAVE*	16
ENGLAND	14
TERRITORY*	14
WASHINGTON*	14
YORK	13
ST	11
TH	11
CENTURY	11
STEAMBOAT *	11
JEFFERSON*	10
PARENTS	10
ROUTE	10
MISSOURI*	9
SIERRA*	9
TRIP	9
FRONTIER*	8
SLAVERY*	8
TALL	8
FORT*	8
LOUIS	8
OREGON*	8
SPANISH	8
ESPECIALLY	7
HAIR	7
HENRY	7
OHIO*	7
DEMOCRACY *	7
GOVERNMENT*	7
OCEAN	7
TH-CENTURY	7
LIBERTY*	6
ORLEANS*	6
CLOTHES	6
PHILADELPHIA*	6
POET	6
REGION	6
SAMUEL	6
WILLIAM	6
ACADEMY	5
FOUNDED	5
SPAIN	5
THOMAS	5
CAPE	4
FINALLY	4
HARVARD	4
HORN	4
JOURNEY*	4
LIKED	4
MIGHTY	4
OVERLAND*	4
PAIR	4
TERRITORIES*	4
AFRAID	3
ATLANTIC*	3
BOTHERED	3
BROWN	3
CLOTHING	3
EUROPEAN	3
FEET	3
HIRED	3
HORRIFIED	3
KANSAS*3	
LOT	3
PENNSYLVANIA*	3
PIONEERS*	3
PLENTY	3
POETRY	3
PROFESSOR	3
TREATED	3
WAGONS*	3

Westward Nar MA
(72 specialized types)

WAGON*	69
PRAIRIE *	33
GRASS*	29
WAGONS*	27
MA*	22
INDIANS*	20
HAIR	15
SLOWLY	15
FEET	13
SUDDENLY	13
STARED	11
BABY	11
BROWN	11
ALIVE	10
KNEES	9
SHUT	9
BARREL*	9
CORNMEAL*	9
DIRT	9
PILE	8
COW *	8
QUIET	8
QUIETLY	8

SMELL	8
THIN	8
COOL	7
DUST	7
HAT*	7
STRAIGHT	7
WHISPERED	7
HUNGRY	6
QUICKLY	6
COAT	6
GUESS	6
LEGS	6
PALE	6
TALL	6
WORSE	6
YELLOW	5
AFRAID	5
AWFUL	5
CHICKENS*	5
DINNER	5
FEVER *	5
FINGERS	5
INSTANT	5
LOUD	5
RAIN	5
SKIN	5
SLIGHT	5
WARM	4
DELICATE	4
ESPECIALLY	4
FOREHEAD	4
GULPED	4
PICKED	4
RIFLE*	4
SACKS	4
THICK	4
THUMB	4
TWISTED	4
WASHED	4
WHEEL	4
WIPED	4
YELL	4
ANGER	3
BISCUITS*	3
CREPT	3
PLENTY	3
PUSHING	3
SLID	3
TIED	3

Westward Nar UA
(108 specialized types)

LAURA	198
PA*	192
MA*	113
MARY	94
WAGON*	87
JACK	44
TALL	23
BROWN	23
PATH	23
FEET	19
EDGE	18
LEGS	18
SLOWLY	15
CARRIE	15
LOG	14
CHAPTER	13
SUPPER	13
YELLOW	13
CLEAN	12
TINY	12
ROOF	12
CLIMBED*	11
HAIR	11
THICK	11
FLAT	11
STOVE	11
FAT	10
NOSES	10
TIED	10
BARE	9
WARM	9
SMOKE	9
TIRED	9
BIT*	8
EMPTY	8
SNUG	8
THIN	8
TIN	8
BOTTOM	8
HOLE	8
PORK	8
QUICKLY	8
SHUT	8
TIGHT	8
CLIMB *	7
CURVED	7
PALE	7
PAN	7
PLATE	7
SLOPE	7
WET	7
OVERHEAD	6
WASHED	6
CRACK	6
FOOT	6
HURRIED	6
HURT	6
MITTENS	6
QUIET	6
SUDDENLY	6
SMOOTH	5
TAILS	5
TONGUE	5
AX *	5
BITS*	5
DISHES	5
FENCE	5
FUN	5
GRAY	5
GREASED	5
PUSHED	5
SLICES	5
STEEP	5
TONIGHT	5
BOILED	4
CLOTHES	4
FRIGHTENED	4
INDIANS*	4
KNEES	4
LOGS	4
LONELY	4
MANES*	4
RAIN	4
ROPES	4
STRETCHED	4
WHISPERING	4
ACHED	3
ARCHED	3
BREATH	3
CIDER *	3
COMBED	3
COMFORTABLE	3
EXCITED	3
FUR	3
HALF-PINT	3
HUGE	3
MELTED	3
NEATLY	3
OAK	3
PICKED	3
POURED	3
ROOTS	3
SHOUTED	3
SMELLED	3
STOMACH	3
STUCK	3
TONGUES	3
TROTTED	3

Mystery Exp MA
(72 specialized types)

BONES*	67
JOHN	66
BONE*	51
EVIDENCE*	51
SKULL*	31
TEETH	17
CLUES*	15
HAIR	14
FOOT	14
PHYSICAL	13
FILE	13
SOLVE*	13
EXAMINING	12
CRIMINALS *	12
EXAMINE	11
TINY	11
RIDGES	11
AREA	10
FEET	10
INVESTIGATION *	10
LEG	10
WEIGHT	10
WILLIAM	10
SUDDENLY	9
ANTHROPOLOGIST	9
CHAPTER	9
SKELETON*	9
FINALLY	8
BAG	8
EXAMINED	8
NARROW	8
QUICKLY	8
SEARCH*	8
BURIED *	7
COLLECTED	7
LABORATORY*	7
LOT	7
MALE	7
SKIN	7
BENT	6
NODDED	6
PLENTY	6
PROBABLY	6
PROCESS	6
RESEARCH	6
SMOOTH	6
UPPER	6
ITEMS	5
TALL	5
CHECK	5
PICKED	5
QUIET	5
WASHINGTON	5
ARCHES	4
BOTTOM	4
CRACK*	4
EMPTY	4
INVOLVES	4
JOB	4
SEARCHES*	4
SHOUTED	4
CORNER	3
COUPLE	3
DAMAGE	3
GRISLY*	3
HIT	3
IMMEDIATELY	3
INVESTIGATE*	3
RIDGE	3
SCATTERED	3
SEARCHING*	3
TOOL	3

Mystery Exp UA
(54 specialized types)

ISLAND	59
FEET	33
FINALLY	28
CREATURE*	18
AMERICAN	13
BURIED*	13
HORROR*	13
MAD	13
EVIL*	12
MYSTERIOUS*	12
HOLE	12
HUGE	11
LABORATORY*	11
ORIGINAL	11
PROFESSOR*	9
FRIGHTENING*	8
LONDON	8
SUDDENLY	8
YORK	8
EMPTY	7
FEATURES	7
HIDDEN*	7
SERIES	7
IMAGINE	6
BURIAL *	6
FINAL	6
FOOTPRINTS*	6
HORRIBLE*	6
ODD	6
PROBABLY	6
PUBLISHED	6
THOMAS	6
WEIRD	6
CANNOT	5
EXCITING	5
EXPEDITION	5

JACK	5
REPTILES	5
TERROR *	5
ENGLAND	4
ROBERT	4
BOTTOM	4
BRITISH	4
CAMERA	4
ELECTRIC	4
ENTIRE	4
HIT	4
JOHN	4
PLANE	4
CONFUSED	3
RID	3
SHELTER	3
TERRIBLE	3
WORSE	3

Mystery Nar MA
(56 specialized types)

SHOP	21
FAT	18
JEANS	16
FUNNY	11
JACKET	11
CORNER	10
HAIR	10
LOT	10
SUDDENLY	9
PROBABLY	8
NOISE	8
AFTERNOON	7
FEET	7
CARD*	7
PICK	7
PICKED	7
SMELL	7
GRABBED	6
BOTTOM	6
GLANCED	6
WHISPERED	6
HEY	5
KID	5
LEGS	5
LOTS	5
PILE	5
POCKETS*	5
PUSHED	5
HESITATED	4
CLEAN	4
COOL	4
DOORWAY	4
FADED	4
FINALLY	4
GUESS	4
HI	4
PARK	4
PRACTICE	4
QUIET	4
REPEATED	4
SIGHED	4
SLIPPED	4
SLOWLY	4
STRAIGHT	4
WORRIED	4
BUSY	3
CRAZY	3
ESPECIALLY	3
FINAL	3
GRINNING	3
INFORMED	3
MINIATURE	3
NERVOUS	3
TALL	3
WORST	3
YELLOW	3

Mystery Nar UA
(90 specialized types)

KAYO	338
ROSIE	327
CAT	99
BREATH	62
BONE*	60
SAMMY	55
CLUB	49
SAUNDERS	35
HOMER	33
BENTON	31
DIAMOND	28
WEBSTER	23
FEET	20
BASEBALL	18
POLICE	17
FUR	14
PROJECT	13
KIDS	13
PROBABLY	12
PARENTS	12
NOSE	11
OAKWOOD	11
NOTEBOOK	10
QUICKLY	10
WHISPERED	10
EXTRA	9
VOCABULARY	9
CUSHMAN	9
UNDERSIGNED	8
WHEREAS	8
APARTMENT	8
LOCKED	8
PET	8
AFTERNOON	7
HURRIED	7
NODDED	7
POCKET*	7
REMOVED	7
TAIL	7
KITCHEN	7
LOT	7
TELEPHONE	7
QUIT	6
RELIEF	6
SORRY	6
FUN	6
HULENBACK	6
CREATURES*	5
FRIGHTMARES*	5
PROJECTS	5
SICK	5
SLOWLY	5
BIKE	5
EDGE	5
HAIR	5
HUGE	5
LEG	5
LICKED	5
PUSHED	5
SLIPPED	5
STUCK	5
WAGGED	5
BICYCLE	4
CHEST	4
GROOM	4
II	4
III	4
WELL-BEING	4
ATTENTION	4
CHEEK	4
GRABBED	4
LEGS	4
PICKED	4
SNIFF	4
VETERINARIAN	4
BACKPACK	3
BET	3
CAIRN	3
CLIMBED	3
DARLING	3
GRINNED	3
HA	3
NICE	3
OKAY	3
PEERED	3
PENCIL	3
PETS	3
PLOPPED	3
REPEATED	3
WORRIED	3

Control Exp MA
(28 specialized types)

FEET	32
AMERICA	16
AREAS	14
AREA	12
PACIFIC	12
HUGE	10
OCEAN	10
CALIFORNIA	8
ALASKA	8
ICE	7
JOHN	7
SWIM	6
ATLANTIC	5
FLORIDA	5
GRADUALLY	5
LOUISIANA	5
ORANGE	5
APART	4
DELAWARE	4
EDUCATION	4
HAWAII	4
LOT	4
OREGON	4
SPLIT	4
WEATHER	4
CREATED	3
FEATURES	3
PREFER	3

Control Nar MA
(58 specialized types)

FEET	21
CLIFF	20
HAIR	18
FOOT	15
KITCHEN	13
SUDDENLY	12
RUG	11
NOSE	10
SKIN	10
CHAIR	10
TIRED	10
DUMB	9
KIDS	9
AFRAID	7
AFTERNOON	7
BREATH	7
EDGE	7
SCREAMED	7
GRAY	6
SLOWLY	6
WHISPERED	6
FINGERS	5
LEGS	5
BOTTOM	5
NODDED	5
QUICKLY	5
STARED	5
BARE	4
CAP	4
FADED	4
GAZE	4
LUNCH	4
MANAGED	4
PUSHED	4
SLAMMED	4
SLID	4
STOMACH	4
TIGHT	4
TIPTOED	4
TONGUE	4
TWISTED	4
WARM	4
WORRY	4
AUTUMN	3
BOTHER	3
DAMP	3
DELIGHT	3
ESPECIALLY	3
HIPS	3
HURRIED	3
KNEES	3
LUCKY	3
PICKED	3
POP	3
TAIL	3
TEETH	3
UNLIKE	3
WARNED	3

About the Author

Dee Gardner is an associate professor of Applied Linguistics and TESOL at Brigham Young University in Provo, Utah, United States. His primary research and teaching interests are in the areas of vocabulary acquisition, literacy development, and applied corpus linguistics.
E-mail: dee_gardner@byu.edu