Cohesion features in ESL reading: Comparing beginning, intermediate and advanced textbooks

Lia Plakans
The University of Iowa
United States

Zeynep Bilki
Kirkwood Community College
United States

Abstract

This study of English as a second language (ESL) reading textbooks investigates cohesion in reading passages from 27 textbooks. The guiding research questions were whether and how cohesion differs across textbooks written for beginning, intermediate, and advanced second language readers. Using a computational tool called Coh-Metrix, textual features were compared across the three levels using Multivariate Analysis of Variance (MANOVA). The results indicated that some features of cohesion yielded significant variation, but with small effect sizes. The majority of cohesion features considered were not different across the textbook levels. Larger effect sizes were found with factors like length, readability and lexical or syntactic complexity.

Keywords: Cohesion, ESL Textbooks, Reading Passages

In language classrooms, teachers, students, and learning materials provide second language (L2) learners with input to scaffold language development. Research has focused on the first two contributors, teachers and students, by investigating effectiveness of teaching methods or teachers' decision-making and by exploring students' language proficiency and individual characteristics, such as age, cognitive style, affect, or first language (L1). The third element, learning materials, has received less empirical attention despite its pervasiveness in school-based language learning.

The writing and publishing of English as a second language (ESL) textbooks impacts teachers and students, which raises the question of what guides the development of textbook content such as reading passages to meet the needs of the classroom. Several likely sources are (a) material writers' "craft knowledge" (Dubin, 1995, p. 15) and intuition (Crossely, Allen, & McNamara, 2011, 2012), (b) readability indices, which measure text difficulty usually through sentence and word length, (c) structural approaches based on graded word lists and grammatical structures (Allen, 2009), and (d) theory and research on L2 learning. Having a deeper research base on the

linguistic characteristics of language learning materials may lead to textbooks that support learners and teachers more soundly. With this goal in mind, our study investigates the nature of reading passages in ESL textbooks.

Second language reading has solid footing in language learning, especially as it facilitates academic literacy acquisition, yielding vast numbers of published ESL reading textbooks. In these textbooks, reading passages are a staple for practice and input. Crandall (1995) states, "the single most important decision you [textbook writers] will make in developing a reading textbook is the choice of [a reading] text or topic" (p. 84). This choice is influenced by difficulty, particularly when distinguishing beginning, intermediate, and advanced reading level textbooks. Since textbooks are marketed as suitable for specific proficiency levels, differentiating them is critical. To modify difficulty, writers design passages with such factors in mind as grammatical structures, lexical sophistication, and length (Crossely, Allen, & McNamara, 2011, 2012; Crossley, Greenfield, & McNamara, 2008), as well as selecting topics to pique students' interest (Crandall, 1995). However, research has shown that other factors, such as cohesion, interact with L2 readers' comprehension (Grabe, 2009). This feature of discourse is critical for readers to make both local and global connections across ideas, clauses, and words in a text (McNamara, Louwerse, & Graesser, 2002). Our study focuses on the differentiation of cohesion in reading passages across textbook levels.

Literature Review

To frame this study, three areas of literature will be reviewed: (a) relevant research on L2 reading at different proficiency levels, (b) the textual feature at the center of our study, cohesion, and (c) research on cohesion in language learning materials.

L2 reading proficiency levels: What distinguishes beginning, intermediate, and advanced?

Reading researchers underscore the importance of L2 proficiency as a dominant source of variance in reading performance (Alderson, 1984, 2000; Alderson & Urquhart, 1984; Bernhardt, 1991, 2011; Carrell, 1991; Clarke, 1980; Grabe, 1991; Taillefer, 1996; Usó-Juan, 2006). Research has shown that low proficiency readers are heavily involved at the word level rather than with discourse level processing in comparison to higher proficiency readers, and, at the same time, they are less accurate in word recognition (Koda, 2004). Such issues contribute to a slower reading rate for low proficiency readers in comparison to higher proficiency readers. In addition to word and discourse level processing, L2 readers also use semantic processing to integrate lexical and contextual information. Previous L2 text processing research (Alptekin & Erçetin, 2010; Horiba, 1996, 2000; Nassaji, 2003; Taillefer, 1996), which examined the performance of L2 readers in different proficiency levels (Nassaji, 2003; Taillefer, 1996) or compared L1 and L2 reading performance (Alptekin & Erçetin, 2010; Horiba, 1996, 2000), have shown that L2 readers draw heavily on their linguistic ability to extract meaning from various L2 texts, initially parsing text into smaller units such as words, phrases, and clauses, based on lexical and syntactic information available and then incrementally integrating them into the larger discourse context. Even when the learners become more proficient, reliance on textual and linguistic processes (e.g., lexical decoding, syntactic parsing, co-referencing) does not decrease

(Taillefer, 1996), but more proficient readers have the ability to shift attention to more abstract, conceptual ideas and make better use of background knowledge, using only as much textual information as needed for confirming and predicting the information in the text (Nassaji, 2003; Taillefer, 1996).

Based on their research, Alptekin and Erçetin (2010) concluded that literal understanding of text was essentially dependent on the level of language proficiency and surface readability features (e.g., syntactic parsing). However, highly proficient L2 users made use of high level automatic processing in integrating propositional units through the use of co-references, logical implications and cause-and-effect relationships. The research on proficiency and reading level has presented a multifaceted picture of the relationship between the processes of reading and readers' L2 proficiency. However, less is known about specifics of processing connections in text and how this interacts with proficiency.

Cohesion in L2 reading

Texts provided to L2 readers are not just a sequential display of isolated words and sentences but are connected syntactically, lexically, and semantically. Therefore, L2 readers need the ability to understand relationships among text elements, which are signaled both explicitly and implicitly through two discourse features—coherence and cohesion. McNamara, Louwerse and Graesser (2002) distinguish coherence and cohesion, explaining that the latter is "grounded in explicit linguistic elements (i.e., words, features, cues, signals, and constituents) and their combinations" (p. 11) while coherence lies in the interplay between text cohesion and the reader, which builds the reader's "mental model" of the text. Coherence is challenging to study as it is greatly affected by reader interpretation, which cannot be captured by only studying the text. In contrast, cohesion is found in the use of devices in the text such as connecting words or repeated word stems.

Scholars have provided a number of taxonomies for cohesion, which share some common categories. For example, Halliday and Hasan (1976) described two major categories of cohesive devices: grammatical and lexical. Grammatical cohesion includes anaphor reference (e.g., pronouns used to refer back to earlier noun phrases), substitution, and conjunctions. Lexical cohesion is captured when the same or related items appear within or across sentences. Louwerse (2002) provides several planes on which to consider different types of cohesion. The first, similar to Halliday and Hasan (1976) and Kintsch (1995), delineates grammatically driven and lexically driven cohesion. Another approach is to view cohesion as made locally, between adjacent clauses as well as globally, between groups of clauses. Thirdly, he suggests distinguishing sources of cohesion, such as conjunctions that are additive, temporal, or causal.

A critical issue with cohesion is how readers process the different approaches to comprehend texts. Louwerse (2002) provides parameters to study this issue by synthesizing earlier taxonomies. He presents three parameters of cohesion processing research: (a) type (causal, temporal, and additive), (b) polarity (positive and negative), and (c) direction (forward, bi-directional, and backward). Type refers to the relationship being illuminated by cohesion, for example an additive cohesive marker would be "in addition" or "and." The polarity refers to the agreement or contrast between the ideas being connected, "however" could be considered

negative, while "moreover" would indicate positive cohesion. Directionality is whether the cohesion marker is connected to ideas to come or to ideas already introduced in a text. In a series of two studies using eye tracking and reading rate, Lowerse found that L1 readers' cognitive processing rates were not occurring quite as he predicted for type and polarity. For some texts, additive cohesion processing was faster than causal, but the reverse was true with other texts. In Lowerse's eye tracking study, polarity was not processed at different rates, but in his reading rate study, negative cohesion was processed faster.

A number of studies have investigated L1 and L2 readers' awareness of cohesion in relation to text comprehension. This research suggests that individual differences exist in the ability to use connective devices such as coreferentials (e.g., a noun in one sentence that refers to a noun in another sentence) (Degand, Lefvére, & Bestgen, 1999) and logical connectors (e.g., *and*, *but*, *then*) (Degand & Sanders, 2002; Geva, 1992; Ozono & Ito, 2003). In a study of cohesion and comprehension, Degand and Sanders (2002) focused on one group of cohesive markers–causal connectives (words such as *because, so, consequently*)–and their impact on L1 and L2 readers of Dutch and French. Their results indicated that the reading comprehension of both groups benefitted from these causal markers.

Horiba (2000) compared reading processes of native and non-native English speakers using think aloud and recall protocols. Related to Louwerse's (2002) parameter of directionality, native speakers were found to use backward inferencing more, while non-native readers read texts similarly whether reading freely or when asked to focus on cohesion. Jonz (1987) conducted a study with L1 and L2 readers using two cloze test instruments, one with fixed ratio deletion and the other with cohesion-based deletion. The results revealed little difference between the two groups on the more common fixed ratio deletion, which is often used in testing reading ability. However, the cohesion-based cloze test was harder for the nonnative readers. Jonz suggests that language proficiency affects how readers recognize and utilize cohesive devices. He reflected that nonnative readers were more text bound thus relied more on cohesion markers.

Bilki (2014) conducted a qualitative study examining how highly proficient L2 readers construct meaning representations in low-cohesive and high-cohesive texts. The results revealed differences between the readers' meaning representation processes at the local and global levels of processing of the high- and low-cohesive text. These differences were most apparent in texts with low text cohesion. The low cohesive text allowed the readers to conduct more elaborative processing compared to their performance with the high cohesive one. All readers in the study processed explicit logical relationships constructed within sentences, mostly contrastive and causal links, but according to the readers' perception, these relationships were not sufficient components for meaning construction over the whole text.

Research has shown that language proficiency greatly impacts L2 reading, which explains the focus on grammar and vocabulary in materials development. However, discourse and processing features, such as cohesion, are part of the current theory of communicative competence, namely discourse competence (Canale & Swain, 1980) and are recognized for their role in reading (Koda, 2004). Therefore, we argue that cohesion across textbook levels should be given consideration in materials development and textbook writing to enhance the discussions of cohesion in text simplification, adaption, and readability (Bilki, 2014; Crossley, Allen, & McNamara, 2011, 2012;

Crossley, Greenfield, & McNamara, 2008; Simensen, 1987; Young, 1999).

Cohesion in language textbooks

A series of studies have investigated discourse features related to cohesion in ESL textbooks. Researchers have compared authentic and simplified ESL texts in beginning (Crossley, Louwerse, McCarthy, & McNamara, 2007) and intermediate levels (Crossley & McNamara, 2008) to see how their linguistic structures differed. In the first study, 105 beginning level texts were analyzed for linguistic structures using Coh-Metrix, a computational tool that measures text features such as cohesive relations, lexical familiarity, length, readability, etc. The second study, a replication and extension of the first, used 224 intermediate level texts, and analyzed the same features, comparing simplified and authentic texts to see if common assumptions about their differences could be substantiated. Although they characterized the linguistic features of these two kinds of texts, they did not compare the two textbook levels. The authors noticed some variation in the linguistic features that suggested level played a role, but it was not the focus of their studies.

In our study, we follow a process somewhat similar to Crossely et al. (2007) and Crossley and McNamara (2008), using Coh-Metrix to measure linguistic features of textbooks. We draw on reading passages from textbooks to target the actual source of ESL materials used in classes. When developing textbooks, materials writers have a sense of what is appropriate at beginning, intermediate, and advanced levels—an intuitive text level schema (Crossely, Allen, & McNamara, 2012). To uncover whether cohesion is part of this inferred formula, we studied the content of published in-use ESL reading textbooks for variation in cohesion features. Our main research interest was how cohesion varies across the three levels of textbooks: beginning, intermediate, and advanced.

Methods

To investigate this topic, a textual analysis was conducted by sampling reading passages from beginning, intermediate, and advanced ESL reading textbooks. Since Halliday and Hasan's landmark book (1976), researchers have detailed features of cohesion. However, a computational tool used to study text, Coh-Metrix, has provided researchers with the means to comprehensively answer many questions in the area of linguistic features and L2 literacy. In our study, passages were run in Coh-Metrix to produce measures for a range of discourse features related to cohesion, which were then compared across levels to reveal significant differences and effect sizes.

Text selection

A total of 162 ESL reading passages were selected from 27 college level ESL textbooks. These passages were categorized into three proficiency levels—beginning, intermediate and advanced—based on the designation by the textbook publishers (see Table 1 for textbook titles and levels and Appendix A for full bibliographic information). The textbooks were selected from a university ESL program's library and frequently used in the program's reading courses.

Table 1. *Textbook titles and levels*

Beginning	Intermediate	Advanced
Facts and Figures	Cause and Effect	Global Outlook 2
Quest 1	Quest 2	Quest 3
For Your Information	Tapestry 2	Tapestry 3
Theme for Today	Insights for Today	Issues for Today
Amazing Stories 1	Amazing Stories 2	Amazing Stories 3
Reading Advantage 1	Reading Advantage 2	Reading Advantage 3
Weaving it Together 1	Weaving it Together 2	Weaving it Together 3
Reading Explorer 1	Reading for a Reason	Reader's Choice
Password 1	Reading Matters	Password 3

Six reading passages were scanned from each textbook, providing 54 samples at each level. Information regarding these passages was recorded in a database. First of all, each passage location was tracked based on where it appeared in the text: the first, second, and final third of each textbook. Because of the exploratory nature of this study, we believed it was necessary to sample evenly within texts for breadth, and choose passages equally across the textbooks. However, reading passages within a textbook may become progressively more difficult linguistically, which may impact the differences being sought between levels.

Secondly, we recorded whether the passages were from authentic sources, adapted from authentic sources or written by the textbooks' authors. Table 2 details the number of readings in these three categories across the three levels, which shows a much higher number of authentic readings in the advanced level sample than the other levels. Authenticity was determined by checking source citations in textbooks for each reading passage selected. The passages that included a citation, footnote, or acknowledgment regarding an outside source for the passage were considered authentic or adapted. Passages with no indication of an original source outside of the textbook were considered non-authentic.

Table 2. *Origin of readings*

	Beginning	Intermediate	Advanced	Total
Authentic	6	5	22	33
Adapted	10	5	2	17
Written for textbook	38	44	39	112

While location and origin varied, we tried to minimize differences in topic and genre. The topics for the reading passages were typical for adult ESL learners, such as general readings in social science, science, and history as well as topics about daily life. In terms of genre, the selected passages were expository and did not include biographies, stories, letters, or poems. Using expository passages may have resulted in the high number of texts written by textbook authors in contrast to authentic texts. Table 3 describes the details of the reading passage sample.

Table 3. *Corpus information*

	Beginning	Intermediate	Advanced
Number of textbooks	9	9	9
Number of passages	54	54	54
Mean number of words	306.24	559.8	824.48
SD	94.33	246.18	376.25
Total words in corpus	16537	30229	44522

Variable selection and analysis

The main purpose of our study was to explore whether important elements of cohesion differ in ESL reading passages across the three textbook levels. Textual features were measured using the Coh-Metrix program. The 162 passages were run in Coh-Metrix 2.0 (www.cohmetrix.com), then 56 variables were selected, which were determined to be salient discourse features related to cohesion or were features that measured descriptive qualities of the texts, which contribute to textual differences across the levels. We used an approach that "cast a wide net" because the study was exploratory, seeking to find what features of the reading passages were significant. Although our focus was on cohesion, if the cohesion measures did not reveal differences across levels, it was important to determine what features were. We ran Multivariate Analysis of Variance (MANOVA) for all 56 variables from Coh-Metrix (see Appendix B for the full list of cohesion-related variables) and the three textbook levels, and then inspected the resulting *F*-values, significance, and effect sizes.

For full descriptions of these measures see the weblink listed above for Coh-Metrix. For a number of variables in Coh-Metrix, the output reports the density of textual features using incidence, ratio or proportion (Graesser et al., 2004). Incidence scores indicate the number of occurrences per 1,000 words. Ratios or proportions are used when one text feature is compared with another, for example "causal cohesion" is the ratio of causal verbs to causal particles. Averages are another measure used in Coh-Metrix, such as with the Latent Semantic Analysis (LSA) measure.

From this screening, we found 24 significant variables, 19 of which had a large effect size[1]. We ran correlations of the 24 to see if they were highly related, using a cut off of $r = .70$ to define a strong relationship between variables (Field, 2009). We found four variable groupings that held correlations that exceeded this. We selected the variable with the largest effect size from the Analysis of Variance (ANOVA) to represent each group.

After following this process, 18 focus variables were left, which are listed in Table 4. These remaining variables were categorized by type: referential cohesion, connectives, situation model, syntactic complexity, descriptive, word information, and readability. Several of these relate directly to cohesion (referential cohesion, connectives, situation model), while other relate somewhat (syntactic complexity & pattern density) or are simply features that could be described as measuring the length and difficulty of a text (descriptive, word information, readability). We

---

[1] Large effect sizes were those larger than 0.138 ($\eta^2 > .138$; Cohen, 1988). We used Cohen's classification of effect sizes (1988) to select variables with large effect sizes; since this study was exploratory, this less conservative cut off was deemed appropriate.

included these latter two groups with the possibility that they may likely distinguish these levels. Their inclusion allows our discussion of textbook differences to consider factors other than cohesion that are more likely to be used by textbook authors, in the case that cohesion is not a strong determinant of level. However, our main focus remained on cohesion while the other factors are supplemental.

Table 4. *Variables included in final study*

Coh-Metrix Variable	Description
Referential cohesion: portion of content words that overlap between adjacent sentences	Referential cohesion refers to the overlap in explicit content words between adjacent sentences, or between all of the sentences in a text
Connectives: Incidence of negative causal connectives	Connectives are cohesive links between ideas and clauses in a text such as negative and positive causal connectives (because, so, although), logical (and, or), and contrastive connectives (although, whereas).
Situation Model: Causal cohesion -- Ratio of causal particles to causal verbs; Incidence of causal verbs and particles; Incidence of intentional action, events, and particles; Temporal cohesion --Tense and aspect repetition	Situation model is described as the features that are present in the reader's mental representation of a text (Kintsch, 1998; Graesser & McNamara, 2011). Causation, intentionality, and temporality are three important dimensions of the situation model; and their contents including intentional cohesion particles (e.g., in order to, so that) and causal particles (e.g., because, so) are used to measure causal and intentional cohesion level of a text.
Syntactic similarity: Sentence syntax similarity all across paragraphs	Refers to the type of syntactic structures used and the repetition of similar patterns. For example, some lower level texts only use simple sentences that follow a simple syntactic pattern (actor-action-object)
Syntactic Complexity: Mean number of modifiers per noun phrase; Noun phrase incidence score	Syntactic complexity refers to syntactic composition of sentences or paragraphs in a text, for example, some sentences in a text are short and have few if any embedded clauses. It tends to be easier to process a text when there are shorter sentences, few words before the main verb of the main clause, and few words per noun-phrase.
Descriptive indices: Average words per sentence; Number of words	Descriptive indices are main descriptive features of a text used to interpret patterns of textual data such as number of words in a text and average words per sentences.
Word information: Average word ferquency for all words; Average minimum word frequency in sentences; Personal proun incidence score; Concreteness in sentences for content words; Concreteness in the text for	Word information refers to the idea that each word in a text is assigned to a syntactic part-of-speech category including content words (e.g., nouns, verbs, adjectives, adverbs) and function words (e.g., prepositions, pronouns). Coh-Metrix assigns only one part-of-speech category to each word on the basis of its syntactic context, computes word frequency scores and also provides an index of how concrete a word is in a text.

content words

Readability: Flesch Kincaid Grade Level; Flesch Reading Ease Score	Readability is a method of assessing texts on difficulty consisting of various readability formulas. Reading ease score is a number from 0 to 100, with a higher score indicating easier reading. Reading grade levels range from 0 to 12. The higher the number, the harder it is to read the text.

To conduct the main comparison analysis, descriptive statistics were completed, and then MANOVA was run to see how the 18 features collectively related to textbook level. ANOVA tests provided further information about the 18 individual variables across levels. Since this number of ANOVA tests is fairly high, we used a Bonferroni adjustment (.05/10) to designate significance at $p < .005$ in order to protect from Type I errors. Lastly, pair-wise comparisons using Tukey HSD (honest significant difference) and Tamhane's T2 allowed us to delve into the differences between each level. Tamhane test was used for post hoc testing only with variables that did not meet the variance assumption.

Results

Sample means and standard deviations (descriptive statistics) for all textual variables as measured for each textbook level are listed in Appendix C.

MANOVA was used to examine the difference in the large effect size variables across reading passages from three textbook levels: beginning, intermediate and advanced. Results indicated that the combined variables resulted in a significant main effect for text level ($F(36, 284) = 6.59$, Wilks's lambda 1 ($\wedge$) = .310, partial η^2 = .455, $p < .05$) with a large effect size. In this section, we will first describe the ANOVA results for features most directly related to cohesion followed by other textual features likely to distinguish text levels. Lastly, we will summarize which features increased or decreased across textbook level, and, the comparative order of the variables related to effect sizes.

Cohesion features and text level

A significant difference was observed between textbook level with regard to proportions of content words that overlap between adjacent sentences, ($F(2, 159) = 8.18$; $p < .05$; partial η^2 = .09), which we categorized as referential cohesion. The follow-up Tamhane test indicated that this feature increased from the beginning level to the intermediate and advanced level but was not significantly different between intermediate and advanced level texts.

For the second kind of cohesion, connectives, a statistically significant effect was found with incidence of negative casual connectives ($F(2, 159) = 11.73$; $p < .05$; partial η^2 = .13). The follow-up Tamhane test indicated that negative causal connectives were statistically different between beginning and intermediate level, and beginning and advanced level texts, but not between the intermediate and advanced level texts. The intermediate and advanced level textbooks had significantly higher number of negative causal connectives than the beginning level textbooks.

Textbook level was also statistically significant for all four situation-model features, ratio of causal particles to causal verbs ($F(2, 159) = 13.37$; $p < .05$; partial $\eta^2 = .14$), incidence of causal verbs and particles ($F(2, 159) = 12.49$; $p < .05$; partial $\eta^2 = .14$), incidence of intentional action events and particles ($F(2, 159) = 25.59$; $p < .05$; partial $\eta^2 = .24$), and mean of tense and aspect repetition scores ($F(2, 159) = 5.61$; $p < .05$; partial $\eta^2 = .07$). The follow-up Tukey HSD revealed that the ratio of causal particles to causal verbs and the incidence of causal verbs, links and particles were statistically different between beginning level and the other two higher level texts. The ratio of causal particles to causal verbs increased from the beginning level to intermediate and advanced level, but incidence score for causal verbs, links and particles decreased from beginning to intermediate and advanced, indicating that the beginning level texts have higher number of causal features. The follow-up test, Tukey HSD, indicated that the variable "mean of tense and aspect repetition" was only significant between the beginning level texts and advanced level texts and decreased from the lower level to the advanced level. Despite the lack of a significant difference between intermediate and advanced level texts, both beginning and intermediate texts' mean scores were lower than the advanced texts. The follow-up Tamhane test indicated that the incidence of intentional actions, events, and particles decreased significantly from the beginning level to the intermediate level and to the advanced level. In sum, based on the causal cohesion analysis in our study, advanced level texts had a higher ratio of causal particles to causal verbs, which indicated that these texts showed less causal cohesion than beginning level texts. Higher ratio results from the texts having many causal verbs, but few causal particles.

As mentioned in the methods section, a large set of Coh-Metrix variables were run in the initial stage of the study, 24 of which can be considered directly related to cohesion (see Appendix B for full list). Only six held statistically significant differences across textbook levels. Four of these were cohesion features related to situation model construction. For two other categories of cohesion, referential cohesion and connectives, only one of the features was significant. None of the measures of LSA yielded differences across textbook level. Given that cohesion only differentiated textbook levels with a few significant features, our results include other textual features that were found to have significant differences with meaningful effects.

Other features and text level

A significant difference was found for syntactic complexity features as well as sentence syntax similarity all across paragraphs ($F(2, 159) = 57.79$; $p < .05$; partial $\eta^2 = .42$), mean number of modifiers per noun phrase ($F(2, 159) = 8.90$; $p < .05$; partial $\eta^2 = .10$) and noun phrase incidence score ($F(2, 159) = 44.15$; $p < .05$; partial $\eta^2 = .36$). Beginning and intermediate texts were not significantly different from each other in terms of the mean number of modifiers before noun phrase*,* but both were significantly lower than the advanced texts. Noun phrase incidence score was significantly different across all three levels, decreasing from the lower level texts to higher level ones. The follow-up Tamhane test indicated that sentence syntax similarity across all paragraphs decreased significantly across the three levels.

Statistical significance was also found with descriptive features: average words per sentences ($F(2, 159) = 80.47$; $p < .05$; partial $\eta^2 = .50$) and number of words in the text ($F(2, 159) = 50.59$; $p < .05$; partial $\eta^2 = .39$). The follow-up test, Tukey HSD, indicated that the descriptive feature,

average number of words per sentences, was statistically different across all three levels, revealing an increase from the beginning to the advanced level texts. The follow-up Tamhane test indicated that there was also a significant increase in number of words from the beginning to the advanced level texts.

Four of five word information features revealed significant differences: concreteness in sentences for content words ($F(2, 159) = 14.95$; $p < .05$; partial $\eta^2 = .16$), concreteness in the text for content words ($F(2, 159) = 12.66$; $p < .05$; partial $\eta^2 = .14$), average word frequency for all words ($F(2, 159) = 14.76$; $p < .05$; partial $\eta^2 = .16$), and personal pronoun incidence score ($F(2, 159) = 7.26$; $p < .05$; partial $\eta^2 = .08$). Only the average minimum word frequency-in-sentence was non-significant ($F(2, 159) = .622$; $p < .05$; partial $\eta^2 = .01$). The follow-up Tukey HSD indicated that the concreteness of content words and average word frequency for all words, decreased significantly from the beginning level to the intermediate level and to the advanced level. The follow-up Tamhane indicated that concreteness in sentences in beginning level texts was significantly higher than both intermediate and advanced texts. No significance difference was found between intermediate and advanced level texts. The follow-up Tukey HSD indicated that personal pronoun incidence score decreased significantly from the beginning level texts to advanced level texts, but no significant difference was found between beginning and intermediate levels and intermediate and advanced levels.

A statistically significant effect was observed in textual readability features, including reading level (Flesch Kincaid Grade Level) ($F(2, 159) = 78.53$; $p < .05$; partial $\eta^2 = .50$) and reading ease (Flesch Reading Ease Score) ($F(2, 159) = 50.38$; $p < .05$; partial $\eta^2 = .39$). The follow-up test, Tukey HSD, indicated that both readability measures were statistically significantly different across all three textbook levels. While reading level increased with textbook level from beginning to advanced, reading ease decreased.

Accuracy of the model (Discriminant Function Analysis)

ANOVA results summarized which textual features increased or decreased across textbook level, and the comparative order of the variables related to effect sizes. To demonstrate how predictive our analysis is, we conducted a discriminant function analysis (DFA). DFA is a statistical procedure that is able to predict how many dimensions we would need to express the relationship between a group of independent variables (the significant Coh-Metrix variables) and the one categorical variable (the level of the reading texts). Using this relationship, we aimed to predict a classification based on the Coh-Metrix variables and assess how well these variables separate the text levels in the classification. First, we generated a discriminant function using the entire original set to predict group membership. Then, we used this discriminant function analysis model to predict group membership of the reading texts using repeated cross-validation. We conducted "Leave-one-out" classification option, which provides a cross-validated component of the classification results. Then, we compared the results of the discriminant analysis in both the original set (original texts) and the cross-validation set (the predicted texts) to see if these results were all statistically significant, which supports the predictions of the analysis.

Table 5 shows the correspondence between the original texts and the predictions (cross-validated data) made by the discriminant function analysis. The results demonstrate that 72.2% of reading

texts were correctly classified into "beginning" "intermediate" and "advanced" levels in the analysis sample ($df = 10$, $n = 162$) $\mu 2 = 168.129$, $p < .001$). For the cross-validated set, 71.0% of reading texts were correctly classified. Maximum chance for these analyses is 33.3%; therefore, our model accuracy rate of 71.0% exceeds this standard. According to our structure matrix results in DFA, all 16 significant Coh-Metrix variables except proportion of content words that overlap between adjacent sentences[2], are important variables that discriminate between three text levels.

Table 5. *Predicted text level vs. original text level results from original and cross-validated set*

Original text level	Predicted text level		
	Beginning	Intermediate	Advanced
Original count			
Beginning	43	11	0
Intermediate	9	36	9
Advanced	0	16	38
%			
Beginning	79.6	20.4	0
Intermediate	16.7	66.7	16.7
Advanced	0	29.6	70.4
Cross-validated count			
Beginning	43	11	0
Intermediate	9	35	10
Advanced	0	17	37
%			
Beginning	79.6	20.4	0
Intermediate	16.7	64.8	18.5
Advanced	0	31.5	68.5

Summary

Results of the study have demonstrated significant differences in textual features across reading passages in beginning, intermediate, and advanced textbook levels including six features directly related to cohesion. Ten features decreased from lower level texts to higher level texts, while eight features increased across levels (see Table 6). The close examination of each textual feature's effect across texts revealed significant differences between levels, particularly between beginning and two higher level texts, intermediate and advanced.

---

[2] We ran both normal and stepwise DFA for our data. According to the structure matrix results coming from the stepwise run, proportion of content words that overlap between adjacent sentences, is less important across levels, but the normal run reveals that the concreteness minimum in sentences for content words is less important (which means that it stays under the cut of point 0.30). Stepwise confirms our results that this feature increased from the beginning level to the intermediate and advanced level but was not significantly different between intermediate and advanced level texts.

Table 6. *Significant differences across text levels (I: Increased, D: Decreased, NSD: No significant difference)*

		Beginning-Intermediate	Beginning-Advanced	Intermediate-Advanced
Readability features	Reading level	I	I	I
	Reading ease	D	D	D
Descriptive features	Average words per sentences	I	I	I
	Number of words	I	I	I
Word information	Concreteness of content words	D	D	D
	Average word frequency for all words	D	D	D
	Concreteness in sentences	D	D	NSD
	Personal pronoun incidence	NSD	D	NSD
Syntactic complexity	Mean number of modifiers before noun phrase	NSD	I	I
	Noun phrase incidence score	D	D	D
	Sentence syntax similarity across all paragraphs	D	D	D
Situation model	Incidence score for causal verbs, links, and particles	D	D	D
	Incidence of intentional actions, events, and particles	D	D	D
	Ration of causal particles to causal verbs	I	I	NSD
	Mean of tense and apect repetition scores	NSD	I	NSD
Referential cohesion	Proportion of content words that overlap between adjacent sentences	I	I	NSD
Connectives	Incidence of negative causal connectives	I	I	NSD

The greatest effect sizes ($\eta^2 > .35$; Cohen, 1988) for differences between text levels were found, not with the cohesion features, but in the other textual features, in particular average words per sentences and total number of words in the text; syntactic complexity; syntactic syntax similarity;

and readability features. One causal cohesion index showed a medium effect size ($\eta^2 > .15 < .35$). The smallest effect sizes were found with referential cohesion, connectives, and three situation model features.

Discussion

This study set out to explore the differentiation of cohesion features in ESL textbook reading passages across beginning, intermediate, and advanced levels. The results indicate that some features differ across textbooks, based on analysis of 162 reading passages. However, of the 24 cohesion features included in the initial MANOVA only six were significant. The categories applied in the first MANOVA came from Coh-Metrix (see Appendix B), which included measures of referential cohesion, LSA (meaning-related connection across a text), connectives, and situation model. While our list of cohesion variables is not exhaustive, 18 of them did not differ significantly across the textbook levels. None of the LSA features varied across levels. This finding is in contrast to Crossley et al. (2012) who found that beginning level texts had a higher degree of global semantic similarity as they contained more repeated or known information compared to higher level texts which included more new information.

Of the nine connectives features in our study only one type differed across levels, which was negative connectives; no significant difference was found with other very common types of connectors such as additive, temporal or logical connectives. It is possible that the quantity of connectives did not differ across levels while the sophistication or difficulty level did; for example, the positive additive connective "in addition" could appear in lower level texts while "moreover" might appear in higher level texts. Another cohesion feature with only one significant feature was referential cohesion, with the overlap of words across a sentences being different at the beginning level from intermediate or advanced. Commonly taught referential markers, such as anaphor reference, which includes pronouns, were not significantly different in reading passages across text levels. Lastly, in contrast to other cohesion categories, the situation model cohesion features included in our analysis were all significantly different (four of four) across textbook levels. In sum, our results found a handful of cohesion features were distinct across levels, but many were not. A general conclusion could be that this is not a feature that materials writers consider when writing reading passages for certain levels of students.

Our results can be interpreted in light of prior characterizations of cohesion. Louwerse (2002) has suggested considering cohesion in terms of three parameters: type, polarity, and direction. Halliday and Hasan (1976) grouped cohesion by grammatical or lexical types. Of the six cohesion features in our study that were significantly different across textbook levels, five appear to fall in the grammatical category: negative causal connectives, causal cohesion, incidence of causal verbs and particles, and incidence of action events and particles. The remaining significant cohesion feature, content word overlap across adjacent sentences, falls into the lexical category. Thus, our study shows a grammatical trend in how ESL textbooks vary cohesion across levels. Another aspect of the type of cohesion features worth noting is the predominance of significant causal connection variables (four of the six), which suggests that material writers differentiate their use of causal connections across textbook level. This finding may connect to the research of Degand and Sanders (2002) who found causal connections were important for L2

readers' comprehension. In terms of polarity (positive or negative), one significant cohesion variable in the study indicated a positive or negative relation–the incidence of negative causal connectives. The other polarity connectives were not significantly different across levels, indicating that this aspect of cohesion may not be considered in differentiating high and low level reading texts. The third parameter, directionality, was not captured by the measures used in our study. Using these categorizations reinforces that material writers are focusing on grammatical differences across text levels, at least in terms of using causal forms; this focus may include grammatical cohesion either intentionally or by default. However, lexical cohesion features are not being used in the same way, suggesting potential for more attention to this type of cohesion markers to distinguish reading levels.

Our results indicated some pattern for the significant cohesion features between textbook levels, particularly in level-to-level comparisons and in terms of increasing or decreasing. For most of cohesion features in the study, the greatest difference occurred between the beginning level texts and the upper two levels. In other words, the reading passages classified as intermediate and advanced were more similar in terms of the cohesion features than either was with beginning reading passages. Three of the cohesion variables increased from beginning to intermediate and advanced texts. However, in other three cases the cohesion features decreased. The decreases were situation model features: incidence of causal verbs, links and particles as well as incidence of intentional actions, events, and particles. Beginning level texts had a higher number of causal lexical features including casual verbs and particles, and both beginning and intermediate level texts had a higher number of intentional actions, events and particles. According to the Coh-Metrix online document (2013), cohesion suffers when the text has many causal verbs but few causal particles that signal how the events and actions are connected. Ratios in our study increased across levels, indicating that higher level texts had many causal verbs, but few causal particles compared to lower level texts. This demonstrates that beginning level texts had more causal particles, which could make them more cohesive.

As cohesion was only minimally distinct across textbook levels in the 162 reading passages, we considered the impact of other textual features that might distinguish reading passages for different levels of readers. These features, sentence and text length differences and readability features, demonstrated significant difference across levels. These findings suggest that such features are primary in distinguishing reading passages across the three levels; material developers may use them to write level-appropriate materials. However, these features may have an indirect connection to cohesion. For example, previous research (Crossley, Greenfield, & McNamara, 2008; Graesser et al., 2004; O'Reilly & McNamara, 2007) states the importance of readability features in understanding cohesion level of the texts. In fact, there is often a reverse relation between cohesion and traditional measures of readability such as grade level and reading ease (Graesser et al., 2004; O'Reilly & McNamara, 2007). Traditional measures of text difficulty rely on sentence length—the shorter the sentence length, the easier the text. Thus, increasing cohesion typically results in an increase in sentence length and therefore increases text difficulty. In our study, while reading level increased with textbook level from beginning to advanced, reading ease decreased across the texts suggesting that advanced level texts should be more cohesive. However, text length alone is not a valid indicator to understand the cohesion level of the text.

From the syntactic perspective, beginning level texts contained a higher level of structural similarity across paragraphs and an increase in the number of noun phrases but a decreased number of modifiers before noun phrases. Our results showed that the syntax in the lower level texts tended toward shorter sentences with few modifiers before noun phrases and duplicated noun phrases. Previous research (Siddharthan, 2006) shows that simplifying a text syntactically, especially for relative clauses and appositives, results in the duplication of noun phrases thus may increase the number of noun phrases in lower level texts. "Syntactic transformations can also change the grammatical function of noun phrases and alter the order in which they are introduced into the discourse. This can result in an altered attentional state at various points in the discourse" (Siddharthan, 2006, p. 101). Thus, the syntax transformation at lower levels could disrupt the flow and cohesiveness of a reading.

The results of our study can be considered in light of implications for textbook development, and further research. In general, our study suggests that cohesion needs more attention in textbook and materials development. The largest effects in terms of differences across textbook levels were found with text features that only indirectly contribute to cohesion such as word count, words per sentence, syntactic similarity, and readability. These features are related to text difficulty, which does impact comprehension. Difficulty has been identified in other studies as differentiating authentic and inauthentic reading passages (Crossley et al., 2007; Crossley & McNamara, 2008). However, research has also pointed to the impact of cohesion on L2 reading (Bilki, 2014; Degand & Sanders, 2002; Horiba, 2000; Jonz, 1987). Therefore, cohesion could be employed by material writers to distinguish features between levels. The finding that cohesion was more detectable at lower levels than between intermediate and advanced suggests that textbook authors are possibly attending to cohesion as a way to distinguish levels in reading passages for lower proficiency readers. However, more differentiation of cohesion could be done at intermediate and advanced levels. In addition, many of the cohesion devices included in the initial screening of the study were not significant; many of these could be considered when writing texts at different levels.

Considering cohesion within the difficulty or level distinction formula has potential; in other words, how does cohesion increase or decrease the ease of reading in a L2? Clearly, more research is needed to answer this question and to make critical decisions in materials development. Research can also delve further into this topic by reporting on how teachers use reading passages and cohesion in teaching L2 readers. Our study serves as a baseline on what appears in currently published and commonly used textbooks, which is simply a starting point for discussion and innovation.

Conclusion

The findings from this study are only a piece of the larger puzzle about L2 reading and cohesion. Thus far, the focus in L2 learners and cohesion has been mostly on writing, not reading. However, based on the results of our study, it seems that through ESL reading textbooks, students are exposed early to textual features that contribute to cohesion in reading passages. In some cases, this exposure continues through the advanced levels, in others it diminishes. Cohesion differences across levels are more grammatical than lexical and seem to be related

causal connectedness. In many cases, the amount of cohesion changes little from beginning levels to advanced levels, suggesting that this feature may be under utilized in textbook writing as a feature to differentiate levels of reading. Studies of readers' processes will help illuminate how students use cohesive features at the different levels, which might further inform materials developers and teachers whether the patterns we found in our study are the best way to introduce students to cohesion or if different approaches would be more conducive to learning.

Several limitations of our study could be incorporated into future research to improve the understanding of cohesion in reading and reading textbooks. First of all, the level designation adopted in our study originated from the textbook publishers. The divisions of beginning, intermediate, and advanced are not uniform in the field; therefore, these levels might not be consistent across textbooks. How publishers designate levels is an important and practical question to attend to. Secondly, our study was exploratory, and therefore, somewhat modest in scope. However, further investigation would benefit from a larger data set that could include different genres of texts and more passages at each level. Lastly, as was alluded to earlier in the discussion, the analysis in our study centered on quantity of cohesion features. More work should be carried out to explore qualitative shifts across levels in terms of cohesion sophistication or variety.

References

Allen, D. (2009). A study of the role of relative clauses in the simplification of new texts for learners of English. *System*, *37*, 585–599. DOI: 10.1016/j.system.2009.09.004

Alderson, J. C. (1984). Reading in a foreign language: A reading problem or a language problem? In J. C. Alderson & A. H. Urquhart (Eds.), *Reading in a foreign language* (pp. 122–135). New York, NY: Longman.

Alderson, J. C. (2000). *Assessing Reading.* Cambridge, UK: Cambridge University Press.

Alderson, J. C., & Urquhart, A. H. (Eds.) (1984). *Reading in a Foreign Language*. London, UK: Longman.

Alptekin, C., & Erçetin, G. (2010). The role of L1 and L2 working memory in literal and inferential comprehension in L2 reading. *Journal of Research in Reading*, *33*, 206–219. DOI: 10.1111/j.1467-9817.2009.01412.x.

Bilki, Z. (2014). *A close observation of second language (L2) readers and texts: meaning representation and construction through cohesion* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses. (Accession Order No. 3638349).

Bernhardt, E. B. (1991). *Reading development in a second language: Theoretical, empirical, and classroom perspectives.* Norwood, NJ: Ablex Publishing.

Bernhardt, E. B. (2011). *Understanding Advanced Second-Language Reading*. New York: Routledge.

Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, *1*, 1–47. DOI*:* 10.1093/applin/I.1.1

Carrell, P. L. (1991). Second language reading: Reading ability or language proficiency? *Applied Linguistics*, *12*, 159–179. DOI: 10.1093/applin/12.2.159

Clarke, M. A. (1980). The short circuit hypothesis of ESL reading–or when language competence interferes with reading performance. In P. L. Carrell, J. Devine & D. Eskey,

(Eds.), *Interactive approaches to second language reading* (pp. 114–125). New York, NY: Cambridge University Press.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Crandall, J. (1995). The why, what, and how of ESL reading instruction: Guidelines for writers of ESL reading textbooks. In P. Byrd (Ed.), *Material writer's guide* (pp. 79–94). New York, NY: Heinle & Heinle.

Crossley, S. A., Allen, D., & McNamara, D.S. (2011). Text readability and intuitive simplification: A comparison of readability formulas. *Reading in a Foreign Language*, *23*, 84–101. DOI: 10.1177/1362168811423456

Crossley, S. A., Allen, D., & McNamara, D. S. (2012). Text simplification and comprehensible input: A case for an intuitive approach. *Language Teaching Research*, *15*, 89–108. DOI: 10.1177/1362168811423456

Crossley, S. A., Louwerse, M. L., McCarthy, P. M., & McNamara, D. S. (2007). A linguistic analysis of simplified and authentic texts. *The Modern Language Journal*, *91*, 15–30. DOI: 10.1111/j.1540-4781.2007.00507.x

Crossley, S. A., Greenfield, J., & McNamara, D. S. (2008). Assessing text readability using cognitively based indices. *TESOL Quarterly*, *42*, 475–493. DOI: 10.1002/j.1545-7249.2008.tb00142.x

Crossley, S. A., & McNamara, D. S. (2008). Assessing second language reading texts at the intermediate level: An approximate replication of Crossley, Louwerse, McCarthy and McNamara (2007). *Language Teaching*, *41*, 409–429. DOI: 10.1017/S0261444808005077

Degand, L., Lefèvre, N., & Bestgen, Y. (1999). The impact of connectives and anaphoric expressions on expository discourse comprehension. *Document Design*, *1*, 39–51. DOI: 10.1075/dd.1.1.06deg

Degand, L., & Sanders, T. (2002). The impact of relational markers on expository text comprehension both in L1 and L2. *Reading and Writing*, *15*, 739–758.

Dubin, F. (1995). The craft of materials writing. In P. Byrd (Ed.), *Material writer's guide* (pp. 13–22). New York, NY: Heinle & Heinle Publishers.

Field, A. (2009). *Discovering statistics using SPSS (3rd ed.).* Thousand Oaks, CA: Sage Publication.

Geva, E., (1992). The role of conjunctions in L2 text comprehension. *TESOL Quarterly*, *26*, 731–747. DOI: 10.2307/3586871

Grabe, W. (1991). Current developments in second language reading research. *TESOL Quarterly*, *25*, 375–397.

Grabe, W. (2009). *Reading in a second language: Moving theory to practice.* Cambridge, UK: Cambridge University Press. DOI: 10.2307/3586977

Graesser, A. C., McNamara, D. S., Louwerse, M. M., & Cai, Z. (2004). Coh-Metrix: Analysis of Text on Cohesion and Language. *Behavioural Research Methods, Instruments and Computers*, *36*, 193–202.

Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English.* London, UK: Longman.

Horiba, Y. (1996). Comprehension processes in L2 reading: Language competence, textual coherence, and inferences. *Studies in Second Language Acquisition*, *18*, 433–473. DOI: 10.1017/S0272263100015370

Horiba, Y. (2000). Reader control in reading: Effects of language competence, text type, and task.

Discourse processes, *29*, 223–367. DOI: 10.1207/S15326950dp2903_3

Jonz, J. (1987). Textual cohesion and second language comprehension, *Language Learning*, *37*, 409–438. DOI: 10.1111/j.1467-1770.1987.tb00578.x

Kintsch, W. (1995). How readers construct situation models for stories: the role of syntactic cues and causal inferences. In Gernsbacher, M.A., & Givón, T. (Eds.), *Coherence in spontaneous text* (pp. 139–160). Amsterdam and Philadelphia: John Benjamins.

Koda, K. (2004). *Insights into second language reading: A cross-linguistic approach.* Cambridge, UK: Cambridge University Press.

Louwerse, M. (2002). An analytic and cognitive parameterization of coherence relations. *Cognitive Linguistics*, *12*, 291–315. DOI: 10.1515/cogl.2002.005

McNamara, D. S., Louwerse, M. M., & Graesser, A. C. (2002). Coh-Metrix (Version 2.0) [Software]. Memphis, TN: University of Memphis, Institute for Intelligent Systems. Available from http://cohmetrix.memphis.edu/cohmetrixpr/index.html.

Nassaji, H. (2003). Higher–Level and Lower–Level Text Processing Skills in Advanced ESL Reading Comprehension. *The Modern Language Journal*, *87*, 261–276. DOI: 0.1111/1540-4781.00189

O'Reilly, T., & McNamara, D. S. (2007). Reversing the reverse cohesion effect: good texts can be better for strategic, high-knowledge readers. *Discourse Processes*, *43*, 121–152. DOI: 10.1080/01638530709336895.

Ozono, S., & Ito, H. (2003). Logical connectives as catalysts for interactive L2 reading. *System*, *31*, 283–297. DOI: 10.1016/S0346-251X(03)00025-3

Siddharthan, A. (2006). Syntactic simplification and text cohesion. *Research on Language & Computation*, *4*, 77–109.

Simensen, A. M. (1987). Adapted readers: How are they adapted? *Reading in a Foreign Language*, *4*, 41–57.

Taillefer, G. (1996). L2 reading ability: Further insights into the short-circuit hypothesis. *The Modern Language Journal*, *80*, 461–477. DOI: 10.1111/j.1540-4781.1996.tb05466.x

Usó-Juan, E. (2006). The compensatory nature of discipline-related knowledge and English language proficiency in reading English for academic purposes. *The Modern Language Journal*, *90*, 210–227. DOI: 10.1111/j.1540-4781.2006.00393.x

Young, D. J. (1999). Linguistic simplification of second language reading material: Effective instructional practice? *The Modern Language Journal*, *83*, 350–366. DOI: 10.1111/0026-7902.00027

Appendix A

Full bibliographic information for textbook titles and levels

Ackert, P., & Lee, L. (2004). *Facts and figures: Reading and vocabulary development* (4th ed.). Boston, MA: Heinle and Heinle.

Ackert, P., & Lee, L. (2004). *Cause and effect: Reading and vocabulary development* (4th ed.). Boston, MA: Heinle and Heinle.

Berish, L., & Thibaudeau, S. (1998). *Amazing stories 1: To tell and retell.* Boston, MA: Houghton Mifflin.

Berish, L., & Thibaudeau, S. (1998). *Amazing stories 2: To tell and retell.* Boston, MA: Houghton Mifflin.

Berish, L., & Thibaudeau, S. (1998). *Amazing stories 3: To tell and retell.* Boston, MA: Houghton Mifflin.

Blanchard, K., & Root, C. (1995). *For your information 1: Basic reading skills.* White Plains, NY:

Longman Publishing Group.
Blass, L., & Whalley, (2005) *Reading for a Reason 2*. New York, NY: McGraw Hill.
Bonesteel, L. (2004). *Password 3: A reading and vocabulary text.* Pearson.
Broukal, M. (1994). *Weaving it together 1* (2nd ed.). Boston, MA: Heinle ELT.
Broukal, M. (1994). *Weaving it together 2* (2nd ed.). Boston, MA: Heinle ELT.
Broukal, M. (1994). *Weaving it together 2* (2nd ed.). Boston, MA: Heinle ELT.
Butler, L. (2009). *Password 1: A reading and vocabulary text.* Upper Saddle, NJ: Prentice Hall.
Douglas, N. (2008). *Reading explorer 1*. Washington D.C.: National Geographic.
Dyer, B., & Bushell, B. (2002). *Global outlook advanced reading 2*. New York, NY: McGraw Hill.
Fellag, L. R. (2000). *Tapestry 3* (2nd ed.). Boston, MA: Heinle and Heinle.
Hartmann, P., & Blass, L. (2006). *Quest: Reading and writing 1* (2nd ed.). New York, NY: McGraw Hill.
Hartmann, P., & Blass, L. (2006). *Quest: Reading and writing 2* (2nd ed.). New York, NY: McGraw Hill.
Hartmann, P., & Blass, L. (2006). *Quest: Reading and writing 3* (2nd ed.). New York, NY: McGraw Hill.
Malarcher, C. (2003). *Reading advantage 1* (2nd ed.). Boston, MA: Heinle and Heinle.
Malarcher, C. (2003). *Reading advantage 2* (2nd ed.). Boston, MA: Heinle and Heinle.
Malarcher, C. (2003). *Reading advantage 3* (2nd ed.). Boston, MA: Heinle and Heinle.
Oxford, R., & Sokolik, M. E. (2000). *Tapestry reading 2*. Boston, MA: Heinle and Heinle.
Silberstein, S., Clarke, M., & Dobson, B. (2008). *Readers choice* (5th ed.). Ann Arbor: MI: University of
        Michigan Press.
Smith, L.C., & Mare, N.N. (2003). *Themes for today 1* (3rd ed.). Boston, MA: Heinle and Heinle.
Smith, L.C., & Mare, N.N. (2003). *Insights for today 2* (3rd ed.). Boston, MA: Heinle and Heinle.
Smith, L.C., & Mare, N.N. (2003). *Issues for today 3* (3rd ed.). Boston, MA: Heinle and Heinle.
Wholey, M. L. (2006). *Reading matters 2*. Boston, MA: Houghton Mifflin.

Appendix B

All cohesion-related variables considered in preliminary ANOVA

*denotes cohesion-related variables that were significant at the *p* < .005 level

Referential cohesion
Argument overlap
Stem overlap
Content word overlap adjacent sentences*
Anaphor reference all distances unweighted
Anaphor reference adjacent unweighted
Argument overlap all distances
Argument overlap adjacent

Latent Semantic Analysis (LSA)
LSA sentences
LSA paragraph to paragraph
LSA sentences all combinations mean
LSA sentence to sentence adjacent mean

Connectives
Connectives all
Incidence of negative logical operators
Incidence of positive logical operators

Incidence of negative additive connectives
Incidence of negative causal connectives*
Incidence of negative temporal connectives
Incidence of positive additive connectives
Incidence of positive causal connectives
Incidence of positive temporal connectives

Situation Model
Causal cohesion*
Incidence of causal verbs and particles*
Incidence of intentional action events and particles*
Mean of tense and aspect repetition scores

Appendix C

Descriptive statistics for 18 variables across textbook levels

Variable category	Variable	# of texts	Level 1 Mean (SD)	Level 2 Mean (SD)	Level 3 Mean (SD)
Referential cohesion	Proportion of content words that overlap between adjacent sentences	162	0.13 (0.05)	0.11 (0.04)	0.10 (0.03)
Connectives	Incidence of negative causal connectives	162	0.07 (0.54)	0.67 (1.24)	1.06 (1.26)
Situation model	Ratio of causal particles to causal verbs	162	0.43 (0.17)	0.60 (0.30)	0.71 (0.28)
	Incidence of causal verbs and particles	162	72.40 (16.54)	63.99 (12.54)	59.66 (10.63)
	Incidence of intentional action events and particles	162	31.86 (13.25)	22.23 (8.66)	17.98 (8.35)
	Mean of tense and aspect repetition scores	162	0.86 (.07)	0.83 (0.09)	0.81 (0.06)
Syntactic similarity	Sentence similarity across all paragraphs	162	0.14 (0.03)	0.11 (0.02)	0.09 (0.02)
Syntactic complexity	Mean number of modifiers per noun-phrase	162	0.77 (0.15)	0.81 (0.14)	0.88 (0.14)
	Noun phrase incidence score	162	317.39 (21.93)	296.93 (22.10)	280.35 (17.19)
Descriptive	Average words per sentence	162	11.23 (2.15)	14.38 (2.06)	17.11 (2.93)
	Number of words	162	306.24 (95.22)	599.80 (248.49)	824.48 (379.78)
	Average word frequency for all words	162	2.40 (0.14)	2.32 (0.15)	2.25 (0.13)

			122.04	62.50	93.74
	Average minimum word frequency-in-sentences	162	(240.24)	(115.18)	(389.20)
			69.49	58.52	48.84
	Person pronoun incidence score	162	(30.70)	(29.32)	(24.05)
			5.52	7.55	9.52
Readability	Flesch Kincaid Grade Level	162	(1.53)	(1.75)	(1.69)
			75.51	66.45	56.60
	Flesch Reading Ease Score	162	(8.65)	(10.75)	(9.87)

About the Authors

Lia Plakans is an associate professor of ESL and Foreign Language Education at the University of Iowa. She conducts research and teaches in the areas of second language reading and writing, language assessment, and language teaching methods. She has been an English language educator in Iowa, Ohio, Texas, and Latvia. E-mail: lia-plakans@uiowa.edu

Zeynep Bilki is an Academic English language instructor at Kirkwood Community College. Her doctoral dissertation at the University of Iowa focused on cohesion and highly proficient second language readers of English. She has taught English as a Foreign Language in Turkey at secondary and tertiary levels. E-mail: zeynepbilki@gmail.com