Syntactic and Lexical Complexity of Reading Correlates with Complexity of Writing in Adults

Yellowlees Douglas[1] & Samantha Miller[2]

[1] Hough Graduate School of Business, Warrington College of Business Administration, University of Florida, Gainesville, USA

[2] Heavener School of Business, Warrington College of Business Administration, University of Florida, Gainesville, USA

Correspondence: Yellowlees Douglas, Warrington College of Business Administration, University of Florida, PO Box 117150 Gainesville, FL 32611-7150, USA. Tel: 1-352-273-3215.

Abstract

This study solicited a cohort of graduate students, aged 23-42, to identify their most frequently-read sources, including online content, newspapers, news magazines, genre and "literary" fiction, and general non-fiction. We analyzed text samples from sources and students' writing, using the Lexile® framework and software that computes 14 features of syntactic complexity. This study validates the accuracy of the Lexile® framework, based on more granular measures of syntactic complexity than the mean sentence length Lexile uses as part of its scoring mechanism. More significantly, this study found strong correlations between lexical and syntactic complexity of student reading materials and students' writing.

Keywords: reading and writing, MBA programs, syntactic priming, organizational communication

1. Introduction

1.1 The Influence of Reading on Writing Skills

Researchers have taught us so much about the relationship between reading and writing. For instance, we know that the acquisition of reading skills often tracks closely with writing skills and that the inter-relationship between markers of sophistication in reading and writing become more pronounced well into students' secondary education (Loban, 1963). We also know that, when acquiring reading and writing skills, good writers spend more leisure time per week reading than poor writers (Monk, 1958). Researchers have also documented the influence of reading on domain-specific knowledge that enhances the complexity and sophistication of students' writing (McCutchen, 1986). Furthermore, extensive studies, beginning in the 1940s, demonstrated that reading deficits and developmental delays predicted similar deficits in writing (Schoneil, 1942). And a rich body of research has identified the characteristics of second-language learners, including measuring the impact of facility in students' native language with syntactic complexity exhibited in their second language (L2) (Lu & Ai, 2015).

A parallel body of literature offers compelling data on the influence on our own speech of the language we find ourselves surrounded by (Niederhoffer & Pennebaker, 2002). In fact, even the fleeting grammatical constructions of conversation cause speakers to engage in linguistic synchrony, with speakers' conversational styles unconsciously mirroring one another's (J. W. Pennebaker, 2011). Speakers taking turns may repeat the same grammatical constructions (Martin J Pickering & Victor S Ferreira, 2008). More strikingly and pertinent to educators, writers may unconsciously mimic the content of their own reading (Ireland & Pennebaker, 2010).

1.2 Existing Literature on Interactions between Reading and Writing

But these two directions in research, surprisingly, fail to converge on the one place where, together, they could shed significant and immensely useful light: does the syntactic and lexical complexity of our reading influence the sophistication of our writing—long after we, as adults, have already become expert writers? This study, conducted among a cohort of students in a Master of Business Administration program, aimed to establish the extent to which our most-frequently read weekly content—from *The Wall Street Journal* to literary fiction to BuzzFeed—might shape or even determine the complexity and sophistication of our written diction and syntax.

1.3 A Dearth of Data on the Influence of Reading on Writing in Adults

Surprisingly little data exists on the links between the reading material of adults and their writing. In fact, this study represents the first known data set that examines in detail the correlations between diction and syntax in reading and the same features in the writing of adults. However, this specific focus has largely remained off most researchers' radar for multiple reasons that involve the studies on reading and writing concentrating almost entirely on elementary and secondary education. In addition, the impact of adults' reading on their writing unfortunately falls into the grey area where developmental psychology, linguistics, and education overlap, the research equivalent of no-man's land.

To date, most studies of the links between reading and writing examine the acquisition of skill in elementary and secondary school students. Developmental psychologists have pegged high performance in reading comprehension to syntactic complexity in writing (Stotsky, 1983). Moreover, researchers have established that reading has a more profound effect on students' writing ability than the inverse (Shanahan & Lomax, 1986). Perhaps more tellingly, the sophistication of writing increased, the more students read at home in their leisure time (Monk, 1958). However, developmental psychologists wrestle with models for understanding the causal relationships between reading and writing, tackling everything from mastery of types of discourse (McCutchen, 1986) to the models for how young learners acquire mastery of reading and writing (Siedow, 1972). Within this line of research, the focus rests on a chicken-egg sort of conundrum: does reading merely give you content and models for how to write—or does facility in reading reflect some larger verbal sophistication that also impacts students' writing? Significantly, even when the focus shifts to the acquisition of skills in reading and writing, reading potentially exerts some influence on the quality of writing. At a developmental level, reading influences writing through writers mimicking in their writing the models gleaned from reading—or through the increasing complexity of reading content informing the sophistication of how readers write.

1.4 Synchrony or Priming between Reading and Writing Samples

Conversely, studies of influences on adults' linguistic patterns concentrate on mimicry in verbal exchanges, revealing strong influences across oral exchanges, regardless of participants' levels of identification with one another (Niederhoffer & Pennebaker, 2002; J. W. Pennebaker, 2011). Moreover, Ireland and Pennebaker (2010) even established high levels of mirroring in the language used by writers exchanging letters. In these studies, however, the Language Style Matching (LSM) described in detail by Pennebaker and his colleagues assessed emotions, evidence of personality traits, and socio-economic traces, all measurable by Pennebaker's own Linguistic Inquiry and Word Count (LIWC) software, which relies on computational linguistics to quantify speakers' use of pronouns, nouns, emotional terms, and function words (James W Pennebaker, Francis, & Booth, 2001).

Nevertheless, language studies demonstrate the impact of something similar to Pennebaker's LSM across the structure of sentences. For instance, the syntactic features of written sentences correlated highly with subjects' written descriptions of images, even when the experimenter's sentences contained unusual features, including double objects (J. K. Bock, 1986). Similarly, speakers repeat syntactic structures from questions, reflecting what Pickering and Ferreira have dubbed "structural priming," with passively constructed sentences triggering passively constructed replies and even unusual inversions of word orders in questions prompting the same inversions in responses (Martin J Pickering & Victor S Ferreira, 2008). Still other studies suggest both lexical and syntactic repetitions between the priming sentences, occurring across verbal and written conditions alike (K. Bock & Griffin, 2000; Pickering, Branigan, & McLean, 2002). Moreover, in laboratory studies, these priming effects proved surprisingly durable, enduring across as many as ten intervening prompts designed to disrupt recency effects (K. Bock, Dell, Chang, & Onishi, 2007). But how durable is the relationship between the content we read and the language we use on the page? And does this mimicry or structural and syntactic priming extend far beyond laboratory settings?

This study aimed to examine the long-term impacts of adults' regular reading materials on their writing. We hypothesized that the content of graduate students' reading would significantly impact their writing in both lexical and syntactic complexity, as well as the amount of time mature students dedicated to reading, possibly reflecting the influences of synchrony or priming between the words and sentence structures adult readers encountered most frequently.

2. Methods

2.1 Data Collection

The entire cohort of students (n=65) enrolled in either a one- or two-year Master of Business Administration program received an email soliciting their participation in an online survey. The brief survey required MBA students to report

1. amount of time they spent reading weekly,

2.   business-related news magazines they read regularly,

3.   newspapers they read regularly,

4.   regularly-read online news sources,

5.   type of reading material they read most frequently in an average week,

6.   sources they used most frequently to access news,

7.   frequency with which they read fiction.

In addition, to complete the survey, students copied and pasted a writing sample from a single paragraph. Since all students in the MBA cohort completed a required, module-long writing course, GEB 5212—Professional Writing in Business, they had all also completed the same assignment, writing a cover letter to apply for a job, a letter many students used in later securing their post-graduate employment. The survey specifically requested that students select the cover letter's second paragraph, one in which every student identified his or her greatest strengths and tied them to requirements or keywords identified in the descriptions of the positions for which they applied. This feature of the survey ensured the least possible variance across the writing samples' length, diction, and writing styles, as students also received coaching on job applications from the school's Graduate Business Career Services Center.

2.2 Reading Materials

Given our participants' status as MBA students, our survey focused on common online news sources, business-related periodicals, popular news-centric periodicals, best-selling and business-focused non-fiction, genre fiction, and critically acclaimed or canonical literary fiction. We used initial Lexile scores from examples of each type of reading material to select the reading materials included in our survey, providing a range of Lexile scores, ranging from those indicating significant ease of reading (620L) to demanding (1590L). Thus, our survey included a range of reading materials our participants were most likely to read regularly (Table 1), as well as indications of our participants' most common sources for reading weekly.

Table 1. Survey options for reading materials

Business-focused Magazines	General Interest Newspapers	Online Sources	Most Frequently Read Content
The Economist	*The Wall Street Journal*	BuzzFeed	Web-based content not published in print
Forbes	*New York Times*	Tumblr	Daily newspapers
Business Week	*USA Today*	Reddit	Weekly magazines
Money	*Gainesville Sun*	Huffington Post	Academic journals
Fast Company	(Don't read any newspaper regularly)	Medium	General interest non-fiction
(Don't read any weekly business magazines)		Ozy	Fantasy, Science-fiction or Mystery fiction
		(Don't read any of these sources regularly)	Literary Fiction
			(Don't read any of these sources regularly)

We also surveyed participants' frequency with which they read fiction more than six times per year and the average amount of time they spent reading non-course-specific material during the year in which they took the survey (2015).

Our analyses using the Lexile® Framework and the Ai's Web-based L2 Syntactic Complexity Analyzer provided scores for Lexile and Ai's fourteen measures for participants' writing samples. We also used the same measures, applied to samples from the survey's reading materials, by tracking a single story across all periodicals to reduce variance across samples. We deliberately limited our analysis of Lexile and syntactic factors to text selections featuring a single, developing news story: the Republican contest for the presidential election of 2016, as the story unfolded during late 2015. This narrowed focus avoided the wide range of variance we initially discovered when we analyzed even the same topic drawn from four different genres of news stories involving science, politics, business, and the arts or style features—all items our participants were likely to read. In contrast, when we analyzed selections

from only a single, developing news story, the Lexile and other scores of textual complexity showed the lowest range of variance across factors within periodicals.

We selected text samples for each reading type based on sources participants were likeliest to read. These sources included, for academic journals, *American Economic Review*, *Administrative Sciences Quarterly,* and other business-related journals, including *American Journal of Sociology*, and *Psychological Science*. Samples for best-selling or critically-acclaimed non-fiction reflected best-sellers, business-centric interests or books included in some participants' course readings (e.g., *Jane Austen, Game Theorist; The Big Short; Thinking, Fast and Slow; Unbroken*). Fiction included critically-acclaimed and best-selling books (e.g., *City on Fire, A Little Life*), while genre fiction samples came from books turned into television series or films (e.g., *World War Z, Songs of Fire and Ice, Outlander*). In all samples, we used an array of sources most likely read by our participants, with our selected online sources reflecting students' initial self-reported destinations for web-based news and feature stories (Tumblr, Reddit, Huffington Post, BuzzFeed, Ozy, and Medium). For the online-only sources, we also tracked the same news story we used for sampling from print and online periodicals, the Republican contest for the 2016 presidential nomination.

2.3 Data Analysis

To assess writing sophistication at the sentence level, this study used two different platforms for analyzing textual complexity: the Lexile® framework and Ai's Web-based L2 Syntactic Complexity Analyzer . Developed to identify the extent to which fluency in native languages influences fluency in L2 writing, the L2 Syntactic Complexity Analyzer provides the most granular and comprehensive identification of writing samples via analysis of fourteen syntactic features, as displayed in Table 2 (Lu, 2010, 2011; Lu & Ai, 2015). To assess lexical coupled with syntactic complexity, this study also employed the Lexile® framework that assesses sematic difficulty and syntactic complexity of text samples, using a corpus of over 600 million words to measure semantic difficulty via the infrequency with which readers would encounter words (Lennon & Burdick, 2004). In addition to its widespread use in elementary and secondary education, the Lexile® framework also provides readability indicators for over 100,000,000 books and articles (MetaMetrics, 2015), including many of the newspapers and journals included in the survey questions. But, in contrast to the comprehensive measurement of Ai's L2 Syntactic Complexity Analyzer, the Lexile® framework analyzes only sentence length. The Lexile® framework operates under the assumption that longer sentences will always contain more syntactic sophistication than shorter sentences, reflected in greater numbers of phrases and clauses, as well as in use of subordination and complex nominals (Izquierdo & Bailey, 1998; Lennon & Burdick, 2004). Together these two means of semantic and syntactic measurement provided a comprehensive assessment of the sophistication and complexity of both our students' writing samples and of the sources they read.

Table 2. Syntactic factors analyzed by Ai's L2 Syntactic Complexity Analyzer

Measure	Code	Definition
Length of production unit		
Mean length of clause	MLC	# of words/# of clauses
Mean length of sentence	MLS	# of words/# of sentences
Mean length of T-unit	MLT	# of words/# of T-units
Amount of subordination		
Clauses per T-unit	C/T	# of clauses/# of T-unit
Complex T-units per T-unit	CT/T	# of complex T-units/# of T-units
Dependent clauses per clause	DC/C	# of dependent clauses/# of clauses
Dependent clauses per T-unit	DC/T	# of dependent clauses/# of T-units
Amount of coordination		
Coordinate phrases per clause	CP/C	# of coordinate phrases/# of clauses
Coordinate phrases per T-unit	CP/T	# of coordinate phrases/# of T-units
T-units per sentence	T/S	# of T-units/# of sentences
Degree of phrasal sophistication		
Complex nominals per clause	CN/C	# of complex nominals/# of clauses
Complex nominals per T-unit	CN/T	# of complex nominals/# of T-units
Verb phrases per T-unit	VP/T	# of verb phrases/# of T-units
Overall sentence complexity		
Clauses per sentence	C/S	# of clauses/# of sentences

We then analyzed the data for both Lexile and Ai's factors, using regression models and Spearman's nonparametric correlations, as well as Pearson's correlations. We subsequently narrowed our regression and correlational analyses to focus specifically on the factors with significant correlations across writing and reading samples: Lexile, Mean Length Sentence (MLS), Mean Length Clause (MLC), and Complex Nominals per T-unit (CN/T). These factors also correspond to the literature analyzing texts' sophistication (Graesser et al., 2014; Graesser, McNamara, Louwerse, & Cai, 2004; Izquierdo & Bailey, 1998) that specifically link mean length of clauses and sentences and use of complex nominal as robust indicators of either writers' sophistication or of difficulty of reading, working most effectively in assessing informational, rather than narrative texts (Graesser et al., 2014). We used Spearman's and Pearson's correlations to examine relationships between

1. Lexile and MLS, MLC and CN/T;

2. Lexile and MLS, MLC, and CN/T in reading and writing samples;

3. Lexile and MLS, MLC and CN/T in average reading type and participants' writing samples;

4. Lexile and MLS, MLC and CN/T and self-reporting reading of fiction regularly;

5. Lexile and MLS, MLC and CN/T and number of hours of weekly reading.

2.4 Choice of Tools for Data Analysis

This study specifically eschewed common readability tools for analyzing texts, including Flesch's Reading Ease Score (Flesch, 1948, 1974) or the Flesch-Kincaid score (Kincaid, 1975), as both tools rely on the number of syllables per word in sentences. Within this framework for assessing difficulty, the word *praxis* receives the same weight in terms of semantic sophistication as another two-syllable word, *model*, despite the semantic challenges thrown up by a word readers typically encounter infrequently—*praxis*, which, by most researchers' definitions, makes any sentence containing *praxis* semantically more challenging than a sentence containing *model*. The Lexile® framework redresses this deficit by relying on its extensive corpus to assess the frequency with which readers are likely to encounter a word, connecting the words readers encounter most frequently with greater ease of reading, as readers process familiar words more easily than unfamiliar words (Goldman & Lee, 2014; Hiebert, 2011; Hiebert & Mesmer, 2013).

In addition, this study also eschewed analyzing data with Coh-Metrix (Graesser et al., 2004) to avoid potentially skewing analyses of students' writing relative to the content of their standard chosen reading. In the writing course from which we obtained student text samples, course instruction specifically requires students to address the same issues Coh-Metrix uses to assess the readability of texts. In particular, the standard course text, *The Reader's Brain: How Neuroscience Can Make You a Better Writer* (Douglas, 2015) stresses that highly readable writing displays at least three of the five factors Coh-Metrix employs in measuring texts' readability: word concreteness, referential cohesion, and causal cohesion (Graesser et al., 2014). As a result, the course methodology on writing could act as a confounder, creating disparities between scoring of students' writing samples and scores from texts drawn from sources like *Forbes,* the *New York Times, Wall Street Journal,* or *The Economist*, created by writers who most likely lack explicit knowledge of the interaction between, say, readability and referential cohesion. In addition, our students' performance on Coh-Metrix would merely assess the efficacy of our coaching our students on coherence and cohesion in their writing—not the influence of their reading on the lexical and syntactic complexity of their writing.

3. Results

From a total sample of size of 65 MBA students, our survey elicited a response rate of 74%, with participants' ages ranging from 23 to 42 years, providing a sample of still-burgeoning as well as developmentally mature reading and writing skills. In addition, this broad age range also ensured our survey included both Millennials and Gen X participants (Twenge, 2014), potentially providing a glimpse into the influences on students' writing from both traditional reading materials—periodicals and books—and authorless, crowd-sourced content online that included popular sources like Tumblr, Reddit, and BuzzFeed.

Our results establish the reliability of Lexile and of syntactic factors most frequently associated with textual sophistication. This study correlated the Lexile® Framework with well-established markers of textual sophistication at the syntactic level: MLC at 0.52 (p=0.0002), MLS at 0.47 (p=0.0013), and CN/T at 0.92 (p=<.0001). On the other hand, reading fiction regularly and reading time failed to show any significant correlations with either Lexile or the three primary measures of syntactic sophistication. In contrast, average reading type correlated significantly on three of the four measures between average reading type and writing samples. In correlations between participants' average reading sources and their writing, Lexile showed the highest correlation (p=0.005), MLS had a strong

correlation (p=0.028) and MLC also correlated significantly (p=0.030). Only CN/T failed to display statistical significance between reading type and writing samples (p=0.067).

4. Discussion

4.1 Lexile Proves a Robust Predictor of Textual Sophistication

This study is the first known to validate the Lexile® framework by correlating its fairly crude tools for assessing reading—word frequency and mean sentence length—to multiple and more granular measures that assess the syntactic complexity of writing. This validation of Lexile®'s accuracy in measuring the syntactic sophistication of texts is important, given the ubiquity of Lexile® scoring of reading materials, from primary through higher education. In addition, many databases now routinely pair Lexile® scores with articles from a wide variety of periodicals, including major newspapers and weekly news-related magazines. Although MLS and MLC are also simple measures of syntactic sophistication, both measures robustly indicate the complexity of sentences, as short sentences are statistically least likely to contain embedded phrases and clauses or subordination (Crossley & McNamara, 2011; Goldman & Lee, 2014; Sheehan, Kostin, & Futagi, 2008; Smith, 1989). In particular, Lexile correlates most strongly with the occurrence of complex nominals. In their simplest form, complex nominals consist of modifier preceding a head noun (Levi, 1978), with either another noun or a pseudo-adjective—an adjective that cannot function as a predicate (Izquierdo & Bailey, 1998). For instance, a *civil engineer* is a type of engineer, rather than an engineer who behaves in a civil manner. Complex nominals enable writers to compress information, the likeliest reason why complex nominal appear most frequently in technical writing (Price, 1974).

In our study, students' writing failed to correlate significantly with reading only on complex nominals per T-unit (CN/T). In this lack of correlation, our participants' writing may reflect more the direct influence of the course in which they created their writing samples than the content of their most frequently-read sources. This graduate business writing course specifically counsels writers to avoid complex nominals containing more than three nouns, a practice advised by some experts on writing (Douglas, 2015; Price, 1974). This practice is also well supported by research on the difficulty with which readers process long complex nominal strings, especially when complex nominals group as many as four or more nouns acting as modifiers, causing even native speakers to misidentify the head noun (Gordon, Hendrick, & Johnson, 2004; Limaye & Pompian, 1991). As a result, this same compression in writing can result in garden path mis-readings that require readers to re-read the noun phrase to identify the correct head noun, making for more challenging reading and, thus, higher Lexile scores. Our readers were most likely to encounter high numbers of complex nominals in academic journals, where writers are likeliest to emulate the writing in other academic journals. Conversely, readers were least likely to find complex nominals in genre fiction that promise immersive narratives, designed for maximum ease of reading (Biber & Gray, 2010; Hadidi & Ghandkaran-Shotorban, 2015; Parkinson & Musgrave, 2014). As a result, participants' writing samples would more directly display the effects of the writing course and its psycholinguistic methods of teaching writing than they would the use of complex nominals in their reading. Participants who completed the survey were also likely to be students who performed well in the course, a performance somewhat reflected by a relative absence of complex nominals in their writing.

4.2 Lexical and Syntactic Complexity of Reading Directly Influence Writing

The complexity of our reading powerfully influences our writing, evident in the strong correlations between the complexity of reading materials and mature students' writing, reflected in Lexile, MLS, and MLC. However, on all other measures, writing samples reflected the direct impacts of their average reported reading materials over the writing course methodology, which attempted to steer students away from writing longer, compound-complex sentences in favor of shorter, more easily-read sentences. The context in which our participants wrote their samples makes this study broadly applicable to writing across disciplines and educational and professional settings. The required writing course's methodology requires students to avoid needless lexical and syntactic complexity, avoiding psycholinguistic challenges to readers like complex nominals or compound-complex sentences that put excessive demands on readers' working memories (Douglas, 2015). Writers can more easily simplify complex sentences than they can modify overly simplistic diction and syntax, as the complexity of writing, particularly in the use of phrases, increases as student writing matures (Staples, Egbert, Biber, & Gray, 2016).

Several compelling implications also arise from our findings, involving the amount of reported weekly reading time and the absence of a correlation between lexical and syntactic measures of sophistication in reading materials and writing samples. First, as students in an MBA program, our participants' reporting reading time is likely lower than the amount of time reported by non-students, as our participants dedicated most of their time to completing course work, reading some textbooks, and working with quantitative data. The scarcity of time students experience in a one-

or two-year MBA program would directly impact their self-reported reading time weekly, accounting for the absence of a correlation between the amount of time spent reading and the syntactic and lexical complexity of their writing. Conversely, the absence of the correlation between time reading and writing complexity could also reflect the result of long-term patterns in reading, stretching across years or decades in adults' lives, that directly shape writers' diction and syntax, rather than any immediate effects of the amount of time adults spend reading in any particular year.

Of the theories attempting to forge causal connections between reading and writing, synchrony or even mimicry (Ireland & Pennebaker, 2010) seems likeliest to account for the influence of reading content on writing. As students accepted into a competitive MBA program, our participants already had high-functioning reading skills, which would have played a relatively weak role in determining their choice of reading content. In fact, our survey revealed that participants regularly read a wide range of content, in terms of its complexity, with participants who read *The Economist* and *New York Times* also reading Reddit, Tumblr, and BuzzFeed weekly. Nevertheless, clearly, the complexity of their most commonly-read sources exerted the strongest influence on our participants' writing, with participants who read web-only content scoring lowest on Lexile and syntactic measures of complexity.

Several other explanations, not commonly seen in the literature on the relationship between reading and writing, might also account for such a strong influence of reading content on writing sophistication. First, this study could also demonstrate the powerful effects of priming outside of a laboratory setting, with the vocabularies and sentence structure from frequently-read content shaping the word choices and syntax in writing. Second, our participants' samples could also reflect the influence of availability or strength of associations between texts participants read and texts they created (Horowitz, Norman, & Day, 1966). In fact, availability could also account for the surprising durability of the effects of priming or the strength of synchrony seen both inside laboratory settings (J. K. Bock, 1986; K. Bock et al., 2007; K. Bock & Griffin, 2000; Hartsuiker & Westenberg, 2000; Niederhoffer & Pennebaker, 2002; J. W. Pennebaker, 2011; M. J. Pickering & V. S. Ferreira, 2008) and in studies of correspondence between writers stretching across decades (Ireland & Pennebaker, 2010).

4.3 Implications for Education

Consistent with some findings on the writing sophistication of students in elementary and secondary education (DeVries, 1970; Heys, 1962; Jenkins, Johnson, & Hileman, 2004), the strong correlations between the types of reading preferred by students in higher education and their writing samples argue for the influence of reading on writing. This study's findings may also explain why, despite increasing attention to students' writing skills, on both admissions essays and required communication requirements or courses, students' writing still fails to live up to employers' expectations (Christiansen, 1965; Council, 2016; Employers, 2011; Hynes & Sigmar, 2009; Korkki, 2007; Kozel, 2010). Graduates' facility in writing may stem less from direct instruction in writing or even frequent practice in writing but from the sophistication of their reading content across their years in higher education and their reading habits, post-graduation(Christiansen, 1965). Moreover, in most schools and colleges of business, textbooks occupy an outsized place in the curriculum, with courses requiring few readings from academic journals or from books. Many educators may shy away from assigning even books written by well-regarded researchers, drawing from their own peer-reviewed research, viewing best-selling books as hewing too closely to "pop" sensibilities—as one colleague labeled Daniel Kahneman's *Thinking, Fast and Slow,* despite the book's foundations in the academic articles that won Kahneman the Nobel Prize in Economics (Kahneman, 2011). As a result, our students' and graduates' writing may reflect the poorly-edited, jargon-ridden sentences of textbooks that focus on quantitative-based research or the simplified prose of many business communication textbooks, which seek to meet the readability standards common to syndicated newspapers (Raze, 1969). Ironically, the reading fodder we supply most frequently to our students may more closely resemble the simplified sentences our students also read on BuzzFeed and Tumblr than the sentences they would encounter in, say, *The Big Short* (Lewis, 2010): Lexile, 1020; MLS, 15.7778; MLC, 14.2; CN/T, 2.3333. In contrast, *Business Communication: Polishing Your Professional Presence* (Shwom & Snyder, 2016) offers significantly lower lexical and syntactic complexity: Lexile, 790; MLS, 10.4667; MLC, 11.2143; CN/T, 1.125. For comparison, the *Huffington Post*, which featured heavily in the preferred reading of our lowest-scoring writers, has lower lexical and syntactic complexity (Lexile, 890; MLS, 16.5; MLC, 8.25; CN/T, 1.7778) than the *New York Times* (Lexile, 1410; MLS, 25.15; MLC, 14.3714; CN/T, 3.1818) but higher scores on all measures than *Business Communication*. Our students may write below institutional and employer expectations because we assign textbooks so poorly written that our students skim them purely for content, or so simplistically written that, as writers, our students' writing reflects more strongly the online content they commonly read or even reflects the dimly-recalled texts assigned during secondary education.

Studies already suggest that merely assigning more writing impacts students' writing skill less directly than the

amount of reading they pursue both inside and outside of classrooms (Christiansen, 1965; Heys, 1962; Loban, 1963; Shanahan & Lomax, 1986; Stotsky, 1983). Our findings suggest that we should incorporate more non-textbook readings into our courses and curricula and also stress the importance of lifelong reading habits and material. If we can tie our students' writing facility directly to their choices of reading materials, we can begin to bolster the sophistication of our students' writing and also make them more mindful of their choices in reading.

5. Conclusions

This study has several key implications that should guide future research. First, Lexile® is a reliable measure of the lexical and syntactic complexity of reading materials, even for adult readers. Second, the syntactic complexity of reading materials exerts a strong impact on writers' word choice and typical sentence structure. Third, some types of eomplexity can make reading more challenging without being conducive to the perception of a text as well written. Speificially, complex nominals, uncommon in most written English prior to the 1970s, rely on noun phrases, noun clauses, and nouns functioning as adjectives—a "noun speak" used in technical writing that slows down reading speeds due entirely to the cognitive loads imposed by complex nominals (Izquierdo & Bailey, 1998; Levi, 1978).

References

Biber, D., & Gray, B. (2010). Challenging stereotypes about academic writing: Complexity, elaboration, explicitness. *Journal of English for Academic Purposes, 9*(1), 2-20. http://dx.doi.org/10.1016/j.jeap.2010.01.001

Bock, J. K. (1986). Syntactic persistence in language production. *Cognitive Psychology, 18*(3), 355–387. http://dx.doi.org/10.1016/0010-0285(86)90004-6

Bock, K., Dell, G. S., Chang, F., & Onishi, K. H. (2007). Persistent structural priming from language comprehension to language production. *Cognition, 104*(3), 437-458. http://dx.doi.org/10.1016/j.cognition.2006.07.003

Bock, K., & Griffin, Z. M. (2000). The persistence of structural priming: Transient activation or implicit learning? *Journal of experimental psychology: General, 129*(2), 177-192. http://dx.doi.org/10.1037/0096-3445.129.2.177

Christiansen, M. (1965). Tripling writing and omitting reading in freshman English: An experiment. *College Composition and Communication, 16*, 122-124. http://dx.doi.org/10.2307/354888

Crossley, S. A., & McNamara, D. S. (2011). Understanding expert ratings of essay quality: Coh-Metrix analyses of first and second language writing. *International Journal of Continuing Engineering Education and Life Long Learning, 21*(2-3), 170-191. http://dx.doi.org/10.1504/IJCEELL.2011.040197

DeVries, T. (1970). Reading, Writing Frequency, and Expository Writing. *Reading Improvement, 7*, 14-19.

Douglas, Y. (2015). *The Reader's Brain: How Neuroscience Can Make You a Better Writer*. Cambridge: Cambridge University Press.

Flesch, R. F. (1948). A new readability yardstick. *Journal of Applied Psychology, 32*, 221-233. http://dx.doi.org/10.1037/h0057532

Flesch, R. F. (1974). *The Art of Readable Writing with the Flesch Readability Formula*. New York: Harper & Row.

Goldman, S. R., & Lee, C. D. (2014). Text Complexity. *The Elementary School Journal, 115*(2), 290-300. http://dx.doi.org/10.1086/678298

Gordon, P. C., Hendrick, R., & Johnson, M. (2004). Effects of noun phrase type on sentence complexity. *Journal of Memory and Language, 51*(1), 97-114. http://dx.doi.org/10.1016/j.jml.2004.02.003

Graduate Management Admissions Council. (2016). How to use the analytical writing assessment score. Retrieved from http://www.gmac.com/gmat/understand-gmat-exam-scores/how-to-use-the-analytical-writing-assessment-score. aspx

Graesser, A. C., McNamara, D. S., Cai, Z., Conley, M., Li, H., & Pennebaker, J. (2014). Coh-Metrix measures text characteristics at multiple levels of language and discourse. *The Elementary School Journal, 115*(2), 210-229. http://dx.doi.org/10.1086/678293

Graesser, A. C., McNamara, D. S., Louwerse, M. M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior research methods, instruments, & computers, 36*(2), 193-202. http://dx.doi.org/10.3758/BF03195564

Hadidi, Y., & Ghandkaran-Shotorban, S. (2015). Investigating Possible Evolutionary Reduction in Subject Nominal Group Complexity in Prose Fiction Text. *Journal for the Study of English Linguistics, 3*(1), 1-12.

http://dx.doi.org/10.5296/jsel.v3i1.7038

Hartsuiker, R. J., & Westenberg, C. (2000). Word order priming in written and spoken sentence production. *Cognition, 75*(2), B27-B39. http://dx.doi.org/10.1016/j.jml.2007.07.003

Heys, F. (1962). The Theme-A-Week Assumption: A Report of an Experiment. *English Journal, 51*, 320-322. http://dx.doi.org/10.2307/810007

Hiebert, E. H. (2011). Beyond Single Readability Measures: Using Multiple Sources of Information in Establishing Text Complexity. *Journal of Education, 191*(2), 33-42.

Hiebert, E. H., & Mesmer, H. A. E. (2013). Upping the ante of text complexity in the Common Core State Standards examining its potential impact on young readers. *Educational Researcher, 42*(1), 44-51. http://dx.doi.org/10.3102/0013189X12459802

Horowitz, L. M., Norman, S. A., & Day, R. S. (1966). Availability and associative symmetry. *Psychological Review, 73*(1), 1-15. http://dx.doi.org/10.1037/h0022661

Hynes, G., & Sigmar, L. (2009). *The importance of business communication courses in employer hiring decisions.* Paper presented at the First annual general business conference, conference proceedings Huntsville, TX.

Ireland, M. E., & Pennebaker, J. W. (2010). Language style matching in writing: synchrony in essays, correspondence, and poetry. *Journal of personality and social psychology, 99*(3), 549-571. http://dx.doi.org/10.1037/a0020386

Izquierdo, B. D., & Bailey, D. (1998). Complex noun phrases and complex nominals: Some practical considerations. *TESL Reporter, 31*(1), 19-29.

Jenkins, J. R., Johnson, E., & Hileman, J. (2004). When is reading also writing: Sources of individual differences on the new reading performance assessments. *Scientific Studies of Reading, 8*(2), 125-151. http://dx.doi.org/10.1207/s1532799xssr0802_2

Kahneman, D. (2011). *Thinking, Fast and Slow*. New York: Farrar, Straus and Giroux.

Kincaid, J. P., Fishburne, R.P., Rogers, R.L., & Chissom, B.S. (1975). *Derivation of new readability formulas (Automated readability index, fog count and Flesch reading ease formula) for Navy enlisted personnel*. (No. RBR-8-75). Millington TN: Naval Technical Training Command Research Branch.

Korkki, P. (2007). Young workers: U nd 2 improve ur writing skills. . *New York Times*. Retrieved from http://www.nytimes.com/2007/08/26/business/yourmoney/26count.htm

Kozel, B. (2010). Writing skills: A crisis as close as your Inbox. *Employment Relations Today, 36*(4), 7-12. http://dx.doi.org/10.1002/ert.20268

Lennon, C., & Burdick, H. (2004). The lexile framework as an approach for reading measurement and success. Retrieved from http://www.lexile.com

Levi, J. N. (1978). *The Syntax and Semantics of Complex Nominals*. New York: Academic Press.

Lewis, M. (2010). *The Big Short: Inside the Doomsday Machine*. New York: W. W. Norton.

Limaye, M., & Pompian, R. (1991). Brevity versus clarity: The comprehensibility of nominal compounds in business and technical prose. *Journal of Business Communication, 28*(1), 7-21. http://dx.doi.org/10.1177/002194369102800102

Loban, W. (1963). *The Language of Elementary School Children*. Urbana, IL: National Council of Teachers of English.

Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics, 15*(4), 474-496. http://dx.doi.org/10.1075/ijcl.15.4.02lu

Lu, X. (2011). A Carpus-Based Evaluation of Syntactic Complexity Measures as Indices of College-Level ESL Writers' Language Development. *Tesol Quarterly, 45*(1), 36-62. http://dx.doi.org/10.5054/tq.2011.240859

Lu, X., & Ai, H. (2015). Syntactic complexity in college-level English writing: Differences among writers with diverse L1 backgrounds. *Journal of Second Language Writing, 29*, 16-27. http://dx.doi.org/10.1016/j.jslw.2015.06.003

McCutchen, D. (1986). Domain knowledge and linguistic knowledge in the development of writing ability. *Journal of Memory and Language, 25*(4), 431-444. http://dx.doi.org/10.1016/0749-596X(86)90036-7

McNamara, D. S., Crossley, S. A., & McCarthy, P. M. (2010). Linguistic features of writing quality. *Written Communication, 27*(1), 57-86. http://dx.doi.org/10.1177/0741088309351547

MetaMetrics. (2015). Lexile Infographic. Retrieved from https://www.lexile.com/about-lexile/lexile-overview/lexile-infographic/

Middleton, D. (2011). Students struggle for words: Business schools put more emphasis on writing amid employer complaints. *The Wall Street Journal*. Retrieved from http://online.wsj.com/article/SB10001424052748703409904576174651780110970

Monk, R. H. J. (1958). *A Study to Determine the Relationship between Children's Home Environments and their School Achievemeni in Written English*. (PhD.), University of Washington, Ann Arbor, MI: University Microfilms.

National Association of Colleges and Employers. (2011). *Job Outlook Survey 2011*. Retrieved from http://www.naceweb.org

Niederhoffer, K. G., & Pennebaker, J. W. (2002). Linguistic synchrony in social interaction. *J. Lang. Soc. Psychol, 21*(4), 337-360. http://dx.doi.org/10.1177/026192702237953

Parkinson, J., & Musgrave, J. (2014). Development of noun phrase complexity in the writing of English for Academic Purposes students. *Journal of English for Academic Purposes, 14*, 48-59. http://dx.doi.org/10.1016/j.jeap.2013.12.001

Pennebaker, J. W. (2011). *The Secret Life of Pronouns: What Our Words Say about Us*. New York: Bloomsbury Press.

Pennebaker, J. W., Francis, M. E., & Booth, R. J. (2001). Linguistic inquiry and word count: LIWC 2001. *Mahway: Lawrence Erlbaum Associates, 71*, 2001.

Pickering, M. J., Branigan, H. P., & McLean, J. F. (2002). Constituent structure is formulated in one stage. *Journal of Memory and Language, 46*(3), 586-605. http://dx.doi.org/10.1006/jmla.2001.2824

Pickering, M. J., & Ferreira, V. S. (2008). Structural priming: A critical review. . *Psychological bulletin, 134*, 427-459. http://dx.doi.org/10.1037/0033-2909.134.3.427

Price, B. D. (1974). Noun overuse phenomenon article. *The Language Quarterly, 2*(4), 29-37.

Raze, T. A. (1969). A study of American newspaper readability. *Journal of Communication, 19*(4), 317-324. http://dx.doi.org/10.1111/j.1460-2466.1969.tb00854.x

Schoneil, F. (1942). *Backwardness in the Basic Subjects*. Toronto: Clarke, Irwin.

Shanahan, T., & Lomax, R. G. (1986). An analysis and comparison of theoretical models of the reading–writing relationship. *Journal of Educational Psychology, 78*(2), 116-123. http://dx.doi.org/10.1037/0022-0663.78.2.116

Sheehan, K. M., Kostin, I., & Futagi, Y. (2008). When do standard approaches for measuring vocabulary difficulty, syntactic complexity and referential cohesion yield biased estimates of text difficulty? Paper presented at the Proceedings of the 30th Annual Conference of the Cognitive Science Society, Washington DC.

Shwom, B., & Snyder, L. G. (2016). *Business Communication: Polishing Your Professional Presence*. New York: Pearson.

Siedow, M. (1972). *Relationships between Syntactic Maturity in Oral and Written Language and Reading Comprehension of Materials of Varying Syntactic Complexity*. (Ph.D.), Indiana University, Dissertation Abstracts International 34 (1973). (4890A)

Smith, D. R. (1989). The Lexile Scale in Theory and Practice. Final Report. Bethesda, MD: National Institute of Health. Retrieved from ERIC: ED307577.

Staples, S., Egbert, J., Biber, D., & Gray, B. (2016). Academic writing development at the university level: Phrasal and clausal complexity across level of study, discipline, and genre. *Written Communication, 33*(2), n.p. (Published online ahead of print.). http://dx.doi.org/10.1177/0741088316631527

Stotsky, S. (1983). Research on reading/writing relationships: A synthesis and suggested directions. *Language arts, 60*(5), 627-642.

Twenge, J. (2014). *Generation Me: Why Today's Young Americans Are More Confident, Assertive, Entitled—And More Miserable than Ever Before* (2nd ed.). New York: Atria.