



Contents lists available at ScienceDirect

English for Specific Purposes

journal homepage: <http://ees.elsevier.com/esp/default.asp>

A multi-dimensional analysis of the Management's Discussion and Analysis narratives in Chinese and American corporate annual reports

Chaowang Ren^a, Xiaofei Lu^{b,*}^a School of Foreign Languages, Guangdong University of Technology, No. 100 Waihuan Xi Road, Guangzhou, Guangdong, 510006, China^b Department of Applied Linguistics, The Pennsylvania State University, 234 Sparks Building, University Park, PA, 16802, USA

ARTICLE INFO

Article history:

Available online 3 January 2021

Keywords:

Management's Discussion and Analysis

Multi-dimensional analysis

Genre features

Comparative analysis

© 2020 Elsevier Ltd. All rights reserved.

1. Introduction

The annual report (AR) is a document used by most public companies to disclose important corporate information to shareholders and the general public. The importance of ARs may be grasped from the following quote from Bill Gates (1996, p. 148): “Warren Buffett follows his own advice: When he invests in a company, he likes to read all of its annual reports going back as far as he can. He looks at how the company has progressed and what its strategy is.”

The importance of narrative disclosures in ARs was highlighted in the guide published in 1998 by the U.S. Securities and Exchanges Commission (SEC), *A plain English handbook: How to create clear SEC disclosure documents* (Beattie, McInnes, & Fearnley, 2004). In particular, the SEC views the Management's Discussion and Analysis (MD&A) narratives as required disclosures in an AR that are essential for helping investors understand the financial status of a firm (Hajek & Henriques, 2017). Additionally, MD&A narratives are also considered a useful source of information for financial analysts (Clarkson, Kao, & Richardson, 1999). In an effort to conform to international practices and norms, the China Securities Regulatory

* Corresponding author. Tel.: +1 (814) 865-4692.

E-mail addresses: paulren@gdut.edu.cn (C. Ren), xxl13@psu.edu (X. Lu).

Commission (2012) introduced the requirement to include MD&A narratives as a separate item (Item 33) in periodic reports and prospectuses in 2002.

Given the importance of MD&A narratives in ARs, much emphasis has been placed on their content, functions, and relationship to corporate performance in the accounting and management literature (e.g., Dutta, Fuksa, & Macaulay, 2019; Merkl-Davies & Brennan, 2007; Rutherford, 2002). Meanwhile, the linguistic features of MD&A narratives have also received increasing attention, as they have implications for readers' perception of the quality of those narratives (e.g., Dutta et al., 2019; Loughran & McDonald, 2016). Studies of this latter type have focused primarily on public companies in the U.S., while few studies have looked at companies in China or compared them against international companies. A systematic analysis of the linguistic features of MD&A narratives of public companies in China and how they compare to those of international public companies, however, will offer useful insights into the extent to which the linguistic practices of Chinese companies align with or differ from international practices and norms when it comes to producing MD&A narratives in English.

The current research aims to address this research gap by comparing the linguistic features of MD&A narratives from listed companies in China and the U.S. using a Multi-Dimensional (MD) analysis (Biber, 1988, 1995) approach. As a corpus linguistic technique, MD analysis is especially appropriate for investigating systematic textual variations between different groups of texts at a macro level, and its benefits have not yet been taken full advantage of by previous studies of the linguistic features of MD&A narratives. Our findings on the systematic linguistic similarities and differences between the two groups of MD&A narratives will have useful implications for understanding the linguistic practices of Chinese companies in producing MD&A narratives in English in relation to those of American companies. They will also be useful for informing the teaching of Business English as English for Specific Purposes (ESP) in the context of China.

1.1. Management's Discussion and Analysis in the annual report

Narrative disclosures in the AR refer to content other than the financial statements, notes and auditor's report in the AR (Rowbottom & Lymer, 2010). MD&A is a required type of narrative disclosure in the AR that describes the company's history of operations, provides necessary background information to explain the financial statements and other statistical data in the AR, and forecasts the company's potential risks and future business prospects (Securities and Exchange Commission, 2017). More business-related than other types of narrative disclosure, MD&A narratives offer an important channel to understand the management's ideas about business performance and corporate development.

It has been claimed that "consistent with the SEC's intention, the MD&A is arguably the most read and most important component of the financial section" (Li, 2010, p. 1050), as it serves to enhance investors' understanding of the company's operation and financial condition. Among the disclosure items in the AR, the MD&A is also relied upon most frequently by financial analysts in the U.S. when preparing their reports (Rogers & Grant, 1997). The value of the MD&A section to both investors and financial analysts speaks to the importance of research to examine its textual features, content quality, and relationship to corporate performance.

1.2. Previous research on narrative disclosures

There is research consensus that narrative disclosures are highly important to AR readers (e.g., Beattie et al., 2004; Rutherford, 2002). Previous research showed that narrative disclosures in the AR were extensively used by professional and nonprofessional stakeholders and investors (e.g., Abrahamson & Amir, 1996; Bartlett & Chandler, 1997). In particular, it has been shown that MD&A narratives offer a useful new source of information to sell-side financial analysts and institutes (Clarkson et al., 1999) and that high quality MD&A narratives help reduce errors and dispersion in financial analysts' earnings forecasts (Barron, Kile, & O'Keefe, 1999).

Given the importance of narrative disclosures, accounting and management scholars have extensively examined their content (e.g., Rutherford, 2002), functions (e.g., Merkl-Davies & Brennan, 2007), and readability (e.g., Moreno & Casasola, 2016), as well as their relationship to the reporting company's financial performance (e.g., Dutta et al., 2019). This body of research has also revealed significant effects of several essential company parameters, including company size (in terms of revenue) and profitability (in terms of profit percent change), on the choice of disclosure items, the amount of information disclosed for each item, and the ways in which items are discussed (e.g., Malafronte, Porzio, & Starita, 2016; Robb, Single, & Zarzeski, 2001; Tarca, Street, & Aerts, 2011; Zimmerman, 2013).

Research on the linguistic features of ARs and narrative disclosures has been conducted mostly from the perspectives of genre and discourse analysis. In line with the construal of genre as structured communicative events motivated by specific communicative purposes and performed by particular discourse communities (Bhatia, 2008), Swales (1990) argued that the AR demonstrates "a shared set of communicative purposes as the principal feature that turns a collection of communicative events into a genre" (pp. 45–47). In an analysis of corporate AR narratives as a genre of accounting narrative, Rutherford (2005) used word frequencies to identify genre rules and reported that the genre tended to utilize "language biased toward the positive", despite the fact that it was expected to "be expressed in neutral terms" (p. 349). Vogel (2010) analyzed the narrative sections of ARs from large companies in different industries from a genre perspective and reported that lexis and syntax were strategically chosen and arranged to achieve self-promotion. Cho and Yoon (2013) performed a comparative genre analysis of corporate earnings calls produced by Korean and first language (L1) English speakers in terms of structural

and linguistic specificity. They reported that Koreans were not adequately aware of the genre expectations of earnings calls and argued for the need to promote their awareness of professional genre specificity in Business English pedagogy.

Research from the discourse analysis perspective views ARs as an example “of how language is used to mediate between specialized and institutionalized forms of communication and more ‘popular’ texts” (Piotti, 2014, p. 13). From this perspective, multiple studies have examined the communicative functions of different linguistic features in ARs and AR narratives (e.g., Garzone, 2004; Malavasi, 2010; Piotti, 2014; Thomas, 1997). For example, Thomas (1997) examined how a combination of verb structures, thematic structures, and cohesion and condensation features were used in the management’s messages in the ARs of a particular company to construct a factual context for the company’s losses. Garzone (2004) identified several differences in the use of first-person pronouns in the ARs of companies in Italy and other countries and attributed those differences to a combination of factors, such as cultural differences in expressing company-investor relationships. Piotti (2014) proposed a communicative account of the use of hedging in ARs and offered a systematic analysis of how hedging devices were used to “modalize” or “modulate” claims in this form of corporate communication (p. 14).

Alongside the increasing popularity of the use of corpus linguistic techniques in analyzing the linguistic features of ESP texts (e.g., Hartig & Lu, 2014), most genre and discourse analysis studies of AR narratives employed corpus linguistic analysis methods in various ways. However, several noteworthy gaps exist in this body of literature. First, there is scant focus on the linguistic features of narrative disclosures of ARs of Chinese companies. As an increasing number of Chinese companies are going public in the U.S. and other international stock markets, it is critical for Chinese public companies to understand how their disclosing practices compare with those of international public companies. Second, while some comparative studies have investigated linguistic variation in the AR narratives of companies of different countries (e.g., Cho & Yoon, 2013; Garzone, 2004), no study has examined linguistic variation by parameters considered to be essential in the profiles of public companies, such as company size and profitability. Given previous findings on the effects of these parameters on the types, amount, and presentation of information in narrative disclosures, an analysis of their potential effects on the linguistic features of narrative disclosures would enrich our understanding of inter-company differences in their linguistic practices in engaging with this professional genre. Third, given the prominence of the genre and discourse analysis perspectives and corpus analytic methods in studies of the linguistic features of AR narratives, MD analysis appears to lend itself well to this line of research. Somewhat surprisingly, however, little extant research in this area has taken advantage of this approach.

1.3. Multi-dimensional analysis

MD analysis was initially proposed by Biber (1988) as a method for examining register variation, based on the theoretical claim that “registers are best described with respect to patterns of linguistic co-occurrence” (Biber, Gray, & Staples, 2016, p. 646). Several comprehensive introductions to MD analysis exist (e.g., Biber, 1988, 1995, 2006). In general, the analysis is performed on a large corpus of naturally occurring texts representing different registers in a discourse domain and proceeds as follows (Biber et al., 2016). First, the distribution of a range of linguistic features in each text in the corpus is analyzed (e.g., in terms of the normalized frequency of occurrence of each feature in each text). The features analyzed are determined either theoretically or empirically to have the potential to characterize or distinguish different registers. Second, a factor analysis is conducted to identify co-occurrence patterns, or dimensions, among those features. Finally, the dimensions are interpreted in terms of their “underlying functional associations” (Biber et al., 2016, p. 646) and texts of different registers are compared along each dimension.

Over the past three decades, MD analysis has developed into the “dominant approach” to the study of register variation (Gardner, Nesi, & Biber, 2019, p. 647). Numerous studies have applied MD analysis to examine the characteristics of and variation among texts of different registers and language varieties in general and specific domains, such as student and adult speech (Reppen, 2001), outsourced call center interactions (Friginal, 2009), blogs (Grieve, Biber, Friginal, & Nekrasova, 2010), newspaper editorials (Huang & Ren, 2020), world Englishes (Xiao, 2009), and legal English (Goźdz-Roszkowski, 2011).

A large body of EAP research has employed MD analysis to examine the linguistic features of texts across and within academic registers, genres, and disciplines. Some MD studies systematically explored variation between spoken and written registers within university language at large (e.g., Biber, 2006), while others looked at variation among different academic disciplines and/or genres (e.g., Gray, 2015) or among different units of a specific disciplinary genre (e.g., Biber & Jones, 2005). In addition to systematically outlining the linguistic characteristics of academic and disciplinary texts, this line of research has also served several pedagogical goals. To better understand learner development or performance and improve second language (L2) assessment, some MD studies examined genre and discipline variation in student writing (e.g., Hardy & Friginal, 2016; Hardy & Römer, 2013) and linguistic variation across learning time, language backgrounds, assessment scores, proficiency levels, exam task types, or writing topics (e.g., Biber et al., 2016; Friginal & Weigle, 2014; Weigle & Friginal, 2015). To inform L2 pedagogy, some MD studies assessed the appropriateness of potential teaching materials for specific pedagogical contexts (e.g., Al-Surmi, 2012) or examined whether EAP instruction resulted in significant improvement in students’ EAP essays (e.g., Crosthwaite, 2016). In a recent study with considerable pedagogical implications, Gardner et al. (2019) identified four dimensions from a corpus of university assignments and mapped academic disciplines, disciplinary groups, levels of study, and genre families onto those dimensions. The authors argued that, by relating linguistic features to situational perspectives on student academic writing, their findings could be used to “inform the teaching of a ‘common academic core’, and to inform the design of programmes tailored to the needs of specific disciplines” (p. 646).

While these MD studies in general, specific, and academic domains differ from the current study in terms of analytical foci and goals, they convincingly suggest that MD analysis can be fruitfully applied to uncover the linguistic variation in MD&A narratives in Chinese and American corporate ARs and that findings on such variation will have useful pedagogical implications for Business English pedagogy. In addition, the proven ability of MD analysis to relate linguistic features to situational perspectives (e.g., Gardner et al., 2019) also aligns well with the prevailing view that discursive and professional practices should be integrated and that it is necessary to go from text to context for a comprehensive and critical view of discursive practices (Bhatia, 2008).

1.4. The current study

In light of the research gaps discussed above, the current study applies MD analysis to examine the linguistic variation in MD&A narratives in Chinese and American corporate ARs, with the specific goal to address the following two research questions:

1. How do the MD&A narratives in Chinese and American corporate ARs vary in terms of their linguistic features?
2. How do company size and profitability relate to the distribution of the linguistic features in the MD&A narratives in Chinese corporate ARs and in American corporate ARs, respectively?

2. Materials and methods

2.1. Corpus

Our data consisted of a corpus of MD&A narratives in the ARs of Chinese and American public companies. The corpus compilation process involved three steps. In the first step, we selected 57 Chinese and 57 American public companies¹ (see the Appendices) representing different company sizes and profitability levels from the 2015 *Fortune Global 500* list, which was published in www.fortune.com and www.fortunechina.com. For each company, we recorded its name, rank on the list, revenue, profit, and profit percent change.

In the second step, we retrieved the English MD&A narratives published in 2013 and 2014 by the 57 Chinese companies from www.cninfo.com.cn and www.hkex.com.hk and those by the 57 American companies from www.annualreports.com and www.sec.gov. The Chinese companies variously included “Management’s Report”, “Board of Directors’ Report”, “Report of Directors”, or “Operating and Financial Review and Prospects” in the English versions of their ARs. Upon close review of the content of these variably named reports and the disclosing requirements from the China Securities Regulatory Commission (CSRC), we determined that they all served the same function and should thus all be included under MD&A.

In the final step, we converted each PDF document to a plain text file and cleaned each file by correcting any textual oddities that resulted from the conversion process and removing any non-textual elements that would not be included in the subsequent linguistic analysis, such as images, graphs, and meta-data information. As summarized in Table 1, the final corpus contained 114 MD&A narratives from Chinese companies (hereafter the Chinese subcorpus), with a total of 992,777 tokens, and 114 MD&A narratives from American companies (hereafter the American subcorpus), with 2,353,408 tokens. We return to a description of the subcorpora used to examine linguistic variation by company size and profitability in Section 3.2.

On average, the MD&A narratives in the Chinese subcorpus were substantially shorter than those in the American subcorpus. A close look at the Chinese and English versions of the MD&A narratives in the ARs of Chinese companies revealed that the English version was translated from the Chinese version in most cases. The translations remained faithful to the narratives in Chinese in both content and language. The quality of the translations appeared very high. This was expected, given the importance of the document and the fact that most of these companies have international business operations and work with international accounting firms to produce their ARs. A further comparison of the English MD&A narratives in the two subcorpora showed that the difference in length resulted primarily from the tendency of American companies to include more narrative information than Chinese companies, particularly information on financial projections and forecasts. This difference may have partially arisen from the different regulations and provisions of the SEC and CSRC. For example, the Securities and Exchange Commission (2017) provides explicit “safe harbor provisions” that “protect forward-looking statements against certain private legal actions alleging material misstatements or omissions” (p. 307). The absence of comparable provisions from the CSRC may explain why financial projections and forecasts were more prevalent and elaborate in the MD&A narratives of American companies.

¹ Altogether, 106 Chinese public companies were included in the 2015 *Fortune Global 500* list, among which 57 released ARs in English. To maximize comparability of the Chinese and American sub-corpora, we selected 57 American public companies from the list as well.

Table 1
Descriptive statistics of the Chinese and American subcorpora.

Subcorpus	Chinese	American
Number of texts	114	114
Number of word types	16,093	21,294
Number of word tokens	992,777	2,353,408
Mean length of text	8,864	21,013

2.2. Analysis

The MD&A narratives in the two subcorpora were analyzed using Nini's (2015) Multidimensional Analysis Tagger (MAT), which replicates Biber's (1988) MD analysis of register variation across speech and writing. The MAT first generates a grammatical analysis of each text using the Stanford Tagger (Toutanova, Klein, Manning, & Singer, 2003) and then identifies the linguistic features used in Biber (1988) from each grammatically tagged text. The frequency counts of all linguistic features are normalized to occurrences per 100 words to account for differences in text length. These features are then used to generate scores for the six functional dimensions presented in Biber's (1988) analysis for each text. Table 2 summarizes the six dimensions from Biber (1988) used in the program.

Table 2
A summary of Biber's (1988) dimensions, adapted from Nini (2015, pp. 5–6).

Dimension	Description
D1	<i>Involved vs. Informational Discourse</i> : A high score indicates the text is affective and interactional, e.g., a casual conversation; the text presents many verbs and pronouns (among other features). A low score indicates the text is informationally dense, e.g., academic prose; the text presents many nouns, long words, and adjectives (among other features).
D2	<i>Narrative vs. Non-Narrative Concerns</i> : A high score indicates the text is narrative, e.g., a novel. The text presents many past tenses and third person pronouns (among other features). A low score indicates the text is non-narrative.
D3	<i>Context-Independent vs. Context-Dependent Discourse</i> : A high score indicates the text is context-independent, e.g., academic prose; the text presents many nominalizations (among other features). A low score indicates the text is context-dependent, e.g., a sport broadcast; the text presents many adverbs (among other features).
D4	<i>Overt Expression of Persuasion</i> : A high score indicates that the text explicitly marks the author's point of view and their assessment of likelihood and/or certainty, e.g., a professional letter; the text presents many modal verbs (among other features).
D5	<i>Abstract vs. Non-Abstract Information</i> : A high score indicates the text is highly technical, abstract, or formal, e.g., scientific discourse; the text presents many passive clauses and conjuncts (among other features).
D6	<i>On-line Informational Elaboration</i> : A high score indicates the text is informational in nature but produced under certain time constraints, e.g., a speech. The text presents many postmodifications of noun phrases (among other features).

The program produces three output files containing the following information both for each input file and for the corpus as a whole: the normalized frequency per 100 tokens for each linguistic feature, the z-score of each feature, and the score for each of Biber's (1988) six dimensions. The program also produces two sets of graphs that allow for the determination of the genres that the corpus is closest to on each dimension and the text type it is closest to overall, respectively, among the genres and text types Biber (1988) analyzed.

We imported the output files containing relevant statistical information from the MAT to SPSS 25.0 for further analysis. To address the first research question, we performed a series of independent-samples *t* tests to determine whether and how the MD&A narratives in the Chinese and American subcorpora vary along the six functional dimensions. To address the second research question, we identified the companies in the top and bottom quarters of each subcorpus in terms of company size and profitability. We then ran a series of independent-samples *t* tests to determine whether these variables led to significant variation in the MD&A narratives along the six dimensions. Because 42 *t* tests were run in the analysis, we adjusted the study-wide alpha value to .05/42 (i.e., .001).

3. Results

3.1. Research question 1: linguistic variation in the MD&A narratives in the Chinese and American subcorpora

Table 3 summarizes the dimension scores for the Chinese and American subcorpora, along with the genre that they are closest to on each dimension and the text type that they are closest to overall. The two subcorpora show scores of the same polarity (i.e., negative or positive)² for all six dimensions, and they are closest to the same genre for five dimensions. Both

² Positive and negative scores on the same dimension represent loadings on the two different ends of the dimension. For example, a positive D1 score represents a loading on the interactive end, while a negative D1 score represents a loading on the informational end.

subcorpora are closest to learned exposition, a type of text characterized by low scores on D1 (informationally dense) and high scores on D3 (more context-independent) and D5 (more technical, abstract, or formal), with such characterizing genres as official documents and academic prose (Nini, 2015). Texts of this type are “typically informational expositions that are formal and focused on conveying information” (Nini, 2015, p. 7). Overall, the MD&A narratives in the two subcorpora demonstrate a high level of consistency in terms of the distribution of the linguistic features along the six dimensions, indicating that they largely conform to the same set of genre expectations. Meanwhile, the American subcorpus has substantially smaller standard deviations for all dimension values, with the exception of D1. These results suggest that the MD&A narratives in the American subcorpus may be linguistically more homogeneous than those in the Chinese subcorpus.

Table 3

Mean dimension scores and closest genres and text type of the two subcorpora.

Dimension	American subcorpus			Chinese subcorpus		
	Mean	SD	Closest genre	Mean	SD	Closest genre
D1	−20.59	2.79	Official documents	−21.00	2.77	Official documents
D2	−4.52	.40	Broadcasts	−2.28	1.89	Academic prose
D3	10.63	1.51	Official documents	10.46	2.39	Official documents
D4	−3.75	.98	Broadcasts	−5.33	1.21	Broadcasts
D5	2.17	.89	Press reportage	.82	1.25	Press reportage
D6	−1.74	.40	General fiction	−2.83	.47	General fiction
Closest text type			Learned exposition			Learned exposition

Table 4 summarizes the results of the independent-samples *t* tests for the dimension scores of the two subcorpora. The results indicate that the MD&A narratives in the two subcorpora differ significantly in their scores for D2, D4, D5 and D6, but not for D1 and D3. We examine the results for each dimension in detail below.

Table 4

Independent-samples *t* test results for the dimension scores of the two subcorpora.

Dimension	American		Chinese		<i>t</i>	df	<i>p</i>
	Mean	SD	Mean	SD			
D1	−20.59	2.79	−21.00	2.77	1.11	226	.267
D2	−4.52	.40	−2.28	1.89	−12.42	226	<.001
D3	10.63	1.51	10.46	2.39	.63	226	.530
D4	−3.75	.98	−5.33	1.21	10.84	226	<.001
D5	2.17	.89	.82	1.25	9.38	226	<.001
D6	−1.74	.40	−2.83	.47	18.75	226	<.001

3.1.1. Dimension 1: involved vs. informational discourse

Low D1 scores mean that the text or corpus is informationally dense (Biber, 1988). The D1 scores for the American and Chinese subcorpora are both very low (−20.59 vs. −21.00), with no significant difference between them ($t = 1.11$, $p = .267$). Both subcorpora are closest to official documents on this dimension. These results indicate that the MD&A narratives in the two subcorpora are both informational in nature, with a comparable degree of informational density.

The key linguistic features characterizing low D1 scores include nouns, long words, and attributive adjectives (Biber, 1988). An examination of the *z*-scores of these features shows that the two subcorpora have comparable average word length (2.38 vs. 2.27), but the American subcorpus uses more attributive adjectives (1.71 vs. 1.09) and fewer nouns (2.63 vs. 3.08) than the Chinese subcorpus. Biber (1988) argued that a higher frequency of attributive adjectives can be associated with a more “careful integration of information in a text” (p. 104) (e.g., *unrecognized tax benefits* vs. *tax benefits*). Meanwhile, a close look at the Chinese subcorpus indicates that its higher noun frequency may be partially related to a lower frequency of first-person pronouns. In particular, Chinese companies commonly use *the Company/Group* instead of *we* for self-reference, suggesting a tendency to appear less interactive and involved (Biber, 1988) than American companies. Examples 1³ and 2 illustrate the frequent use of *the Company/Group* for self-reference in the Chinese subcorpus and the higher use of attributive adjectives in the American subcorpus.

- Ex. 1. **The Group** will develop reliably and orderly to explore the wind power resources along with the publication of the policy of wind power development supported by the State and the improvement of the technologies of the offshore wind power development. **The Group's offshore** wind power projects totaling 300.0 MW have been listed in the approved plans of the National Energy Administration. (China Huadian, 2014)
- Ex. 2. **We** have made acquisitions that included a significant amount of goodwill and other intangible assets. When **we** complete an acquisition, **we** apply the acquisition method of accounting, which among other things, requires the recognition of goodwill (which represents the excess cost of the

³ All excerpts are quoted verbatim and may contain some expressions that are non-native-like.

acquisition over the fair value of net assets acquired and identified intangible assets). (AETNA, 2014)

3.1.2. Dimension 2: narrative vs. non-narrative concerns

Low D2 scores mean that the text or corpus is non-narrative (Biber, 1988). The D2 scores for the American and Chinese subcorpora are both low (−4.52 vs. −2.28). However, an independent-samples *t* test reveals a statistically significant difference between the two scores ($t = -12.42, p < .001$), indicating that the American subcorpus is further along the non-narrative pole, to the extent that the MD&A narratives in the American and Chinese subcorpora are closest to two different genres on D2 (broadcasts vs. academic prose).

Low D2 scores are characterized by low usage of third-person pronouns and past tense verbs (Biber, 1988). An examination of the z-scores of these features reveals that the use of third-person pronouns is comparably low in the American and Chinese subcorpora (−.128 vs. −.13), but the former uses substantially fewer past tense verbs than the latter (−.7 vs. .43). A close examination of the MD&A narratives in the two subcorpora revealed that Chinese companies tended to place more emphasis on past performance than American companies, resulting in higher usage of past tense verbs, as illustrated in Example 3. Meanwhile, American companies tended to discuss ongoing efforts and future prospects more than Chinese companies, resulting in greater use of other verb forms, such as present tense verbs, non-finite forms, and modal verbs, as illustrated in Example 4.

- Ex. 3. The increase **was** mainly attributed to the increased allocation of resources in negotiated deposits, debt investment schemes of good credit rating and high quality, asset management products and asset securities products. It also **benefited** from gains in the domestic capital market which **contributed** to a significant increase in investment income. (People's Insurance Company of China, 2014)
- Ex. 4. Part of this strategic initiative **has** been to leverage our scale to **improve** the value we **provide** to customers and shareholders and **achieve** \$1 billion in annual efficiencies, up to \$400 million of which **will be** reinvested in technology, platforms and functionality to **improve** our current operations and develop new capabilities. We also **continue** to **balance** our product mix between protection products and more capital-intensive products in order to **maintain** predictable operating earnings and cash flows. (METLIFE, 2014)

3.1.3. Dimension 3: context-independent vs. context-dependent discourse

High D3 scores indicate that the text or corpus does not depend on context for interpretation (Biber, 1988). The D3 scores for the American and Chinese subcorpora are both very high (10.63 vs. 10.46), with no significant difference between them ($t = .63, p = .530$). Both subcorpora are closest to official documents on this dimension. These results indicate that the MD&A narratives in the two subcorpora are equally context-independent.

A key linguistic feature characterizing high D3 scores is nominalizations (Biber, 1988). The z-scores of this feature are comparably high for the American and Chinese corpora (3.11 vs. 3.22). Nominalizations (e.g., *realization, reduction, resolution, expectation*) package information that could otherwise be expressed in a clause in a less context-dependent way, as they allow for greater flexibility to omit context-dependent information such as the agent, time, and location. As illustrated in Examples 5 and 6, the high frequency of nominalizations in both corpora is accompanied with low frequencies of temporal (e.g., *now*) and locative (e.g., *here*) adverbials, in line with Biber's (1988) discussion of the characteristic features of D3.

- Ex. 5. Capitalizing on the **opportunities** arising from the new trend of **urbanization**, the **Transportation** Finance SBU actively explored business **opportunities** in the **construction** of integrated **transportation** hub and **transportation** resource **integration** through direct financing, **transaction** financing and structured financing to promote its **transformation** towards **investment** banking. (China Minsheng Bank, 2013)
- Ex. 6. We currently expect to spend approximately \$300 million for capital **expenditures** during fiscal 2014. Several of the larger 2014 capital **expenditures** include the **replacement** of a **distribution** center, the **creation** of a national **distribution** center, the **implementation** of track-and-trace **authentication** technology, and **information** system **investments** to support increased order volume and future **growth**. (AmerisourceBergen, 2013)

3.1.4. Dimension 4: overt expression of persuasion

Low D4 scores mean that the text or corpus does not explicitly mark the authors' point of view and their assessment of likelihood and/or certainty (Biber, 1988). The D4 scores for the American and Chinese subcorpora were both negative (−3.75 vs. −5.33), and both are closest to broadcasts on D4, indicating that the authors' point of view is not overtly marked in either subcorpus. However, an independent-samples *t* test reveals a statistically significant difference between the two scores ($t = 12.84, p < .001$), indicating that the Chinese subcorpus is significantly less overt than the American subcorpus in presenting the authors' views.

A key linguistic feature characterizing low D4 scores is the limited use of modal verbs. Biber (1988) categorized modal verbs as possibility modals, such as *can* and *may*, necessity modals, such as *should* and *must*, and predictive modals, such as *will* and *shall*. Table 5 summarizes the z-scores of the different types of modals in the two subcorpora. These results show that the American subcorpus uses substantially more modals than the Chinese subcorpus, particularly possibility modals. The

American companies' more prominent use of modal verbs in the MD&A narratives aligns with their greater emphasis on prospects than Chinese companies, as illustrated in Examples 7 and 8.

- Ex. 7. Financial Statements regarding the realization of certain net deferred tax assets, the possibility that tax audit resolutions over the twelve months ending January 31, 2015, **could** reduce unrecognized tax benefits by an amount within a certain range or beyond that range and the reasons for that reduction, the expectation that any change **will** not have a significant impact on the Company's Consolidated Financial Statements and the possibility that the resolution of a group of related matters **might** result in a material liability to Walmart. (Wal-Mart Stores, 2014)
- Ex. 8. These core strengths in talent and culture have enabled Lenovo to flexibly serve global, large enterprise customers as well as small-to-medium businesses and individual consumers across product lines, and these are the same strengths that **will** enable us to expand and deepen our customer reach through the Group's clear strategy. Additionally, Lenovo's "global-local" model of business leadership and talent allowed the Group to take advantage of far-reaching industry trends, while at the same time leveraging the unique strengths of local leadership to drive success in key markets. (Lenovo, 2014)

Table 5

Z-scores of modal verbs in the American and Chinese subcorpora.

Subcorpus	Necessity modals	Possibility modals	Predictive modals	All modals
American	-.81	-.61	-.76	-3.75
Chinese	-.91	-1.39	-.6	-5.33

3.1.5. Dimension 5: abstract vs. non-abstract information

High D5 scores mean that the text or corpus conveys information in a technical, abstract, and formal way (Biber, 1988). The D4 scores for the American and Chinese subcorpora are both positive (2.17 vs. .82), and both are closest to press reportage on D5, indicating that the MD&A narratives in both subcorpora are technical, abstract and formal. An independent-samples *t* test reveals a statistically significant difference between the two scores ($t = 9.38, p < .001$), indicating that the American subcorpus is significantly more abstract than the Chinese subcorpus.

Two key linguistic features characterizing high D5 scores are conjuncts and passive clauses, including agentless passives and by-passives (Biber, 1988). Table 6 summarizes the z-scores of these features in the two subcorpora. These results show that the American subcorpus uses more agentless passives (e.g., *The resolution was revoked*), by-passives (e.g., *The resolution was passed by the board*), and conjuncts (e.g., *therefore*) than the Chinese corpus. A close examination of the two subcorpora also suggests that the American subcorpus more frequently uses conjuncts together with passive clauses to indicate "the complex logical relations among clauses" that characterize abstract, formal discourse (Biber, 1988, p. 112), as illustrated in Examples 9 and 10. The lower frequencies of conjuncts and passive clauses in the Chinese subcorpus may have partially resulted from L1 influence. Specifically, Chinese has been found to frequently drop conjuncts and rely on context to reflect the logical relationships between clauses (Tse, 2010); it has also been found to use the passive voice less frequently than English (Xiao, McEnergy, & Qian, 2006).

- Ex. 9. **Further**, a bill passed by the Japanese legislature in 2012 will increase the consumption tax rate during 2014 and 2015. Insurance commissions are subject to consumption tax for individuals exceeding certain earnings thresholds; **however**, the tax is not charged on employee compensation or insurance premiums. The increase in this tax is expected to lead to increased costs for insurers. (Prudential Financial, 2013)
- Ex. 10. The four high-end office buildings we hold or operate in Beijing and Shanghai are all landmarks in core administration area or central commercial district; **therefore**, they have attracted many well-known brand names to become their tenants. In 2013, the company seized opportunities, maintained high occupancy rates, and the rental rate continued to increase. (SINOCHEN, 2013)

Table 6

Z-scores of passive clauses and conjuncts in the two subcorpora.

Subcorpus	Agentless passives	By-passives	Conjuncts
American	-.04	1.14	1.03
Chinese	-.57	-.09	.46

3.1.6. Dimension 6: on-line informational elaboration

Low D6 scores mean that the text or corpus is not produced on-line or under severe time constraints (Biber, 1988). The D6 scores for the American and Chinese subcorpora are both low (-1.74 vs. -2.83), and both are closest to general fiction. An independent-samples *t* test reveals a statistically significant difference between the two scores ($t = 18.75, p < .001$), indicating that the Chinese subcorpus contains significantly fewer features of texts produced on-line.

Low D6 scores are characterized by limited usage of demonstratives and *that*-clauses, including *that* complements to verbs (e.g., *We anticipate that ...*) or adjectives (e.g., *We are confident that ...*) and *that* relative clauses in object positions (e.g., *the situation that we are in*) (Biber, 1988). Table 7 summarizes the z-scores of these features in the two subcorpora. These results show that the Chinese subcorpus uses fewer demonstratives and *that*-clauses than the American subcorpus. These differences suggest that, although MD&A narratives represent a formal, planned type of discourse, American companies tend to employ relatively more demonstratives and subordinate features than Chinese companies to achieve cohesion, to unpack information that could otherwise be presented in a more dense manner with phrasal features, or to elaborate information relative to the authors' stance (Biber, 1988), as illustrated in Examples 11 and 12.

- Ex. 11. **This** ASU requires that government guaranteed residential real estate mortgage loans that meet specific criteria be recognized as other receivables upon foreclosure; previously, **these** assets were included in foreclosed assets. Government guaranteed residential real estate mortgage loans that completed foreclosure during 2014 and met the criteria specified by ASU 2014. (Wells Fargo, 2014)
- Ex. 12. **This** ensured that Dongfeng Motor could complete the registration of its subsidiary in the Shanghai FTZ, and obtain approval for investment and other relevant matters as scheduled. In March 2014, the Bank assisted Dongfeng Motor in finishing the share acquisition project, thereby building the Bank's brand as the premier bank of Shanghai FTZ related business. (Bank of China, 2014)

Table 7Z-scores of *that* clauses and demonstratives in the two subcorpora.

Subcorpus	<i>That</i> clause as verb complements	<i>That</i> clause on object positions	<i>That</i> clause as adjectival complements	Demonstratives
American	-.83	-.22	-.29	-.91
Chinese	-1.03	-.56	-.45	-1.81

3.2. Research question 2: linguistic variation by company size and profitability

This section reports results on whether, within each of the two subcorpora, the MD&A narratives varied by company size and profitability.

3.2.1. Company size

Following the practice of the *Fortune* magazine, we operationalized company size using total revenue. American companies generally ranked higher on the 2015 *Fortune Global 500* list and have larger revenue than Chinese companies. However, this difference should not affect our analysis, as our goal is to examine whether there is variation by company size within each subcorpus. Specifically, in each subcorpus, we selected the top and bottom 14 companies in terms of total revenue for comparison. As shown in Table 8, the MD&A narratives in the American subcorpus do not vary significantly by company size on any dimension and those in the Chinese subcorpus differ significantly on D4 ($t = 4.60, p < .001$). These results show that larger Chinese companies are even less overt in expressing the authors' views than smaller ones.

Table 8Independent-samples *t* test results for the dimension scores for the top and bottom 14 largest companies in the two subcorpora.

Dimension	Subcorpus	Top 14		Bottom 14		<i>t</i>	df	<i>p</i>
		Mean	SD	Mean	SD			
D1	Chinese	-20.21	3.42	-20.83	2.61	-.76	50.49	.450
	American	-20.25	2.22	-21.47	2.55	-1.44	44.42	.149
D2	Chinese	-2.57	1.64	-1.81	1.43	1.85	52.96	.071
	American	-4.69	.43	-4.59	.57	-.27	43.68	.599
D3	Chinese	10.02	2.27	8.92	1.90	-1.97	52.40	.054
	American	10.07	.99	10.97	1.19	-.70	44.71	.082
D4	Chinese	-5.98	.98	-4.61	1.23	4.60	51.51	<.001
	American	-3.67	.48	-3.68	.37	-1.75	53.34	.957
D5	Chinese	.95	1.02	.91	1.40	-1.48	49.35	.883
	American	1.82	.56	1.94	1.22	1.80	52.03	.771
D6	Chinese	-2.83	.46	-2.89	.61	-.44	50.29	.661
	American	-1.51	1.09	-1.84	.74	.44	47.24	.198

3.2.2. Profitability

The 2015 *Fortune Global 500* list includes information on the profit percent change from 2013 to 2014 for each company, which we used in the current study to operationalize profitability. In each subcorpus, we selected the top and bottom 14 companies in terms of profitability for comparison. As shown in Table 9, the MD&A narratives do not vary significantly by profitability on any dimension in either subcorpus.

Table 9Independent-samples *t* test results for the dimension scores for the top and bottom 14 most profitable companies in the two subcorpora.

Dimension	Subcorpus	Top 14		Bottom 14		<i>t</i>	df	<i>p</i>
		Mean	SD	Mean	SD			
D1	Chinese	-20.35	2.86	-20.21	2.92	.13	25.99	.897
	American	-19.92	3.45	-20.18	2.38	-.23	23.11	.823
D2	Chinese	-2.09	1.83	-2.01	1.68	.14	25.79	.909
	American	-4.67	.44	-4.47	.32	1.38	23.94	.181
D3	Chinese	10.01	1.58	9.89	1.68	-.20	25.90	.847
	American	10.59	.75	10.25	1.08	-.95	23.14	.351

Table 9 (continued)

Dimension	Subcorpus	Top 14		Bottom 14		t	df	p
		Mean	SD	Mean	SD			
D4	Chinese	−4.98	1.36	−5.30	1.53	−.60	25.65	.557
	American	−3.71	.82	−3.78	.89	−.22	25.84	.830
D5	Chinese	.74	1.03	.82	1.10	.20	25.89	.842
	American	2.13	.91	2.13	.49	.02	19.92	.982
D6	Chinese	−2.91	.31	−2.63	.59	1.56	19.49	.134
	American	−1.85	.39	−1.80	.41	.36	25.94	.723

4. Discussion

The results of our comparative MD analysis of the MD&A narratives in Chinese and American corporate ARs revealed several substantive findings regarding the linguistic features and variation of MD&A narratives produced by Chinese and American companies. The scores of the MD&A narratives on the six functional dimensions derived from Biber's (1988) seminal MD analysis of register variation indicated the following linguistic features of MD&A narratives: they are informationally dense (low D1 scores), non-narrative (low D2 scores), contextually independent (high D3 scores), non-explicit in expressing the author's point of view (low D4 scores), technical, abstract, and formal (high D5 scores), and well planned (low D6 scores). Overall, the MD&A narratives are closest to learned exposition in terms of their linguistic features. The consistent polarity of the dimension scores of the MD&A narratives across the two subcorpora (i.e., negative for D1, D2, D4, D6 and positive for D3 and D5) constitutes the first piece of evidence that MD&A narratives have fairly stable and consistent genre features.

With respect to the first research question, our findings show that the MD&A narratives produced by Chinese and American companies differed significantly in their scores for four dimensions. First, the Chinese subcorpus was significantly more narrative than the American subcorpus, as indicated by the significantly higher D2 score. The different D2 scores resulted from a stronger tendency for Chinese companies to place greater emphasis on narrating past events (e.g., *It benefited from ...*) than commenting on ongoing efforts (e.g., *We continue to ...*) or future prospects (e.g., *We will reinvest ...*).

Second, the Chinese subcorpus was significantly less explicit in expressing the author's point of view, as indicated by the significantly lower D4 score. While additional research is needed to identify factors that contribute to the tendency of Chinese companies to disclose information in a less persuasive manner, this tendency may reflect a decreased awareness of or attention to the promotional and persuasive functions of the MD&A narratives (e.g., Malavasi, 2010). Meanwhile, in the U.S., the Securities Act Regulation section 230.175 (rule 175) “provides a ‘safe harbor’ for voluntary disclosure of a ‘forward-looking’ statement” (Till, 1980, p. 608), which protects management from liability for making financial projections in good faith. This protection has likely allowed American companies to place greater emphasis on prospects, resulting in more prominent use of modal verbs in the MD&A narratives. The *Contents and Formats of Annual Reports* (2012 Revision) published by the China Securities Regulatory Commission (2012) contains six principles that highlight the requirement for Chinese companies to disclose decision-related information, such as information concerning major investment projects, mergers and acquisitions, and research and development programs, but lacks comparable provisions for financial projections and forecasts. While Chinese companies also included statements about future plans, they were less inclined to predict future outcomes, resulting in less prominent use of modal verbs in the MD&A narratives.

Third, the Chinese subcorpus was significantly less technical, abstract, and formal than the American subcorpus, as indicated by the significantly lower D5 score. This result hints at both content and style differences between the two subcorpora: the higher degree of abstractness of the American subcorpus appears to align with its lower degree of narrativeness, and, as mentioned earlier, its more technical and formal style is characterized by its more widespread use of passive clauses and conjuncts.

Fourth, the Chinese subcorpus also showed significantly fewer features of on-line informational elaboration, as indicated by the significantly lower D6 scores. Less on-line information elaboration often results from more shared background knowledge (Biber, 1988). The low D6 scores for both subcorpora are thus well justified, as the MD&A narratives are intended for professional and semi-professional readers, who are expected to have basic background knowledge of the reporting companies. The lower D6 scores of the Chinese subcorpus suggest that Chinese companies may have assumed greater background knowledge than American companies. Meanwhile, Biber (1988) pointed out that features characterizing high D6 scores, such as demonstratives and *that* clauses, have the “secondary use” for “the expression of opinions, attitudes, or personal statements of individuals” (p. 160). Thus, the higher D6 score of the Chinese subcorpus may also be related to its lower D4 score, which marks a lower degree of explicitness in expressing the authors' views.

Regarding our second research question, our findings show that the dimension scores of the MD&A narratives did not vary by company size and profitability in the American subcorpus; however, in the Chinese corpus, some variation was observed by company size. Specifically, larger Chinese companies were more implicit in expressing the authors' views (lower D4 score) than smaller ones. While previous studies have not examined the relationship between these variables and the linguistic features of narrative disclosures, some have reported a positive association between company size and level of disclosures (Wallace & Naser, 1995) and positive (Owusu-Ansah & Yeoh, 2005) or negative associations (Wallace & Naser, 1995) between

profitability and the extent of voluntary disclosures. The finding that the linguistic features of the MD&A narratives in the American subcorpus did not vary significantly in relation to these variables can be interpreted as a second piece of evidence that MD&A narratives have fairly stable and consistent genre features. The finding on the linguistic variation by company size within the Chinese subcorpus echoes the larger standard deviations observed for its dimension scores and paints a less stable and consistent picture of the Chinese companies' engagement with the genre features of MD&A narratives. Somewhat surprisingly, with the exception of D3, on all other dimensions that company size exhibited significant effects, larger Chinese companies were more distant from their American counterparts than smaller Chinese companies.

The findings of the current study have useful pedagogical and practical implications in Business English (BE), a branch of ESP that has become increasingly important in China. As more Chinese companies go public in international stock markets, knowledge of different genres of BE texts is increasingly emphasized in BE programs, and knowledge of the genre features of narrative disclosures in ARs in particular will become an important component of the professional literacy of finance and accounting management personnel. Our findings on the linguistic features and variation of MD&A narratives in the Chinese and American ARs can inform BE teachers and students as well as business professionals of the expected features of the genre of MD&A narratives, the linguistic differences between the MD&A narratives produced by Chinese and American companies, and the linguistic variation among MD&A narratives produced by Chinese companies of different sizes. By conducting the comparative analysis, we did not assume a deficiency approach to interpreting the differences and variation observed. Nevertheless, these findings can be used in different ways in teaching and learning the genre features of MD&A narratives, depending on the pedagogical perspective and needs of the BE program.

The grammatically tagged version of the corpus of MD&A narratives produced in this study can be used to inform the design of pedagogical materials in advanced BE courses. As McEnergy, Xiao, and Tono (2006) noted, MD analysis systematically links linguistic features and their co-occurrence patterns to the communicative functions they enact. Teachers can be trained to identify occurrences of salient linguistic features for each of the six functional dimensions in grammatically tagged texts and to select relevant texts to illustrate how the co-occurrence of those features achieves the different communicative functions characterizing MD&A narratives as a professional genre. They may design genre analysis activities for students to analyze the uses of relevant linguistic features in carefully selected texts and the communicative functions they are used to realize. They can also design comparative analysis activities for students to examine differences in the use of relevant linguistic features across different texts, the impact of those different uses on the functions the texts communicate, and the possible sources of those differences. Such activities can promote learners' awareness of the genre features of MD&A narratives as well as the types of linguistic differences and variation among MD&A narratives produced by L1 and L2 English writers (e.g., Dong & Lu, 2020).

5. Conclusion

This study has investigated the linguistic features and variation of the MD&A narratives in ARs of Chinese and American public companies on the 2015 *Fortune Global 500* list. Our findings have shed useful light on the genre features of MD&A narratives, on the linguistic differences between MD&A narratives produced by Chinese and American companies, and on the linguistic variation by company size among the MD&A narratives produced by Chinese companies. Our findings have useful implications for BE pedagogy and practice in the context of China. Our analysis also demonstrated the applicability and usefulness of MD analysis as a research methodology for revealing linguistic variation in BE as ESP research.

This study has several limitations, some of which constitute useful avenues for future research. First, the corpus compiled can be expanded in multiple ways to enrich the research scope. For example, the scale of the corpus can be increased by including more Chinese and American companies to verify our findings. Data from different time points can be collected to track changes in the genre features of MD&A narratives produced by Chinese and American companies. Data from companies from countries with other L1 backgrounds can be collected to examine linguistic variation of MD&A narratives across a larger repertoire of L1 backgrounds. Second, some differences may have arisen from L1 influence in the translation process (e.g., the lower frequency of the conjuncts and passive clauses in the Chinese subcorpus), the English language competence of the translators or writers (e.g., the lower frequency of the various types of *that*-clauses in the Chinese subcorpus), and differences in corporate and national culture (e.g., Garzone, 2004). Our study design did not allow us to draw conclusions on the effects of these factors, and future research could systematically explore their impact. Finally, the relationship of the genre features of MD&A narratives to other important aspects of MD&A narratives, such as the level and quality of disclosure, would be highly useful to investigate.

Acknowledgements

We thank the editors and anonymous reviewers for their constructive comments and suggestions. This research was funded by two grants (No. GD15XGL13 and No. GD16XGL21) from the Guangdong Provincial Fund for the Philosophy and Social Science Planning Project to the first author.

Appendices

Appendix A. 2015 Fortune 500 companies in the Chinese subcorpus

No.	Rank in 2015	Rank in 2014	Company Name	Revenues		Annual Revenues Change		Profit		Annual Profit Change		Ticker
				(\$M)	%			(\$M)	%			
1	2	3	SINOPEC GROUP	446,811.00	–2.30			5,177.00	–42.00			SINOPEC
2	4	4	CHINA NATIONAL PETROLEUM	428,620.00	–0.80			16,359.50	–11.60			PETROCHINA
3	18	25	INDUSTRIAL & COMMERCIAL BANK OF CHINA	163,174.90	9.70			44,763.90	4.80			ICBC
4	29	38	CHINA CONSTRUCTION BANK	139,932.50	11.60			36,976.60	5.90			CCB
5	36	47	AGRICULTURAL BANK OF CHINA	130,047.70	12.70			29,126.40	7.70			ABC
6	37	52	CHINA STATE CONSTRUCTION ENGINEERING	129,887.10	17.20			2,079.10	12.20			CSC
7	45	59	BANK OF CHINA	120,946.00	14.50			27,525.10	7.90			BOC
8	55	55	CHINA MOBILE COMMUNICATIONS	107,529.40	–0.10			10,451.20	13.60			CM
9	60	85	SAIC MOTOR	102,248.60	11.10			4,540.10	12.50			SAIC
10	71	–	CHINA RAILWAY ENGINEERING	99,537.90	9.10			959.90	9.20			CR
11	72	79	CHINA NATIONAL OFFSHORE OIL	99,262.20	3.40			8,592.60	11.60			CNOOC
12	79	80	CHINA RAILWAY CONSTRUCTION	96,395.20	0.70			1,154.20	17.00			CRCC
13	94	98	CHINA LIFE INSURANCE	87,249.30	7.80			1,687.50	183.70			CL
14	96	128	PING AN INSURANCE	86,021.80	25.60			6,374.90	39.20			PINGAN
15	105	107	SINOCHEM GROUP	80,635.00	6.20			562.70	–25.50			SINOCHEM
16	109	113	DONGFENG MOTOR GROUP	78,978.60	6.70			1,600.40	10.50			DFM
17	160	154	CHINA TELECOMMUNICATIONS	62,147.60	0.20			2,037.70	31.00			CT
18	165	187	CHINA COMMUNICATIONS CONSTRUCTION	60,119.20				1,467.10				CCC
19	174	208	PEOPLE'S INSURANCE COMPANY OF CHINA	57,047.50	15.10			2,127.60	61.10			PICC
20	186	160	CITIC GROUP	55,325.70	–9.30			4,715.00	–23.40			CITIC
21	190	217	BANK OF COMMUNICATIONS	54,464.20	12.70			10,687.40	5.50			BOCO
22	196	165	SHENHUA GROUP	52,731.10	–11.90			4,376.00	–26.30			SH
23	198	133	CHINA MINMETALS	52,383.10	–22.30			–374.10	–193.10			CMM
24	207	248	BEIJING AUTOMOTIVE GROUP	50,566.00	16.70			819.90	–10.70			BAIC
25	218	211	BAOSTEEL GROUP	48,323.40	–2.00			952.90	3.00			BSIS
26	224	221	CHINA HUANENG GROUP	47,401.40	–0.60			423.90	–0.60			HN
27	231	286	LENOVO GROUP	46,295.60	19.60			828.70	1.40			LENOVO
28	235	350	CHINA MERCHANTS BANK	45,613.80	33.70			9,074.30	7.80			CMB
29	240	227	ALUMINUM CORP. OF CHINA	45,445.00	NA			–1,758.20	NA			CHALCO
30	253	313	POWER CHINA	43,009.70	16.90			1,071.80	28.70			CHINAPOWER
31	270	267	CHINA NATIONAL BUILDING MATERIALS GROUP	40,644.40	–0.90			477.80	9.80			CNBM
32	271	338	INDUSTRIAL BANK	40,594.70	16.60			7,650.50	14.10			IB

(continued on next page)

(continued)

No.	Rank in 2015	Rank in 2014	Company Name	Revenues		Profit	Annual Profit Change		Ticker
				(\$M)	%		(\$M)	%	
33	272	401	COFCO	40,524.50	31.80	123.70	197.30		COFCO
34	276	357	SINOPHARM	40,105.70	20.50	439.00	27.00		SINOPHARM
35	281	330	CHINA MINSHENG BANKING	39,921.90	12.40	7,229.80	5.10		CMBC
36	288	278	SINOMACH	39,722.50	0.80	−288.90	−215.80		CMEC
37	296	383	SHANGHAI PUDONG DEVELOPMENT BANK	38,683.80	22.80	7,632.30	14.70		SPDB
38	326	354	CHINA METALLURGICAL GROUP	35,807.50	6.30	280.60	NA		MCC
39	328	384	CHINA PACIFIC INSURANCE GROUP	35,669.80	13.60	1,793.20	19.10		CPIC
40	343	297	CHINA GUODIAN	34,627.40	−8.50	488.50	−14.50		GDKJ
41	345	368	CHINA HUADIAN	34,487.70	6.00	1,080.70	35.00		HDGJ
42	354	381	JIANGXI COPPER	33,778.20	6.80	52.80	−43.10		JXC
43	362	366	GUANGZHOU AUTOMOBILE INDUSTRY GROUP	33,237.40	1.40	284.20	36.70		GAC
44	366	382	CHINA ELECTRONICS	33,084.90		228.70			CEC
45	380	432	SHAANXI YANCHANG PETROLEUM (GROUP)	31,391.00	11.50	1,085.30	−31.30		YCPETRO
46	390	398	CHINA NONFERROUS METAL MINING (GROUP)	30,456.30		−12.30			CNMG
47	391	465	CHINA ENERGY ENGINEERING GROUP	30,322.10	17.70	389.00	51.70		CEEC
48	392	396	CHINA DATANG	30,206.90	−2.50	11.70	−94.10		DTP
49	402	348	SHOUGANG GROUP	29,668.90	−13.50	3.30	NA		SGGJ
50	403	393	CHINA POWER INVESTMENT	29,584.70	−4.80	234.00	−49.80		CHINAPOWER
51	420	–	CHINA EVERBRIGHT GROUP	28,155.30	19.70	1,475.10	392.10		GDKG
52	432	451	CHINA OCEAN SHIPPING	27,483.00	2.50	541.60	NA		COSCO
53	437	–	CHINA AEROSPACE SCIENCE & TECHNOLOGY	27,190.40	17.50	1,431.60	5.60		CA
54	457	–	CHINA POLY GROUP	26,046.60	21.50	1,020.50	−1.40		POLY
55	464	–	HNA GROUP	25,646.40	36.40	206.90	26.40		HNAI
56	477	466	ZHEJIANG GEELY HOLDING GROUP	24,986.40	−3.00	275.50	133.20		GEELY
57	500	310	WUHAN IRON & STEEL	23,720.9	−35.8	54.5	347.2		WIS

Appendix B. 2015 Fortune 500 companies in the American subcorpus

No.	Rank in 2015	Rank in 2014	Company Name	Revenues		Profit	Annual Profit Change		Ticker
				(\$M)	%		(\$M)	%	
1	1	1	WAL-MART STORES	485,651	2.00	16,363	2.10		WMT
2	5	5	EXXON MOBIL	382,597	−6.10	32,520	−0.20		XOM
3	12	12	CHEVRON	203,784	−7.50	19,241	−10.20		CVX
4	14	14	BERKSHIRE HATHAWAY	194,673	6.90	19,872	2.00		BRK-A
5	15	15	APPLE	182,795	7.00	39,510	6.70		AAPL
6	16	29	MCKESSON	181,241	31.30	1,476	16.90		MCK
7	21	21	GENERAL MOTORS	155,929	0.30	3,949	−26.10		GM
8	23	19	PHILLIPS 66	149,434	−7.30	4,762	27.80		PSX
9	24	27	GENERAL ELECTRIC	148,321	1.40	15,233	16.70		GE
10	27	26	FORD MOTOR)	144,077	−1.90	3,187	−55.50		FORM
11	30	35	CVS HEALTH	139,367	9.90	4,644	1.10		CVS

(continued)

No.	Rank in 2015	Rank in 2014	Company Name	Revenues		Annual Revenues Change		Profit		Annual Profit Change		Ticker
				(\$M)	%			(\$M)	%			
12	33	34	AT&T	132,447	2.90			6,224	–65.90			T
13	34	30	VALERO ENERGY	130,844	–5.00			3,630	33.50			VALERO
14	35	39	UNITEDHEALTH GROUP	130,474	6.50			5,619	–0.10			UNHEX
15	41	42	VERIZON COMMUNICATIONS	127,079	5.40			9,625	–16.30			VZ
16	46	88	AMERISOURCEBERGEN	119,569.10	34.10			276.5	–36.30			ABC
17	50	37	FANNIE MAE	116,461	–7.30			14,208	–83.10			FNM
18	52	60	COSTCO WHOLESALE	112,640	7.10			2,058	0.90			CSTC
19	53	50	HEWLETT-PACKARD	111,454	–0.80			5,013	–2.00			HP
20	54	74	KROGER	108,465	10.30			1,728	13.80			KR
21	61	57	J.P. MORGAN CHASE & CO.	102,102	–3.90			21,762	21.40			JPM
22	66	62	EXPRESS SCRIPTS HOLDING	100,887.10	–3.60			2,007.60	8.80			ESI
23	80	66	BANK OF AMERICA CORP.	95,181	–6.40			4,833	–57.70			BAC
24	82	71	INTERNATIONAL BUSINESS MACHINES	94,128	–5.60			12,022	–27.10			IBM
25	83	81	MARATHON PETROLEUM	91,417	–2.70			2,524	19.50			MPC
26	84	67	CARDINAL HEALTH	91,084	–9.90			1,166	249.10			CAH
27	85	90	BOEING	90,762	4.80			5,446	18.80			BOEING
28	86	82	CITIGROUP	90,646	–3.20			7,313	–46.50			CITI
29	88	112	AMAZON.COM	88,988	19.50			–241	–188.00			AMZN
30	90	89	WELLS FARGO	88,372	0.30			23,057	5.40			WFC
31	95	104	MICROSOFT	86,833	11.50			22,074	1.00			MSFT
32	100	92	PROCTER & GAMBLE	84,537	0.40			11,643	2.90			PG
33	101	102	HOME DEPOT	83,176	5.50			6,345	17.80			HD
34	104	87	ARCHER DANIELS MIDLAND	81,201	–9.60			2,248	67.50			ADM
35	114	117	WALGREENS BOOTS ALLIANCE	76,392	5.80			1,932	–21.10			WBA
36	117	116	TARGET	74,520	2.70			–1,636	–183.00			TGT
37	118	121	JOHNSON & JOHNSON	74,331.00	4.20			16,323.00	18.00			JNJ
38	120	120	ANTHEM	73,874.10	3.40			2,569.70	3.20			ANTM
39	121	131	METLIFE	73,316	7.50			6,309	87.30			MET
40	124	162	GOOGLE	71,487	17.90			14,444	11.80			GOOG
41	135	146	COMCAST	68,775	6.40			8,380	22.90			CMCSA
42	149	151	UNITED TECHNOLOGIES	65,100	3.40			6,220	8.70			UTC
43	152	127	AMERICAN INTERNATIONAL GROUP	64,406	–6.20			7,529	–17.10			AIG
44	168	182	UNITED PARCEL SERVICE	58,232	5.00			3,032	–30.60			UPS
45	169	174	DOW CHEMICAL	58,167	1.90			3,772	–21.20			DOW
46	170	223	AETNA	58,003.20	22.60			2,040.80	6.60			AET
47	172	153	BUNGE	57,161	–8.60			515	68.30			BG
48	176	192	LOWE'S	56,223	5.30			2,698	18.00			LOW
49	180	167	CONOCOPHILLIPS	55,997	–5.80			6,869	–25.00			COP
50	182	195	INTEL	55,870	6.00			11,704	21.70			INTC
51	183	213	ENERGY TRANSFER EQUITY	55,691	14.20			633	223.00			ETE
52	187	181	CATERPILLAR	55,184	–0.80			3,695	–2.50			CAT
53	194	264	PRUDENTIAL FINANCIAL	54,123	30.50			1,381	NA			PRU
54	211	191	PFIZER	49,605	–7.80			9,135	–58.50			PFE
55	214	232	WALT DISNEY	48,813	8.40			7,501	22.20			DIS
56	215	228	SCHLUMBERGER	48,580	7.10			5,438	–19.20			SLB
57	216	266	HUMANA	48,500.0	17.4			1,147.0	–6.8			HUM

References

- Abrahamson, E., & Amir, E. (1996). The information content of the president's letter to shareholders. *Journal of Business Finance & Accounting*, 23(8), 1157–1182.
- Al-Surmi, M. (2012). Authenticity and TV shows: A multidimensional analysis perspective. *Tesol Quarterly*, 46(4), 671–694.
- Barron, O. E., Kile, C. O., & O'Keefe, T. B. (1999). MD&A quality as measured by the SEC and analysts' earnings forecasts. *Contemporary Accounting Research*, 16(1), 75–109.
- Bartlett, S. A., & Chandler, R. A. (1997). The corporate report and the private shareholder: Lee and Tweedie twenty years on. *The British Accounting Review*, 29(3), 245–261.
- Beattie, V. A., McInnes, W. M., & Fearnley, S. (2004). *Through the eyes of management: Narrative reporting across three sectors*. London: Centre for Business Performance.
- Bhatia, V. K. (2008). Genre analysis, ESP and professional practice. *English for Specific Purposes*, 27(2), 161–174.
- Biber, D. (1988). *Variation across speech and writing*. Cambridge: Cambridge University Press.
- Biber, D. (1995). *Dimensions of register variation: A cross-linguistic comparison*. Cambridge: Cambridge University Press.
- Biber, D. (2006). *University language: A corpus-based study of spoken and written registers*. Amsterdam: John Benjamins.

- Biber, D., Gray, B., & Staples, S. (2016). Predicting patterns of grammatical complexity across language exam task types and proficiency levels. *Applied Linguistics*, 37(5), 639–668.
- Biber, D., & Jones, J. K. (2005). Merging corpus linguistic and discourse analytic research goals: Discourse units in biology research articles. *Corpus Linguistics and Linguistic Theory*, 1(2), 151–182.
- China Securities Regulatory Commission. (2012). *The standards concerning the content and formats of information disclosure by companies offering securities to the public No.2 - Contents and formats of annual reports (2012 revision)*. Retrieved from http://www.csrc.gov.cn/pub/csrc_en/laws/overRule/Announcement/201210/t20121025_216124.html. (Accessed 1 March 2019).
- Cho, H., & Yoon, H. (2013). A corpus-assisted comparative genre analysis of corporate earnings calls between Korean and native-English speakers. *English for Specific Purposes*, 32(3), 170–185.
- Clarkson, P. M., Kao, J. L., & Richardson, G. D. (1999). Evidence that management discussion and analysis (MD&A) is a part of a firm's overall disclosure package. *Contemporary Accounting Research*, 16(1), 111–134.
- Crosthwaite, P. (2016). A longitudinal multidimensional analysis of EAP writing: Determining EAP course effectiveness. *Journal of English for Academic Purposes*, 22, 166–178.
- Dong, J., & Lu, X. (2020). Promoting discipline-specific genre competence with corpus-based genre analysis activities. *English for Specific Purposes*, 58, 138–154.
- Dutta, S., Fuksa, M., & Macaulay, K. (2019). Determinants of MD&A sentiment in Canada. *International Review of Economics & Finance*, 60, 130–148.
- Friginal, E. (2009). *The language of outsourced call centers: A corpus-based study of cross-cultural interaction*. Amsterdam: John Benjamins.
- Friginal, E., & Weigle, S. (2014). Exploring multiple profiles of L2 writing using multi-dimensional analysis. *Journal of Second Language Writing*, 26(2), 80–95.
- Gardner, S., Nesi, H., & Biber, D. (2019). Discipline, level, genre: Integrating situational perspectives in a new MD analysis of university student writing. *Applied Linguistics*, 40(4), 646–674.
- Garzone, G. (2004). Annual company reports and CEO's letters: Discoursal features and cultural markedness. In C. N. Candlin, & M. Gotti (Eds.), *Intercultural aspects of specialized communication* (pp. 311–341). Frankfurt: Peter Lang.
- Gates, B. (1996). What I learned from Warren Buffett. *Harvard Business Review*, 74(1), 148–150.
- Goźdz-Roszkowski, S. (2011). *Patterns of linguistic variation in American legal English*. Frankfurt: Peter Lang.
- Gray, B. E. (2015). *Linguistic variation in research articles: When discipline tells only part of the story*. Amsterdam: John Benjamins.
- Grieve, J., Biber, D., Friginal, E., & Nekrasova, T. (2010). Variation among blogs: A multi-dimensional analysis. In A. Mehler, S. Sharoff, & M. Santini (Eds.), *Genres on the web: Corpus studies and computational models* (pp. 45–71). New York: Springer-Verlag.
- Hajek, P., & Henriques, R. (2017). Mining corporate annual reports for intelligent detection of financial statement fraud - A comparative study of machine learning methods. *Knowledge-Based Systems*, 128, 139–152.
- Hardy, J., & Friginal, E. (2016). Genre variation in student writing: A multi-dimensional analysis. *Journal of English for Academic Purposes*, 22, 119–131.
- Hardy, J., & Römer, U. (2013). Revealing disciplinary variation in student writing: A multi-dimensional analysis of the Michigan Corpus of Upper-level Student Papers (MICUSP). *Corpora*, 8(2), 183–207.
- Hartig, A. J., & Lu, X. (2014). Plain English and legal writing: Comparing expert and novice writers. *English for Specific Purposes*, 33, 87–96.
- Huang, Y., & Ren, W. (2020). A novel multi-dimensional analysis of writing styles of editorials from China Daily and the New York Times. *Lingua*, 235, 1–16.
- Li, F. (2010). The information content of forward-looking statements in corporate filings - A naïve Bayesian machine learning approach. *Journal of Accounting Research*, 48(5), 1049–1102.
- Loughran, T., & McDonald, B. (2016). Textual analysis in accounting and finance: A survey. *Journal of Accounting Research*, 54(4), 1187–1230.
- Malafrente, I., Porzio, C., & Starita, M. G. (2016). The nature and determinants of disclosure practices in the insurance industry: Evidence from European insurers. *International Review of Financial Analysis*, 45, 367–382.
- Malavasi, D. (2010). The multifaceted nature of banks' annual reports as informative, promotional and corporate communication practices. In P. E. Allori, & G. Garzone (Eds.), *Discourse, identities and genres in corporate communication: Sponsorship, advertising and organizational communication* (pp. 211–234). Bern: Peter Lang.
- McEnery, T., Xiao, R., & Tono, Y. (2006). *Corpus-based language studies: An advanced resource book*. London: Routledge.
- Merkl-Davies, D. M., & Brennan, N. M. (2007). Discretionary disclosure strategies in corporate narratives: Incremental information or impression management? *Journal of Accounting Literature*, 27, 116–196.
- Moreno, A., & Casasola, A. (2016). A readability evolution of narratives in annual reports: A longitudinal study of two Spanish companies. *Journal of Business and Technical Communication*, 30(2), 202–235.
- Nini, A. (2015). *Multidimensional analysis tagger (version 1.3) - Manual*. Retrieved from <http://sites.google.com/site/multidimensionaltagger>. (Accessed 1 May 2019).
- Owusu-Ansah, S., & Yeoh, J. (2005). The effect of legislation on corporate disclosure practices. *Abacus*, 41(1), 92–109.
- Piotti, S. (2014). *Exploring corporate rhetoric in English: Hedging in company annual reports. A corpus-assisted analysis*. Milan, Italy: EDUCatt.
- Reppen, R. (2001). Register variation in student and adult speech. In S. Conrad, & D. Biber (Eds.), *Variation in English: Multi-dimensional studies* (pp. 187–199). London: Longman.
- Robb, S. W. G., Single, L. E., & Zarzeski, M. T. (2001). Nonfinancial disclosures across Anglo-American countries. *Journal of International Accounting*, 10(1), 71–83.
- Rogers, R. K., & Grant, J. (1997). Content analysis of information cited in reports of sell-side financial analysts. *Journal of Financial Statement Analysis*, 3, 17–31.
- Rowbottom, N., & Lymer, A. (2010). Exploring the use and users of narrative reporting in the online annual report. *Journal of Applied Accounting Research*, 11(2), 90–108.
- Rutherford, B. A. (2002). The operating and financial review: Could do better, could say more. *Accounting and Business*, 5(3), 40–41.
- Rutherford, B. A. (2005). Genre analysis of corporate annual report narratives: A corpus linguistics-based approach. *Journal of Business Communication*, 42(4), 349–378.
- Securities and Exchange Commission. (2017). *Financial reporting manual*. Retrieved from <https://www.sec.gov/corpfin/cf-manual>. (Accessed 23 March 2019).
- Swales, J. (1990). *Genre analysis: English in academic and research settings*. Cambridge: Cambridge University Press.
- Tarca, A., Street, D. L., & Aerts, W. (2011). Factors affecting MD&A disclosures by SEC registrants: Views of practitioners. *Journal of International Accounting, Auditing and Taxation*, 20(1), 45–59.
- Thomas, J. (1997). Discourse in the marketplace: The making of meaning in annual reports. *Journal of Business Communication*, 34(1), 47–66.
- Till, K. (1980). SEC safe harbor for forecasts - A step in the right direction? *Duke Law Journal*, 29(3), 607–640.
- Toutanova, K., Klein, D., Manning, C., & Singer, Y. (2003). Feature-rich part-of-speech tagging with a cyclic dependency network. In *Proceedings of the 2003 conference of the North American chapter of the association for computational linguistics on human language technology* (pp. 252–259). East Stroudsburg, PA: Association for Computational Linguistics.
- Tse, Y.-K. (2010). Parataxis and hypotaxis in the Chinese language. *International Journal of Arts and Sciences*, 3(16), 351–359.
- Vogel, R. (2010). Company annual reports: Analysis of the genre with regard to communication strategies. *Discourse and Interaction*, 3(1), 87–111.
- Wallace, R. O., & Naser, K. (1995). Firm-specific determinants of the comprehensiveness of mandatory disclosure in the corporate annual reports of firms listed on the stock exchange of Hong Kong. *Journal of Accounting and Public Policy*, 14(4), 311–368.
- Weigle, S. C., & Friginal, E. (2015). Linguistic dimensions of impromptu test essays compared with successful student disciplinary writing: Effects of language background, topic, and L2 proficiency. *Journal of English for Academic Purposes*, 18, 25–39.
- Xiao, R. (2009). Multi-dimensional analysis and the study of world Englishes. *World Englishes*, 28(4), 421–450.
- Xiao, R., McEnery, T., & Qian, Y. (2006). Passive constructions in English and Chinese: A corpus-based contrastive study. *Languages in Contrast*, 6(1), 109–149.
- Zimmerman, J. L. (2013). Myth: External financial reporting quality has a first-order effect on firm value. *Accounting Horizons*, 27(4), 887–894.

Chaowang Ren is an assistant professor in the School of Foreign Languages at Guangdong University of Technology. He completed his Ph.D. (2017) in Business English Studies at Guangdong University of Foreign Studies. His research interests include corpus linguistics, Business English and discourse analysis.

Xiaofei Lu is Professor of Applied Linguistics and Asian Studies at The Pennsylvania State University. His research interests are primarily in corpus linguistics, English for Academic Purposes, second language writing, and intelligent computer-assisted language learning. He is the author of *Computational Methods for Corpus Annotation and Analysis* (2014, Springer).