SciencePG
Science Publishing Group

A Multidimensional Analysis of Letters to the Shareholders by Chinese and American Advanced Equipment Manufacturing Companies

Liao Shunzhu

English Department, School of Foreign Languages, Guangdong University of Technology, Guangzhou, China

Email address:
3074665545@qq.com

Abstract: Initially proposed by Biber, the multidimensional analysis is used to examine the linguistic variation among different registers, shedding light on the linguistic features and differences among different genres. Through the tool of MAT, this study compared the register features in the letters to the shareholders by Chinese and American advanced equipment manufacturing companies (AEMCs). It is found that: 1) the letters to the shareholders by Chinese and American AEMCs are both informational, non-narrative, contextually independent, non-overt in expressing views, abstract and well-planned, belonging to the text type of learned exposition; 2) In comparison, the letters to the shareholders by Chinese AEMCs are relatively more informational, narrative and abstract, but less dependent on context and lower in on-line informational elaboration while letters by American AEMCs show more features of interaction with the readers, focusing on the present actions, context-dependence and on-line informational elaboration; 3) The key linguistic features that account for these differences include nouns, nominalizations, first-person pronouns, present tense verbs, past tense verbs, phrasal coordination, phrasal coordination, prediction modals, conjuncts, passives, demonstratives and that-clauses. 4) The future Return on Equity (ROE) of the Chinese AEMCs is related to the dimensions of "explicit/context-dependent references" and "on-line informational elaboration". 5) In comparison to the previous study, industry is found to have an effect on the linguistic variation of letters to the shareholders. This study gave a comprehensive view of the linguistic features and differences in the letters to the shareholders by Chinese and American AEMCs and had some pedagogical implications for business English teaching and learning.

Keywords: Multi-dimensional Analysis, Linguistic Variation, Letters to the Shareholders, Chinese and American Equipment Manufacturing Companies

1. Introduction

The letter to the shareholders, also called the CEO's letter or the Chairman's statement, is considered the most widely read section in the annual report [1]. As a promotional genre, it does not only report the market environment, the company's financial and operational performance, future strategies and goals of the company [2], but also intends to create a credible and confident corporate image [3].

Given its importance, the letter to the shareholders has been studied from various aspects including genre [4], rhetorical features [3], linguistic features [5], and the relationship between language usage and company performance [6, 7]. Comparative studies between Chinese and American letters to the shareholders are also attracting attention. For example, Wang and Li [8] studied the intertextuality in the Chinese and American letters to the shareholders and found that the amount of performance and forward-looking information in the Chinese letters correlated negatively with the Return on Equity (ROE) and the earnings per share (EPS) while the frequency of reporting verbs in English letters correlated negatively with EPS. Huang Ying's study [9] on the Chinese and western banks showed that Western letters to the shareholders resorted to more credible and affective appeals by using more interactional meta-discourse markers while Chinese letters to the

shareholders made more logical appeals by using more interactive meta-discourse markers, indicating a lack of register awareness of the Chinese writers of English.

It can be seen from these studies that language in Chinese and American letters to the shareholders differs in various ways and the correlation between linguistic features and company performance varies in cultures. However, these studies only focus on some of the linguistic features and very few studies have been given to the overall register features and advanced equipment manufacturing industry. Therefore, this study aims to further their research by exploring the register features with a multidimensional (MD) analysis on the letters to the shareholders by Chinese and American advanced equipment manufacturing companies (AEMCs), with the purpose to reveal the key linguistic features and see whether industry and company's future performance relate to these differences.

2. Multidimensional Analysis

Initiated by Douglas Biber [10] to examine the linguistic variation across spoken and written English, MD analysis is based on the theoretical claim that "registers are best described with respect to patterns of linguistic co-occurrence" [11]. Register variation must be studied from "sets of co-occurring features" rather than some "individual markers" [10] (p21). MD analysis starts by analyzing the distribution of a number of linguistic features and then a factor analysis is applied to identify sets of co-occurrence patterns. Each set of co-occurring linguistic features in every factor constitutes a dimension which is interpreted based on their "underlying functional associations" [11]. Finally, register variation is described by comparing each dimension. The whole process is assisted with the corpus-based methodology to annotate the texts and analyze the statistics. MD analysis is considered a practical way that shifts the researchers' attention from "a particular language feature" [12] to the notion of "language variation as a continuum" [12], providing a quantitative and qualitative method for the "holistic" [13] analysis of the complexity of texts.

Over the last three decades, MD analysis has shed light on the complexity of linguistic variation across different registers and domains, such as learners' discourse [12], blogs [14], academic discourse [15], translation [16] and contrastive study [17]. It has provided pedagogical implications for language teaching and learning. For instance, Aguado-Jiménez et al.'s study [12] has looked into the feasibility of carrying out MD analysis of learner language and found that students given register awareness-raising training were better prepared to understand register and the connections between individual linguistic features. Gardner et al. related the situational variables with clusters of linguistic features and identified four dimensions that can be selected to "inform the teaching of a 'common academic core'" [18].

The application of MD analysis in business English study has been few but has started to attract some attention. Jiang & Xu's study [19] showed that MD analysis can be used to differentiate business English from general English, economic news from general news, and economic academic discourse from general academic discourse. Wang & Bu [20] compared the Chinese and American corporate annual report narratives through MD analysis and found that Chinese and American corporate narratives differed significantly in four dimensions. They pointed out that the findings were meaningful in that they may help the Chinese companies improve the annual report narratives and suggested that future research could focus on the relation between register features and company profitability. Hu & Tan [21] compared 120 letters to the shareholders by Chinese and American enterprises and discussed the cultural and linguistic impact on the linguistic variation. Ren & Lu's recent study [22] applied MD analysis to compare register features of the Management's Discussion and Analysis narratives in Chinese and American corporate annual reports and investigated the relation of company size and profitability with the register features. These studies have given insights into the language in business English, annual reports in particular, providing important implications for the teaching and writing of business English. However, most of them were written in Chinese, making it difficult for international researchers to gain inspiration from them. Moreover, the linguistic variation in the letters to the shareholders by Chinese and American companies in a specific industry has not yet been explored and the relation between register features and company profitability has not yet been discussed. Therefore, this study attempts to address this gap by laying our emphasis on the industry of advanced equipment manufacturing companies and relation between future profitability and register features.

This study employs Biber's 1988 MD analysis [10] of 67 linguistic features and 6 dimensions. These six dimensions are: D1 "involved vs informational production", D2 "narrative vs non-narrative concerns", D3 "explicit vs situation-dependent reference", D4 "overt expression of persuasion", D5 "abstract vs non-abstract information" and D6 "online informational elaboration". Each dimension is determined by a set of linguistic features, some with positive loadings and others with negative loadings. Take D2 for example. The features with positive loadings, such as past tense verbs, third-person pronouns and perfect aspect verbs, suggest that the text is narrative while the features with negative loadings, such as present tense verbs and attributive adjectives, indicate that the text is non-narrative.

3. Research Design

3.1. Research Questions

This study intends to examine the linguistic variation in the letters to the shareholders by the Chinese and American AEMCs and addresses the following two research questions:

1) In what way do the letters to the shareholders in the ARs of the Chinese and American AEMCs differ in terms of the six functional dimensions?

2) How does future profitability relate to the register features in the letters to the shareholders by the Chinese and American AEMCs respectively?

3.2. Data

Our data consisted of 300 letters to the shareholders between 2014 and 2016 from annual reports (ARs) of 50 Chinese AEMCs and 50 American AEMCs. The letters are all written in English. The Chinese ARs were collected from www.hkex.com.hk or the respective company official websites, and the American ARs were collected from www. annualreports.com or the respective company official websites. The letters to the shareholders were retrieved from the ARs and converted into text files.

Through cleaning the noises of the text, such as graphs, images, meta-data, misspelling, mispunctuation, etc, we compiled two corpora, namely the Corpus of Chinese AEMCs (CAEMC) and the Corpus of American AEMCs (AAEMC). To examine whether future profitability is related to linguistic features of the letter to the shareholders, we first retrieved the financial data of Return on Equity (ROE) from Wind database and then compiled four subcorpora based on the comparison to the next year's ROE, that is, we compared the result of the previous year's ROE with that of the next year to see whether the future profitability grew or declined. For example, we compared the 2014 ROE with the 2015 ROE. Those companies that grew in ROE are included in the Growth subcorpus, while those that declined in ROE in the Decline subcorpus. Table 1 gives an overview of the two corpora and four subcorpora complied.

Table 1. Overview of the corpora and subcorpora.

Corpora /subcorpora	Number of texts	Tokens	Types	STTR
CAEMC	150	181,300	7,596	41.90
Chinese Growth Subcorpus	76	89,974	5,692	63.26
Chinese Decline Subcorpus	74	91,326	5,754	63.01
AAEMC	150	202,784	9,942	49.03
American Growth Subcorpus	75	96,856	6,940	71.65
American Decline Subcorpus	75	105,928	7,495	70.76

3.3. Methods

This study uses Nini's Multidimensional Analyzer Tagger (MAT) [23] to analyze the linguistic features of the corpora and subcorpora. MAT is a program that replicates Biber's taggers for the MD analysis of English texts. Test on LOB and the Brown corpus has shown that MAT can successfully replicate Biber's analysis [23].

The program automatically tags the grammatical features of each text by using the embedded "Standford POS Tagger" and then calculates the frequency of the 67 linguistic features used in Biber [10]. The frequency counts are standardized to occurrences per 100 tokens. The z-scores of each linguistic variation are calculated based on means and standard deviations presented in Biber [10]. Finally, the dimension scores are determined by using the mean z-scores of variables higher than 1 and the closest genre is given for each

dimension based on the dimension score. The dimension score can be positive or negative, indicating the loadings on different extremes of the dimension. For example, a positive D2 score indicates a loading on a narrative end, while a negative D2 represents a loading on a non-narrative end.

Three spreadsheets are generated by the program. They contain statistical information for each text and the corpus as a whole, namely, the standardized frequency per 100 tokens for each linguistic feature, the z-score of each feature, the score for each dimension. Seven graphs are also produced to show the text type the corpus is closest to on each dimension.

This study first ran the program on the two corpora and four subcorpora, and then imported the statistical information of dimension score and z-score of the 67 linguistic features to IBM SPSS 26.0. Independent-sample t tests were run to determine whether CAEMC and AAEMC, Chinese Growth subcorpus and Chinese Decline subcorpus, American Growth subcorpus and American Decline subcorpus vary across the six dimensions and what linguistic features account for the differences. PowerConc 1.0 [24] is used to study the context of some salient linguistic features.

4. Results and Discussion

4.1. Register Features of Letters to the Shareholders in CAEMC and AAEMC

Table 2 shows the mean dimension scores, closet text types and independent-sample t test results of the two corpora. Both corpora show scores loaded on the same end, either negative or positive, for all six dimensions. It can be concluded from the table that letters to the shareholders by Chinese and American AEMC are both informational, non-narrative and abstract, less dependent on context, weak in overt persuasion and well-planned. In terms of the genres they are closest to, they are the same in five dimensions and both belong to the text type of learned exposition in general. Learned exposition is a text type that includes official documents, press reviews and academic press. It is characterized by low scores on D1 and high scores on D3 and D5, featuring texts that are highly informational and formal [23]. This finding confirms Hu & Tan's previous study [21] on letters to the shareholders by Chinese and American Companies, indicating that letters to the shareholders present fairly stable and consistent genre features regardless of the industry sectors.

The independent-sample t test results indicate that the two corpora differ significantly in D1, D2, D3 and D6 ($p<0.001$). In D5, a significance at $p<0.05$ is also identified. It indicates that the letters to the shareholders by Chinese AEMCs are relatively more informational, narrative and formal, but less dependent on context and lower in on-line informational elaboration while letters to the shareholders by American AEMCs show more features of texts that are interactive, non-narrative, context-dependent, non-abstract and loosely planned. This result differs from Hu & Tan's findings [21] in which significance at $p<0.001$ is detected in D1, D3 and D6,

but not in D2 (p=0.018). It indicates that the industry sector may be a factor to affect the linguistic variation.

Table 2. *Mean dimension scores, closest text types and independent-sample t test results of CAEMC and AAEMC.*

Dimension	CAEMC			AAEMC			t	df	p	MD
	Mean	SD	Closest genre	Mean	SD	Closest genre				
D1	-23.77	2.69	Official documents	-17.95	5.25	Official documents	-12.10	222.04	0.000	-5.82
D2	-1.70	2.05	Academic prose	-3.09	1.31	Official documents	6.97	253.18	0.000	1.38
D3	12.28	3.63	Official documents	9.43	3.21	Official documents	7.19	298.00	0.000	2.85
D4	-2.13	2.22	Press reportage	-2.10	2.41	Press reportage	-0.11	298.00	0.916	-0.03
D5	0.87	2.07	Press reportage	0.36	1.91	Press reportage	2.22	298.00	0.027	0.51
D6	-2.61	0.70	General fiction	-1.50	0.94	General fiction	-11.66	275.85	0.000	-1.12
Closest text type			Learned exposition			Learned exposition				

4.2. Linguistic Features That Account for the Differences

Independent-sample *t* tests of the z-scores of the 67 linguistic features in the two corpora show that 51 linguistic features differ significantly (p<0.05), accounting for 76% of the total. Key linguistic features of each dimension are examined in detail below.

4.2.1. D1: Involved vs Informational Production

The D1 scores for the Chinese and American corpora are negative and low (-23.77 vs -17.95)[1], with a statistical significance at p<0.001, indicating that letters to the shareholders in both corpora are informationally dense but letters to the shareholders by American AEMCs are relatively more interactional and affective.

The key linguistic features that have a positive weight on information density include nouns, prepositions, attributive adjectives, TTR and word length [10] (p104). The z-score and *t* test results show that CAEMC uses more nouns (2.69 vs 2.31) and prepositions (0.89 vs 0.02) and that the average word length (2.43 vs 2.12) is longer. There is no significant difference in the use of attributive adjectives, but the TTR (-0.38 vs 0.53) is much higher in AAEMC. It indicates that CAEMC is inclined to use more information-loaded nouns and longer words with specific and specialized meanings and prepositions to link the nouns, but the language lacks variety and richness. The top ten nouns used in CAEMC are *year, company, development, market, industry, growth, management, power and production,* referring to companies' achievement or analysis of the market, production and industry. A close investigation of long words in CAEMC shows that most of the long words come from compound words such as *customer-orientation*, product names or nominalization such as *equipment* and *performance*, words typical in the equipment manufacturing industry. A comparison to Hu & Tan's study [21] also shows that the mean deviation between CAEMC and AAEMC (-5.82) is comparatively lower than that found in Hu & Tan's study (-8.62) [21]. The possible reason is that words in the equipment manufacturing industry tend to be specialized and long.

Meanwhile, AAEMC is more interactional and involved in that it uses more 1st person pronouns (-0.50 vs 1.41), present tense verbs (-1.8 vs -1.0) and private verbs (-1.27 vs -0.98). While Chinese companies tend to use *company/group* for self-reference, American companies tend to use *we, us* and *our*. The use of the first-person pronouns is often considered a language strategy of impression management. Thomas [25] argued that the first-person pronoun was employed by management to indicate their roles in the success of the company. However, the difference may be largely due to the cultural difference between China and America. According to Hofstede's cultural dimension theory [26], America is a small-power-distance culture where power is distributed equally within institution, while China is a large-power-distance culture where management's authority and power are expected to be respected. Therefore, Chinese management uses *company* and *group* rather than first-person pronouns to assert its authority. Moreover, present tense verbs and private verbs (e.g. *expect*, *believe*) are relatively high in AAEMC, suggesting that American management tends to interact with the readers in the immediate context and express their private attitudes and thoughts.

4.2.2. D2: Narrative vs Non-narrative Concerns

The negative and low D2 score (-1.70 vs -3.09) of the two corpora means that both corpora belong to non-narrative texts. However, a statistically significant difference is found between the two corpora (t=6.97, p<0.001), so significant that they are closest to two different genres (academic prose vs official documents), suggesting that AAEMC is more non-narrative and CAEMC leans more towards the narrative end.

The *t* test result shows that the difference mainly comes from the more frequent use of present participial clauses and past tense verbs in CAEMC and the high frequent use of present tense verbs in AAEMC. According to Biber [10] (p109), present participial clauses are often combined with the past tense verbs to vividly describe past events. A close examination of CAEMC reveals that most of the past participial clauses collocate with the past tense verbs as shown in Example 1. The most frequent past tense verbs used in CAEMC include *be, continue, amount, record, have, achieve, reach*. *Be* is used to show a relational or attributive description of the events while other verbs often collocate with figures to suggest the past achievement of the company. In comparison, AAEMC uses more present tense verbs such as *be, have, believe* and *expect* to describe the present situation and express

---

1 All scores compared here and in the following are the dimension scores or z-scores of CAEMC vs AAEMC.

its confidence and attitude, as illustrated in Example 2. Such difference may be attributed to the different cultural views of time. Chinese culture values the past and decision is often made based on the lessons from the past, while American culture places more importance on the present.

Ex. 1 However, during the Reporting Period, the international futures price of crude oil halved, leading to the reduction of the price of domestic oil products for 11 times in a row. (Beijing Jingcheng Machinery Electric Company Limited, 2014)

Ex. 2 We believe these metrics reflect the alignment of the Aerojet Rocketdyne product portfolio with a number of our government customers' highest priorities. (Aerojet Rocketdyne Holdings Inc., 2014)

4.2.3. D3: Explicit vs Situation-dependent Reference

The D3 scores for both corpora are positive and high (12.28 vs 9.43), with a statistically significant difference at p<0.001, indicating that they are both referentially explicit discourse depending less on context but CAEMC, in comparison, is more context-independent.

A high score on this dimension is related to linguistic features such as *wh*-relative clauses, phrasal coordination and nominalization and a low score is related to the use of adverbs [10] (p110). The *t* test result shows that CAEMC uses more phrasal coordination (using *and* to link two or more verbs, adjective or adverbs)(6.43 vs 5.55) and nominalization (e.g. *implementation, reduction*)(3.32 vs 2.31), as illustrated in Example 3, while AAEMC uses more time and place adverbs such as *today, now, recently, below, above*, as shown in Example 4. While time and place adverbs are used for text-internal or external referents, phrasal coordination and nominalization are the informational package that offers greater flexibility for information flow without referencing to context-dependent information such as agent, time and location [22].

Ex. 3 The Group will align the implementation of this strategic plan with the pace of the "Made in China 2025" strategy and ensure its effective execution, while striving to achieve the goal of reform and transformation to a global automobile company by 2020. (BAC Motor Corporation Limited, 2015)

Ex. 4 Today, the Foundation focuses on its mission to alleviate the root causes of poverty by supporting microfinance and access to clean water and better sanitation worldwide. (Caterpillar Inc., 2015)

4.2.4. D4: Overt Expression of Persuasion

Both corpora on this dimension have a negative and low score (-2.13 vs -2.10), with no statistically significant difference, suggesting that Chinese and American AEMC managements do not explicitly express their attitudes and assess the likelihood. Given that the letter to the shareholders is an official document that discloses the company's information to the investors and that explicit opinion or assessment of the likelihood may lead to controversy, it is understandable that both corpora present a low score on this dimension.

However, a close examination of the key features that determine the dimension score shows that AAEMC uses more infinitives (0.21 vs 1.3) and CAEMC uses more prediction modals (0.74 vs -0.17). Infinitives are often used as adjectives and verb complements in which the head adjectives or verbs often express the author's attitudes or stance [10] (p111), as illustrated in Example 5. The most frequent verbs and adjectives that collocate with the infinitives are *continue, able, want, positioned* and *need*, expressing the ability or desire to accomplish something. When it comes to prediction modals, the most frequent collocates on the left 3 and right 3 of modals *will, would, shall* in CAEMC are *group, company, industry, market, development, business*, indicating an emphasis on the market and industry prediction, as can be seen in Example 6. In comparison, the collocates in AAEMC are *we, that, which, growth* and *action,* indicating a prediction of the future action, as shown in Example 7.

Ex. 5 We expect to generate strong cash flow in 2015 and beyond to fund this plan. (AGCO Corp., 2014)

Ex. 6 The Chinese automobile industry will also enter a new monumental period by undergoing transformation and evolution. (BYD Company Limited, 2016)

Ex. 7 These actions and others will improve business continuity and productivity, and expand our access to new talent and technology. (Boeing Co., 2014)

4.2.5. D5: Abstract vs Non-abstract Information

Both corpora on this dimension have a positive score slightly over zero (0.87 vs 0.36), meaning that information is presented in a technical, abstract and formal way. A weak statistical significance at p=0.027 is found between the two corpora, indicating CAEMC presents the information in a more abstract way.

The key linguistic features on this dimension are conjuncts, passives, past participial clauses, and other adverbial subordinators [10] (pp111-112). The *t* test result shows that CAEMC differs from AAEMC in that the former uses more conjuncts (1.04 vs 0.24) and agentless passives (-0.6 vs -0.77) while the latter uses more other adverbial subordinators (0.47 vs 1.12) such as *since, while*. This finding is contradictory to Ren & Lu's findings [22] in which they point out the L1 influence results in the low frequency of conjuncts and passives in English written by Chinese. However, this finding suggests that the authors of the letters to the shareholders by the Chinese AEMC are aware of the language difference between Chinese and English and intentionally conform to the language standards of English-speaking countries. Another possibility is that many of the English texts in CAEMC are the translations of the original Chinese version. Since the agents are often omitted in the original Chinese texts, agentless passives are frequent in the English translation. Furthermore, a further examination of the conjuncts in the CAEMC and AAEMC shows that American companies tend to use conjuncts to list examples, contrast and show addition while Chinese companies tend to use conjuncts to show addition, contrast and cause-and-effect.

4.2.6. D6: Online Informational Elaboration

The D6 scores for both corpora are negative (-2.61 vs -1.50), with a statistically significant difference at p <0.001, showing that the Chinese and American corpora are well-planned, informationally integrated and not produced within a real-time constraint but American corpus, in comparison, is more informationally elaborated and contains more features of texts produced in real-time constraint

The major difference between the two corpora lies in the more frequent use of demonstratives (e.g. *these, this*) (-1.76 vs -0.66) and *that* clauses on object positions (-0.53 vs -0.2) in the American corpus and more frequent use of phrasal coordination (6.43 vs 5.55) in the Chinese corpus. Demonstratives improve the cohesion of the text and *that* clauses elaborate on the information often relative to the personal stance of the author [10] (pp113-114). Example 8 illustrates the use of both demonstratives and *that* clause in the American corpus. Phrasal coordination is a device for "idea unit expansion" and "information integration" [10] (p114), so the Chinese corpus is informationally dense and integrated, as illustrated in Example 9.

Ex. 8 This bold goal requires that we set and achieve higher expectations for our work and recognize that we now compete with companies in other sectors for capital, influence, talent and positioning in a business environment that is getting tougher, more complex and more global every day. (Boeing Co., 2016)

Ex. 9 The Company believes that good corporate governance is one of the key factors that drives the Company's success and balances the benefits among its shareholders, customers and employees, and will therefore continue to raise and improve its corporate governance standard. (CHTC Fong's Industries Company Limited, 2014)

4.3. Register Features and Future Profitability

To address the second question, we performed the independent-sample t tests on the dimension scores for the Chinese Growth subcorpus and the Chinese Decline subcorpus, and for the American Growth subcorpus and the American Decline subcorpus. The result is shown in Table 3. We can see that letters to the shareholders in the Chinese subcorpora vary significantly on D3 and D6 in terms of future profitability while the American subcorpora do not vary on any dimension.

The result indicates that when the ROE is expected to increase, Chinese letters to the shareholders tend to be more referentially explicit, less context-dependent and more informationally integrated, showing fewer features of real-time constraint texts. The specific linguistic features that vary in the Chinese subcorpora include higher TTR and the more frequent use of phrasal coordination and present participial clauses in the Chinese Growth subcorpus and more analytic negation (*not* and *n't*) in the Chinese Decline subcorpus. It can be concluded that those Chinese companies that expect a ROE growth in the next year produce letters to the shareholders with higher informational density, narrative information and informational integration. The possible reason is that those Chinese growth companies tend to impress the readers with more information so as to create a positive image. An examination of the concordance lines of the words *not* and *n't* in the Chinese Decline corpus shows that many of the analytic negations are used to describe the adverse industrial and economic environment, as illustrated in Example 10, indicating a language strategy of attributing the prospective decline to the external environment.

Ex. 10 During the year, the metal and mineral industry did not have enough momentum to drive significant rebound in demand. (China Dynamics (Holdings) Limited, 2015)

As shown in Table 3, the t test results for the American Growth subcorpus and Decline subcorpus reveal no statistically significant difference on any dimension, which indicates that letters to the shareholders by American AEMCs have relatively stable and consistent register features.

Table 3. *Independent-sample t test results for the dimension scores in the Chinese and American subcorpora.*

dimension	Subcorpus	Growth		Decline		t	df	p	MD
		Mean	SD	Mean	SD				
D1	Chinese	-23.93	2.64	-23.61	2.74	-0.73	148	0.464	-0.32
	American	-17.92	5.37	-17.98	5.16	0.08	148	0.937	0.07
D2	Chinese	-1.44	2.18	-1.98	1.88	1.63	148	0.105	0.54
	American	-2.92	1.20	-3.26	1.39	1.61	148	0.109	0.34
D3	Chinese	12.86	3.69	11.68	3.51	2.00	148	0.048	1.17
	American	9.30	3.20	9.56	3.23	-0.49	148	0.623	-0.26
D4	Chinese	-2.40	2.29	-1.86	2.13	-1.50	148	0.135	-0.54
	American	-1.98	2.39	-2.23	2.45	0.62	148	0.53	0.25
D5	Chinese	0.95	2.06	0.79	2.09	0.45	148	0.655	0.15
	American	0.54	2.06	0.18	1.74	1.18	148	0.238	0.37
D6	Chinese	-2.74	0.66	-2.48	0.72	-2.32	148	0.022	-0.26
	American	-1.46	1.06	-1.54	0.80	0.52	137.93	0.607	0.08

5. Conclusion

By applying Biber's MD analysis to examine the linguistic features and variation of the letters to the shareholders by Chinese and American AEMCs, the present study has come to the following findings:

Firstly, though the letters to the shareholders by Chinese and American AEMCs belong to the same text type of learned exposition, they vary significantly on D1, D2, D3 and D6

with a weak significant difference on D5. It indicates that the letters to the shareholders by Chinese AEMCs are informationally denser, relatively more narrative, more context-independent and informationally integrated while the letters to the shareholders by American AEMCs show more features of interaction with the readers, non-narrativeness, context-dependence and on-line informational elaboration.

Second, the key linguistic features that account for these differences include nouns, nominalizations, first-person pronouns, present tense verbs, past tense verbs, phrasal coordination, prediction modals, conjuncts, passives, demonstratives and *that*-clauses.

Third, the letters to the shareholders by Chinese companies vary significantly on D3 and D6 in terms of the future ROE, while the letters to the shareholders by American AEMCs remain stable and consistent regardless of the future performance of ROE.

Fourth, what industry the companies belong to may affect the linguistic features of the letters to the shareholders. The mean deviation on D1 is lower than that found in Hu & Tan's study [21], suggesting that the use of specialized terms of the industry affects the variation. Moreover, a more significant difference on D2 was found, indicating that Chinese AEMCs place more emphasis on narrating the past performance and situation due to the rapidly changing environment in the industry.

This study has given a comprehensive analysis of the linguistic features and variation of the letters to the shareholders by Chinese and American AEMCs. Our findings may help English learners, translators and business professionals to understand the register features and linguistic differences of the letters to the shareholders by Chinese and American AEMCs, and help them produce proper letters to the shareholders that achieve the communicative functions.

This study also has some limitations, some of which may become the scope for future research. First, our corpora can be expanded to include more samples in different companies and different years. It may be useful to examine the linguistic variation from a diachronic perspective. Second, though our study has found a significant difference in the Chinese subcorpora in terms of D3 and D6 by using the ROE, it may be worthwhile to examine more financial indicators.

Acknowledgements

References

[1]   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.

[2]   Mobasher, A., & Ali, A. M. (2015). Exploring the macrostructure of management forewords of corporate annual reports. *International Journal of Applied Linguistics and English Literature*, 4 (1), 14-25.

[3]   Hyland, K. (1998). Exploring corporate rhetoric: Metadiscourse in the CEO's letter. *The Journal of Business Communication*, 35 (2), 224-244.

[4]   De Groot, E. B., Korzilius, H., Nickerson, C., & Gerritsen, M. (2006). A corpus analysis of text themes and photographic themes in managerial forewords of Dutch-English and British annual general reports. *IEEE Transactions on Professional Communication*, *49* (3), 217-235.

[5]   Garzone, G. (2005). Letters to shareholders and chairman's Statements: Textual variability and generic integrity. *Genre variation in business letters*, 179-204.

[6]   Conaway, R. N., & Wardrope, W. J. (2010). Do their words really matter? Thematic analysis of US and Latin American CEO letters. *The Journal of Business Communication*, *47* (2), 141-168.

[7]   Clatworthy, M. A., & Jones, M. J. (2006). Differential patterns of textual characteristics and company performance in the chairman's statement. *Accounting, Auditing & Accountability Journal*, 19 (4), 493-511.

[8]   Wang, L., & Li, Z. (2018). A multi-dimensional contrastive study on the intertextuality in Chinese and American business discourse. *Foreign Language Learning Theory and Practice*, 03, 56-62.

[9]   Huang, Y. (2012). Contrastive Analysis of Distribution and Clustering of Metadiscourse Markers: CEO Letters in English Banking Annual Reports. *Foreign Language and Literature*, 28 (04), 84-90.

[10]   Biber, D. (1988). *Variation across speech and writing*. Cambridge: Cambridge University Press.

[11]   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.

[12]   Aguado-Jiménez, P., Pérez-Paredes, P., & Sánchez, P. (2012). Exploring the use of multidimensional analysis of learner language to promote register awareness. *System*, *40* (1), 90-103.

[13]   Biber, D. (1992). On the complexity of discourse complexity: A multidimensional analysis. *Discourse Processes*, *15* (2), 133-163.

[14]   Grieve, J., Biber, D., Friginal, E., & Nekrasova, T. (2010). Variation among blogs: A multi-dimensional analysis. In *Genres on the Web* (pp. 303-322). Springer, Dordrecht.

[15]   Crosthwaite, P. (2016). A longitudinal multidimensional analysis of EAP writing: Determining EAP course effectiveness. *Journal of English for Academic Purposes*, *22*, 166-178.

[16]   Kruger, H., & Van Rooy, B. (2016). Constrained language: A multidimensional analysis of translated English and a non-native indigenised variety of English. *English World-Wide*, *37* (1), 26-57.

[17]   Huang, Y., & Ren, W. (2020). A novel multidimensional analysis of writing styles of editorials from China Daily and The New York Times. *Lingua*, *235*, 102781.

[18] 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.

[19] Jiang, J., & Xu. J. (2015). A corpus-based multi-dimensional analysis of business English registers. *Foreign Language Teaching and Research*, 47 (2), 225-236+320.

[20] Wang, L., & Bu, H. (2018). A multi-dimensional contrastive analysis of functional features of Chinese and US corporate discourse. *Journal of PLA University of Foreign Languages*, 41 (02): 96-103+161.

[21] Hu, C., & Tan, J. (2020). A multi-dimensional analysis of register features of letters to shareholders by Chinese and American enterprises. *Foreign Language and Their Teaching*, 03, 66-75+149.

[22] Ren, C., & Lu, X. (2021). A multi-dimensional analysis of the Management's Discussion and Analysis narratives in Chinese and American corporate annual reports. *English for Specific Purposes*, *62*, 84-99.

[23] Nini A. (2015). *Multidimensional Analysis Tagger (Version 1.3) – Mannual*. retrieved from https://sites.google.com/site/multidimensionaltagger. (assessed 1 March 2021).

[24] Xu, J., Liang, M. & Jia, Y. (2012). BFSU PowerConc 1.0. National Research Centre for Foreign Language Education, Beijing Foreign Studies University.

[25] Thomas J. (1997). Discourse in the marketplace: The making of meaning in annual reports. *The Journal of Business Communication*, 34 (1), 47-66.

[26] Hofstede G, Hofstede G J, Minkov M. (2005). *Cultures and organizations: Software of the mind*. New York: Mcgraw-hill.