2010

The predictive ability of corporate narrative disclosures: Australian evidence

Yinan Dong
Edith Cowan University

2010

The Predictive Ability of Corporate Narrative Disclosures: Australian Evidence

Yinan Dong
Edith Cowan University

Edith Cowan University

Copyright Warning

EDITH COWAN UNIVERSITY

The Predictive Ability of Corporate Narrative Disclosures: Australian Evidence

---

by

Yinan DONG

B. Bus (Accounting)

Supervisor: Professor Malcolm SMITH

Date of submission: November 2010

A thesis submitted in partial fulfillment of the requirements for the award of Bachelor of Business with Honours at the Business and Law Faculty, Edith Cowan University, Joondalup campus.

Declaration

I certify that this thesis does not, to the best of my knowledge and belief:

(i) incorporate without acknowledgement any material previously submitted for a degree or diploma in any institution of higher education;

(ii) contain any material previously published or written by any person except where due reference is made on the text; or

(iii) contain any defamatory material.

(iv) I also grant permission for the library at Edith Cowan University to make duplicate copies of my thesis as required.

Signature: Yinan DONG

Date: 22nd November 2010

Abstract

The main objective of this study is to contribute to the academic literature by investigating the relationship between narrative disclosures and corporate performance based on Australian evidence. The research design takes as its starting from the content analysis of discretionary narrative disclosures conducted by Smith and Taffler (2000), and extends their research by combining thematic content analysis and syntactic content analysis.

This study focuses on the discretionary disclosures (the Chairman's Statement) of Australian manufacturing companies. Based on the Earnings per Share (EPS) movement between 2008 and 2009, 64 sample companies are classified into two groups: good performer and poor performer.

This study is grounded on signalling theory and agency theory, and links with the impression management strategy. Based on two branches of impression management (rationalisation and enhancement), six groups of variables are collected to examine narrative disclosures from both quantity ("what to disclose") and quality ("how to disclose") perspectives. Manual coding and two computer-based software programs are employed in this study.

This study finds that the word-based and theme-based variables based on discretionary disclosures are significantly correlated with corporate performance. Moreover, word-based variables can successfully classify companies between good performer and poor performer with an accuracy of 86%. However, there is no significant relationship between corporate performance and report size, use of long words (as a proxy for jargon), FLESCH readability score, or persuasive language.

The main value of this study is to build a classification model based on Australian evidence for continuing companies, since most prior research focuses on UK, US and New Zealand companies and is based on a healthy/failed distinction.

Acknowledgements

First of all, a big thank you to my fantastic supervisor Professor Malcolm Smith, for his ideas, help, and patience throughout my study. Without his guidance and input, this study would not have even got off the ground. I have gained not only accounting and research knowledge from him, but also an enthusiasm for research.

Many thanks to Dr. Greg Maguire and Dr. Jo McFarlane who helped me with my writing skills and motivated me to keep improving. I would also thank Dr. Peter Standen, Dr. Theo Christopher, and all staff from SOAR and Office of Teaching and Learning staff for their kindness and support whist I was down there.

I would like to thank Yun Ren, my good friend and also my study mate for all my university time. Thanks to her company and help when I felt my research was in limbo. Thanks also to all my friends who gave me encouragement and happiness.

Most importantly, my special thanks to my parents for all their love in so many ways over the years during my life and education journey, and for giving me the belief that I can achieve my dream. This thesis is dedicated to them.

Thank you all.

Table of Contents

List of Figures

List of Tables

Chapter One: Study Introduction

1.1 Research background

There are two kinds of narrative disclosures in the annual report: compulsory disclosures and discretionary disclosures. Compulsory disclosure information is such as Director's Report; and discretionary disclosures information includes Chairman's Statement (also called President Letter, Letter to Stakeholder, etc.), Management Discussion and Analysis (MDA), Operating and Financial Review (OFR), Notes to the Financial Statements. This study will only concentrate on discretionary disclosures, investigate the relationship between corporate discretionary narrative disclosures and financial performance characteristics.

Discretionary narrative disclosure is a way that companies voluntarily report their information, which can be quantitative or qualitative, financial or non-financial, using formal or informal channels. It is a unique advertisement for companies designed to elicit responses from its readers such as buying more stock, lending more money, refraining from selling currently held stock, or supporting management (Tennyson, Ingram, & Dugan, 1990).

Since corporate managements have the choice to select disclosure content and style, they can use this communication channel to provide specific information to influence or manipulate a broad range of outside information users. From 1880s, a growing number of companies have voluntarily disclosed information in the annual report (Hackston & Milne, 1996). Since then, discretionary disclosures have drawn an increasing amount of attention for accounting researchers (Meek, Roberts & Gray, 1995). Andersen (2000, p. 7) surveyed on UK companies, and found that the narrative disclosures of the annual report have increased from 45% in 1996 to 57% in 2000. Meanwhile, narrative disclosures have become "longer and more sophisticated" over the past decades (Merkl-Davies & Brennan, 2007, p. 118). Therefore, it is essential to study narrative disclosures based on the current data. This study here would

concentrate on non-financial discretionary narrative disclosures (the Chairman's Statement) by formal channels (annual reports) between 2008 and 2009 fiscal years. To start with, two areas of research significance will be discussed below.

1.2 Research significance

Merkl-Davies and Brennan (2007) reviewed and synthesised previous research on discretionary narrative disclosures, and stated that there are two assumed purpose of narrative disclosures in prior research: to provide incremental information to help outside information users making better decisions; or to behave opportunistically to impair the ability of outside information users to make rational decisions based on information asymmetries. The research significance of the two alternative approaches will be discussed in detail as follows.

1.2.1 Provision of incremental information

Compared with financial disclosures, narrative disclosures contain complementary and incremental information (Smith & Taffler, 1995). Financial disclosures are intended, as the Financial Accounting Standards Board (FASB, 1978) stated "to assist investors and creditors in projecting the amount, timing, and uncertainty (risk) of future dividends and interest payments" (para.21). The major limitation of the financial statement is that the information is a review of past corporate performance which has already happened. As "old news is no news", information users are more interested about the corporate "future" information, such as "the firm's perception of the importance of economic and industry-specific factors, and references to current action, future strategies and intended policies" (Smith & Taffler, 1995, p. 1195).

In narrative disclosure sections of the annual report, companies would disclose information such as company and industry general background, past performance results, expectations of future performance, and potential opportunities and challenges. Such narrative information is a valuable ingredient to outside information users to make judgments and decisions.

1.2.2 Impairment of information asymmetry

It is assumed that management has superior information compared to outside information users, on the prediction of corporate future performance (Healy & Palepu, 2001); that is referred to as information asymmetry. This information asymmetry can be reduced by providing more disclosures or by increasing the disclosure quality, since narrative disclosures may provide valuable incremental information to outside users (Healy & Palepu, 2001). Meanwhile, impairing information asymmetry can benefit companies by mitigating the negative selection costs (Verrecchia, 2001), and help to build an efficient capital market as well (Healy & Palepu, 2001).

However, companies are not always in favour of increasing disclosure transparency, as poorly performing companies' managements tend to hide negative information by disclosing opportunistically. In these poorly performing companies, managements have a strong incentive to control and manipulate information users' impressions and perceptions by selecting the discretionary disclosure content and the disclosure approach. Under this impression management strategy, companies intend to influence the information users' decisions, and get benefits by providing favourable information. The detail of impression management will be discussed in the theory chapter, Chapter Three.

Because of the impression management, the quality of narrative disclosure has aroused the public's attention (Clarke & Dean, 2007; Donoher, Reed, & Storrud-Barnes, 2007). If managements use narrative disclosures as part of an impression management strategy, the value of narrative disclosures will be undermined, and the judgments of outside information users may be negatively influenced. Thus, the study of discretionary narrative disclosures serves a vital part in accounting research (Merkl-Davies & Brennan, 2007). Healy and Palepu (2001) provide a framework for analysing corporate disclosures in a capital markets setting, and they argued that due to information asymmetry and agency conflicts between

management and outside information users, the study of disclosure is essential (Healy & Palepu, 2001). The following paragraphs will illustrate the research significance from both points of outside information users and regulation authorities.

1.3 Various information users

1.3.1 Outside information users

The narrative disclosure in the annual report is an important instrument for companies to communicate their performance, risk and opportunity to outside information users. Sell-side analysts cited almost twice the amount of information provided by narrative disclosures compared with the financial statement (Rogers & Grant, 1997); auditors use narrative information as supplementary information to analyse and corroborate corporate going concern decisions (Smith & Taffler, 2000); and Bryan (1997) suggested corporate disclosures can assist in assessing corporate short-term prospects, and help investors to reduce their investment risk. In summary, narrative studies can help public users make better decisions.

1.3.2 Regulatory authority

Based on the current changes of economy and market, as accompanied by accounting scandals (such as Enron's bankruptcy, Parmalat and WorldCom), regulatory authorities "worldwide have been showing an increasing interest in expanding disclosures in annual reports in addition to those required in the financial report" (Hrasky, 2008, p. 5; Clarke & Dean, 2007; Donoher et al., 2007). For improving corporate disclosure accountability and transparency, regulatory bodies set up relevant regulations and rules to strengthen the disclosure information quality.

One of the extremely influential regulations, the Jenkins Report, was published in US (American Institute of Certified Public Accountants (AICPA), 1994). In 2002, the US government emphasised the necessity to improve the information quality of disclosures in the Sarbanes-Oxley Act. In terms of UK, the government implemented

a review to advance disclosure information (Department of Trade and Industry 〔DTI〕 2004). Although regulators of different countries have not fully addressed the format and content of annual reports, for reducing the information asymmetry, Australian standard setters do ask companies to "include, either by law or custom, other financial and non-financial information" as an obligation (Australian Auditing Standards Board 〔AUASB〕, 2006: para.7, cited by Hrasky, 2008, p. 13). Moreover, the Corporate Law Economic Reform Program Act 2004 (also known as CLERP 9) has addressed the advancement of continued corporate disclosure (Parker, 2005).

Since regulatory authorities give companies self-determination rights to some extent, they would like to know how companies use the rights, and whether the auditing and accounting regulations work perfectly. Healy and Palepu (2001) stated, if the regulations are imperfect, managements are more likely to use their superior knowledge of corporate performance to conceal negative information. As they stated "Management motives for making discretionary disclosures and their credibility are, therefore, interesting empirical questions" (Healy & Palepu, 2001, p. 420). Thus, it is essential to investigate narrative disclosures, and to help regulatory authorities know whether they need additional regulation and supervision in order to improve corporate transparency and management credibility.

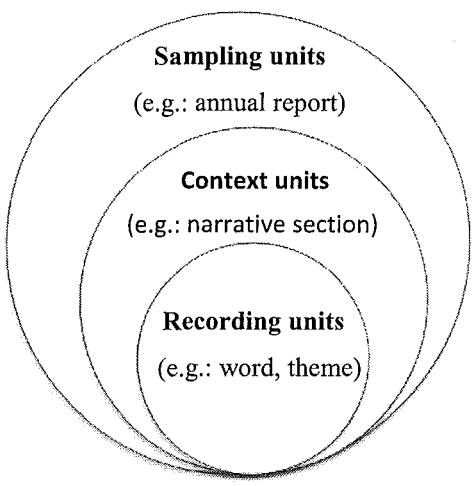
To sum up, the study of narrative disclosures helps public and information intermediaries (such as financial analysts and rating agency) to know how complete the corporate information is, to uncover managements' superior information, to get a transparent and reliable understanding of corporate profiles; and be guided in making better and unbiased decisions.

1.4 Three data units of narrative

In terms of narrative disclosure information, there are three data units: sampling unit, context unit, and recording unit (also known as text unit) (Krippendorf, 1980). As Figure 1 shows, within each sampling unit is a context unit, and within each context unit is a recording unit. Context unit is the largest informational segment which can be searched in order to identify a recording unit, and the information content of recording unit is often interpreted in conjunction with all other recording units within the context unit (Jones & Shoemaker, 1994). For example, Smith and Taffler (2000) used content analysis and examined discretionary narrative disclosures of UK corporate annual reports, analysed both by word- and theme- bases. They found that there is an association between the content of the Chairman's Statement and corporate performance. In their study, the annual report is a kind of sampling unit; the narrative disclosure (the Chairman's Statement) is one of the context units; the word and theme used for analysis could be seen as two kinds of recording unit.

In content analysis research, the corporate annual report is the most popular sampling unit, and there are various recording units (text units), such as word, phrase, theme (Neuman, 2006). Among these recording units, thematic content analysis usually uses word and theme units; while for syntactic content analysis, the most common recording units are sentence, word, and syllable.

Figure 1 Three data units of narrative



1.5 The structure of the thesis

The first chapter initially introduced background information of narrative discourses, followed by two study significance, with illustration from different information users' point of views. In the end, three research data units were introduced.

The next chapter of this thesis reviews relevant narrative disclosure research based on different analysis approaches. Chapter Three discusses the theoretical framework of this thesis, develops six related hypotheses, and outlines the research framework of this thesis. The research method of this study is described in Chapter Four, followed by details of the research results, test of hypotheses, reliability and validity in Chapter Five. The study discussion, values, limitations and some concluding comments for further study are presented in the last chapter, Chapter Six.

Chapter two: Literature Review

2.1 Introduction

The previous introduction chapter outlined the structure of this thesis, discussed the relevant background information regarding discretionary narrative disclosures, study significance and information users, and three data units of narrative study. This chapter reviews previous narrative research literature, and it aims to get a clear outline of narrative disclosure study in order to develop an appropriate research method for this research.

The initial discussion of this chapter explains a classification of narrative study, and gives a brief introduction. Then, this chapter focuses on content analysis study, and discusses two approaches of content analysis. The discussion includes a review of relevant research, statistical analysis, research device introduction, explanation of reliability and validity, and critical analysis, followed by conclusions.

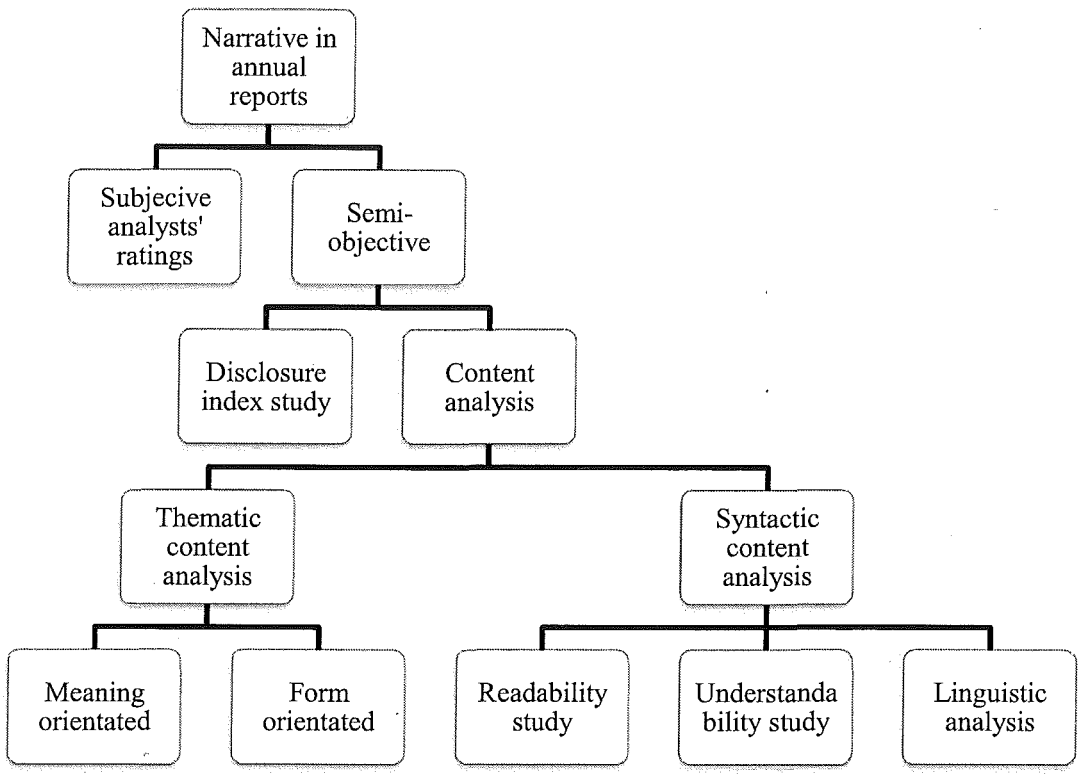
2.2 Classification of analysis approaches

Previous research has two main objectives which regard to corporate narrative disclosures with a focus on either the corporate actual performance, or their external social influences. This research focuses only on the former research objective. It will examine the association between the corporate narrative disclosures and the corporate performance characteristics (good/bad performance). This section will focus on the related literature in this field.

There are various approaches to analysing the quantity and quality of a narrative in an annual report, especially the relationship between narrative disclosures and corporate performance characteristics. Beattie, McInnes, and Fearnley (2004) identified two major classifications of narrative analysis approach: subjective ratings and semi-objective approaches. The latter approach includes the disclosure index study and

content analysis, with content analysis composed of three approaches: thematic content analysis, readability studies and linguistic analysis (Beattie et al., 2004). Moreover, Jones and Shoemaker (1994) grouped readability studies and linguistic analysis into one category called syntactic content analysis. The summarisation of the above-mentioned two classifications is outlined in Figure 2, and all the five approaches will be introduced subsequently.

Figure 2 Narrative analysis approaches



The subjective analysts' ratings approach was created by the Association of Investment Management and Research （AIMR） (formerly the Financial Analysts Federation （FAF）). The reliability of this approach has been criticised by many researchers (Lang & Lundholm, 1993; Healy & Palepu, 2001) as it involves several biases. Moreover, the publication of these ratings stopped in 1997, and only focused on US companies. There are both spatial and temporal limitations. For these reasons, this study will not adopt this approach to measure the quality of narrative information about Australian companies.

The disclosure index study is grounded in the assumption that a disclosure's quantity and quality of the disclosures are positively related, and this approach uses the amount of disclosure to reflect the disclosure quality. The disclosure indices were defined by Dixon, Coy and Tower (1991). There are usually three levels of coding scheme in this approach (Botosan, 1997; Robb, Single, & Zarzeski, 2001). The coding schemes may vary from research to research, but all have the same principle by seeking to transfer disclosure's quality into quantified measurement. This approach has been criticised by Marston and Shrives (1991) since it cannot reflect the disclosure quality, and to some extent it is judged to be subjective.

Content analysis is a well-developed social technique for "gathering and analysing the content of text" (Neuman, 2006, p. 322). It is defined by Krippendorff (1980) as "a research technique for making replicable and valid inferences from data to their context" (p. 21). Content analysis has been used frequently in the humanities and social sciences, but relatively rare in accounting research. Compared with other types of scientific evaluation, the distinguishing feature of content analysis is that it is unobtrusive because documents can be evaluated without the knowledge of the communicator (Jones & Shoemaker, 1994).

There are two subsets of content analysis: thematic analysis and syntactic analysis. Thematic analysis identifies specific trends, attitudes, or content categories from the text and then draws inferences from them; while syntactic analysis, on the other hand, centres upon the difficulty of reading and understanding the textual message (Jones & Shoemaker, 1994). Hrasky (2008) summarised that thematic analysis looks at "what the narrative is written", which focuses on the verbal side of narrative disclosures. In terms of syntactic analysis, it focuses on assessing aspects of "how the narrative is written". No matter which subsets are employed, they all require encoding and scoring of the classified narrative data. During these procedures, high levels of validity and

reliability are required. The following subsections will discuss relevant researches, and demonstrate the validity and reliability of content analysis.

2.2.1 Thematic content analysis

As mentioned above, the recording units (text units) vary with different forms of content analysis. According to Jones and Shoemaker (1994), the most common recording units are "themes", followed by "words". Based on the two different text units, thematic content analysis can be classified into two categories (Smith & Taffler, 2000): meaning orientated (subjective) analysis, which is based on a theme variable; and form orientated (objective) analysis, which is based on a word variable. The two types of thematic content analysis will be introduced as follows.

Smith and Taffler (2000) suggested that "meaning orientated (subjective) analysis focuses on analysis of the underlying themes in the texts under investigation" (p. 627). It needs prior specification of categories and judgments. Moreover, "theme clusters of words with different meanings or connotations that are taken together refer to some theme or issue" (Weber, 1990, p. 37).

Form orientated (objective) analysis involves "routine counting of words or concrete references" (Smith & Taffler, 2000, p. 627). It is an objective analysis because the analysis procedure is "relying upon interrelationships in the data rather than subjective decisions by readers to identify content" (Tennyson, et al., 1990, p. 398).

There is an argument about which approach is the more reliable of the two. Different researchers hold different opinions. Krippendorff (1980) argued that a theme-based meaning orientated approach is preferable because it determined the hidden messages conveyed in the narrative disclosures. On the contrary, Weber (1990) stated that the word category that decided by co-variation among high-frequency words is more reliable than themes. Moreover, a word-based approach can reduce the need for researcher intervention, and thus, avoiding researcher bias. To sum up, both approaches

to content analysis are important, and both can be used to predict corporate performance (Smith & Taffler, 2000). Moreover, a combination of keywords and themes in the Chairman's Statement is able to improve the degree of discrimination in the classification of financially healthy and failed companies (Smith & Taffler, 2000).

A number of researchers have adopted either meaning orientated (word-based) or form orientated (theme-based) content analysis to examine the relationship between narrative and corporate performance. The research found that the disclosure information is significantly different between companies with different corporate performance. Ingram and Frazier (1983) conducted an explanatory study that stated the correlation between narrative disclosures and corporate performance across three industries. Tennyson et al. (1990) adopted a word-based, thematic content analysis and provided the usefulness of narrative disclosures in explaining financial distress. The pioneering research of Smith and Taffler (2000) examined the association between narrative disclosures and financial performance (healthy/failed) based on 66 UK manufacturing companies. They employed both word-based and theme-based content analysis methods, and found that the Chairman's Statement alone could distinguish between healthy and failed companies as accurately as carefully developed financial ratio based z-score models. Furthermore, they suggested that the use of narrative indicators is likely to contribute to reduce Type II error rates of around 20 per cent (Smith & Taffler, 2000). Moreover, some studies find that poorly performing companies have a tendency to disclose more positive information, use more positive keywords (Brennan, Guillamon-Saorin, & Pierce, 2009), or emphasise the managerial optimism about corporate future performances (Matsumoto, Pronk, & Roelofsen, 2006, cited in Merkl-Davies & Brennan, 2007). Rutherford (2005) counted the frequency of 90 keywords, and came to the conclusion that poorly performing companies tend to emphasise and overstate the positive information regardless of whether or not it is misleading.

2.2.2 Syntactic content analysis

The primary strength of thematic analysis when used with accounting narratives is its ability to identify the motivations and concerns of accounting communicators, while the importance of syntactic content analysis is highlighted by the fact that it can furnish objective benchmarks to narrative study. Pennebaker (2002) stated that since the writing style provided richer information than the content, the study focuses on how people talking about a given topic became far more important than the study topic of what people are talking about. Furthermore, this approach is arguably less problematic than thematic inference because word, syllable, and sentence counts can be performed relatively objectively (Jones & Shoemaker, 1994). Generally, there are two syntactic content analysis approaches which are commonly used: the readability study and the understandability study. Moreover, more complex linguistic studies have been considered more widely by researchers recently (de Beaugrande & Dressler, 1981; Roseberry, 1995; Sydserff & Weetman, 1999).

Readability & understandability studies

This approach assesses corporate performance by testing the cognitive difficulty of the text. It is necessary to assess how well the narrative message is presented, because there may be an information gap between producer and user. This information gap may lead to negative decision-relevance consequences. For successful disclosure information to be conveyed, there are two requirements that need to be satisfied: text-centred readability (the complexity of the display) and reader-centred understandability (the capability of users in discerning the appropriate meaning) (Smith & Taffler, 1992b).

Many prior researchers (Adelberg & Razek, 1984; Jones, 1988) treated readability and understandability as synonymous and did not make any distinction between the two. However, the experimental research of Smith and Taffler (1992b) suggested that the difference between "readability" and "understandability" was marked and measureable.

In their research, they adopted the LIX score, the FLESCH score, and the CLOZE test respectively and found the level of association between LIX and FLESCH scores was high, while their correlation with the CLOZE test was low. This proved "readability" and "understandability" to be two different concepts that in conflict with the assumptions in the prior literature (Smith & Taffler, 1992b). The CLOZE test is an excellent predictor of textual content. However, it has been doubted recently on its role as a measurement of "understandability", since it correlated poorly with other recognised measures of unberstandability (Jones, Smith, & Whale, 2010). For this reason, only the readability studies will be reviewed in this paper.

Merkl-Davies and Brennan (2007, p. 133) summarised, there are four categorises of study in the readability research field:

> "(1) *reading difficulty of annual report narrative, (2) variability of readability of different narrative sections of annual report, (3) association between the reading difficult of annual report narratives and various firm characteristics, most commonly firms performance, and (4) studies focusing on methodology development".*

The purpose of this paper is to investigate the relationship between narrative disclosures and corporate performance, therefore, only the third category will be addressed here.

An important step in this research field has already been undertaken by Smith and Taffler (1992a & 1992b). In their study, they used the FLESCH score and LIX scores as indicators of readability, and found that the narrative discourse quality is positively related with corporate performance: good financial performance is associated with a clear Chairman's Statement narrative, which is reflected by high levels of readability. This research indicated that readability can be used to predict corporate performance.

Although "readability and understandability" study is a dominant narrative research method, it has been criticised as having four limitations, listed below (Jones & Shoemaker, 1994; Beattie, et al., 2004):

- The measurement of reading difficulty is designed for children' writings and is already out of date. It may be inappropriate for evaluating the adult-based and technical accounting narratives.
- Readability scores focus on word- and sentence- level features and not on whole-text aspect.
- The readability formula takes no account of the interests and motivations of the reader.
- Even if these first three major criticisms are set aside, many of the prior syntactic studies lack robustness, and do not reveal the actual comprehension process.

Linguistic analysis

For addressing these criticisms, Sydserff and Weetman (1999) introduced a new method – the texture index of linguistic analysis. They adopted this texture index from applied linguistics originally as an alternative to readability formulas which offers practical validation for application of a texture index, however, this approach is able to "capture much richer set of text characteristics and is shown not to be associated with readability scores" (Beattie, et al., 2004, p. 212). Therefore, this approach itself can be seen as a powerful tool for analysis of accounting narratives.

Compared with prior readability studies, a linguistic analysis approach provides a unit-by-unit analysis with valid theory and advanced methodology. Moreover, two indexes, topicality and intertextuality, allow the reader to be involved in the study (de Beaugrande & Dressler, 1981). However, this approach is more time-consuming than computer-based readability study. Thus, only satisfying validity is not attractive enough to take the place of readability formula, unless it can provide some narrative

information that cannot be captured by readability formula. In terms of further study, Sydserff and Weetman (1999) suggested "explore more precisely the relation between textual difficulty, as measured by readability formulas; and ratings of texture, as measured by the texture index" (p. 478). Beattie and her colleagues recommend "weightings for each text characteristic" (Beattie et al., 2004, p. 213).

2.3 Statistical analysis

Generally, two statistical techniques have been used in developing prediction models. They are multiple discriminant analysis (MDA) and the logistic regression. The multivariate technique such as linear discriminant analysis (LDA) is able to distinguish healthy (non-failed) and failed companies with a high degree of accuracy. LDA, especially z-score (Altman, 1968) is commonly employed to discriminate corporate status.

Smith and Taffler (1992a) have suggested that based on the information conveyed by a Chairman's Statement, LDA might be adopted to identify whether a company could potentially fail. This assumption has been successfully proved by the same authors in 1995. In that research (Smith & Taffler, 1995), they used an appropriately weighted linear discriminant model (z-score), and confirmed that the narrative statement alone could be used as a significant indicator of corporate performance. Moreover, Smith and Taffler (2000) implemented LDA and Fisher discriminant analysis, and concluded that both word-based and theme-based content analyses were able to correctly predict corporate performance, and suggested that the accuracy of existing models might be improved by combining the variables from financial ratio and word-based ratio models.

Some researchers argue that MDA is not statistically optimal because of two shortcomings. However, these critics are doubted or have been remedied by some researchers. Firstly, the opposition argues this approach is inappropriate if the joint distribution of the independent variables is not multivariate normal, whereas logistic

regression does not restrict the distribution of independent variables with such severity (Tennyson et al., 1990). Nevertheless, the logistic regression approach used by Tennyson et al. (1990) research was questioned by Smith and Taffler (2000), since their empirical results were so disappointing. Secondly, MDA is criticised because the z-score may over-predict failed companies, as demonstrated by excessive Type two errors. This is a major deficiency with previous z-score prediction models (Smith & Gunalan, 1996). To address this problem, Smith and Gunalan (1996) examined the companies whose z-score profiles were similar to failed companies, and those which were able to reverse the bankruptcy trend. They selected the matched failed and recovered UK companies, and built a discriminant model to distinguish between the two company groups. This model has provided a useful discriminant between failed companies and recovery candidates, and improved predictability.

On the contrary, there are two advantages of MDA technique. First of all, compared with univariate study, the MDA technique is advanced "by considering an entire profile of characteristics common to the relevant firms, as well as the interaction of these properties" (Altman, 1968, p. 592). Furthermore, MDA reduces "the analyst's space dimensionality" (Altman, 1968, p. 592). Because of the two superior characteristics, the MDA technique is widely used by researchers in classification study.

To sum up, although there is some negative side in MDA technique, a number of researchers have proved that the predictability of the MDA technique is still robust in this performance predicting research field (Dames, 1979). Therefore, this study will adopt MDA technique (z-score) for statistic analysis.

2.4 Study devices

Both thematic and syntactic content analyses can be used to analyse and predict corporate performance. However, the biggest problem of using the two methods is the

bias during classification and coding processes. Generally, there are two types of coding – manual coding and computer coding. Compared with computer coding, manual coding is more prone to measurement errors, and bias, while computer coding is unable to use intuition to resolve ambiguities caused by symbolic meanings (Jones & Shoemaker, 1994). To reduce the coding bias, manual coding asks that all coders follow common assumptions about the coding of words over time; and computer coding requires more logical and sophisticated software packages (Jones & Shoemaker, 1994).

For this research, both thematic content analysis and syntactic content analysis will be involved to address "what" and "how" narrative information is disclosed by companies with different financial performance. A manual coding approach will be adopted in thematic analysis as it can improve the reliability in terms of "how" information can be disclosed; whereas a computer-based coding approach will be employed, in addition to manual methods, in syntactic thematic analysis. The Linguistic Inquiry and Word Count (LIWC) software will be used; the attraction of using this software is that by simply counting functional and emotional words in a given speech or text sample, a researcher could presumably get cues about the writers' thought processes, emotional states and motivation, and measure people's need states (Pennebaker, 2002; Tausczik & Pennebaker , 2010).

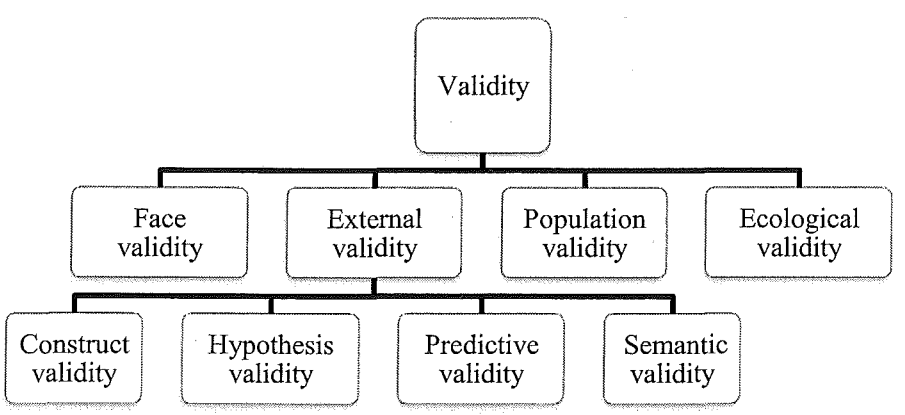
LIWC is a word count strategy developed by Pennebaker and his colleagues in 2002. It uses "a number of judge-defined dictionaries that categorize words into each of over 70 linguistic or psychologically-relevant categories" (Pennebaker, Mehl, & Niederhoffer, 2003, p. 553). These psychologically meaningful categories include negative and positive emotion words, articles, prepositions, pronouns, and cognitive words. The significance of LIWC is that it helps researchers to link daily word use to a broad array of real-world behaviours by providing linguistic analysis of each text (Pennebaker, 2002; Tausczik & Pennebaker, 2010).

2.5 Reliability and validity

Reliability and validity are paramount in content analysis. Krippendorff (1980) indentified three measurements of reliability: stability, reproducibility or inter-coder reliability, and accuracy. Among the three types of measurements, inter-coder reliability is the most commonly used one. There is no set answer for the question of how high the level of reliability must be (Krippendorff, 1980), but Krippendorff suggested that "inter-coder reliability correlations in excess of 80 per cent should be sought" (Smith & Taffler, 2000, p. 637). Moreover, there are two methods to evaluate reliability: coefficient of agreement, and Scott's pi. The former method does not include the likelihood of random agreement, and the latter method is recommended by many researchers. In this research, an independent check will be used to verify reliability, as suggested by Krippendorff (1980).

Validity relates to how well the results of the study mirror reality (Jones & Shoemaker, 1994). Weber (1990) classified validity into four categories: face validity, external validity, population validity, and ecological validity. For external validity, there are four branches: construct validity, hypothesis validity, predictive validity, and semantic validity. The classification is show in Figure 3. Research should aim for high levels of all these validities.

Figure 3 Classification of validity



2.6 Critical analysis and research gaps

There is an extensive literature on performance prediction, but most studies have used financial ratios as variables; little attention has been paid to the predictive ability of corporate narrative disclosures. Although these financial ratio models can successfully distinguish bankrupt companies from healthy companies with up to 85%-95% accuracy (Tennyson et al., 1990), narrative information can potentially provide a different scope and incremental value to predictive ability (Smith & Taffler, 1995). Moreover, most relevant research has used UK and US evidence, and there is no model directly applied to Australian manufacturing companies. As disclosure regulations vary between different countries, it is necessary to develop a predictive model based on Australian evidence.

2.7 Conclusions

This thesis will adopt both thematic content analysis and syntactic content analysis. In terms of the thematic content analysis, word-based and them-based variables will be collected manually and taken into consideration; and for the syntactic content analysis, a readability study will be conducted. Although the validity of readability is problematic in syntactic content analysis, the readability formula (FLESCH) is inexpensive and still helpful (Sydserff & Weetman, 1999). Moreover, as there are increased demands of developing objective methods of both thematic and syntactic content analysis (Sydserff & Weetman, 1999), a computer-based measurement (Linguistic Inquiry and Word Count software) will be used in this research.

The initial discussion of this chapter introduced the classification of narrative disclosures. Then, this chapter centred upon an extensive literature about the association of narrative disclosures and corporate performance, especially in the area of content analysis study. A summary of each content analysis approach and related relevant literature is listed in Table 1. The following chapter will discuss the underlying theoretical perspective of this research study, and develop hypotheses.

Table 1 Summary of content analysis research

	Relevant research	Significance	Criticism
Thematic content analysis			
Meaning orientated	• Ingram and Frazier (1983) • Smith and Taffler (2000)	Determining the hidden messages (Krippendorff, 1980)	Reliability questioned during classification coding processes (Jones & Shoemaker, 1994)
Form orientated	• Ingram and Frazier (1983) • Tennyson, et al., 1990 • Smith and Taffler (2000) • Rutherford (2005) • Aerts (2005)	Less researcher bias involved (Weber, 1990)	
Syntactic content analysis			
Readability study	• Smith and Taffler, 1992b • Courtis (1995)	Still helpful and prevalent Courtis (1998)	• Out of date • Not on whole-text aspects • No account of reader's motivations • Lack robustness (Sydserff & Weetman, 1999; Beattie, et al., 2004)
Understandability study	• Smith and Taffler (1992a) • Smith and Taffler (1992b)	Excellent predictor of textual content (Smith & Taffler, 1992a&b; Jones, Smith, & Whale, 2010)	"CLOZE" poorly related with understandability (Jones, Smith, & Whale, 2010)
Linguistic analysis	• de Beaugrande and Dressler (1981) • Roseberry (1995) • Sydserff and Weetman (1999) • Pennebaker (2002) • Pennebaker, Mehl, and Niederhoffer (2003) • Tausczik and Pennebaker (2010)	• Unit-by-unit analysis • Sound theory • Takes reader into consideration (Sydserff & Weetman, 1999) • Get cues about the writers' thought, emotion, motivation, and need by simply counting words. (Pennebaker, 2002)	• Time consuming • Not attractive enough to replace readability formula (Sydserff & Weetman, 1999; Beattie, et al., 2004)

Chapter Three: Theories

Merkl-Davies and Brennan (2007) identified there are five theories which provide a theoretical perspective for in this research area: agency theory, signalling theory, legitimacy theory, stakeholder theory, and institutional theory. In their research, they described each theory, and discussed the characteristic of each theory. Following their discussion, two theories, agency theory and signalling theory will be used in this thesis. Signalling theory is used to focus on good performing companies; in contrast, agency theory is used to focus on poorly performing companies. The selection reasons and differences with the other three theories are listed as follows (Merkl-Davies & Brennan, 2007):

- This thesis assumes outside investors are users of narrative disclosures, which is consistent with the characteristics of both agency theory and signalling theory;
- This thesis focuses on corporate financial performance, instead of their social or environment performance;
- This thesis focuses on impression management as an every-day occurrence, while the other three theories are often used under a non-routine reporting context;
- The sample of this thesis is selected from the population of listed Australia manufacturing companies, and for a large sample size study, agency theory and signalling theory are more prevalent;
- This thesis adopts a content analysis method, while the other three theories are commonly used in case studies.

3.1 Signalling theory

Merkl-Davies and Brennan (2007) summarised that this theory "focuses on the behaviours of managers in well-performing companies who signal this superiority by greater transparency in their disclosures and presentation of information" (p. 124).

Ross's (1977) examination of capital markets found that the good performing companies tended to disclose more information. Meanwhile, this tendency forces other companies in the same industry to provide more information in order to maintain their credibility in the capital market (Ross, 1977; Smith, Jamil, Johari & Ahmad, 2006). This situation was explained as signalling theory in that if the company does not disclose its information, the public would assume that the company was too negative to make disclosures. Under signalling theory, Watts and Zimmerman (1986) developed a "signalling hypothesis" which states that the corporate good performance would encourage management to make more disclosures, which indicates that the narrative disclosures can reflect corporate performance.

Grounded in signalling theory, corporate performance is not only related with the quantity of disclose, but also related with the quality of disclosure. This was shown by the research of Smith and Taffler (1992a). They developed their hypotheses based on signalling theory and stated that better corporate performance is positively associated with readability level and understandability level.

3.2 Agency theory

Modern companies delegate decision making from one party (the principal) to another party (the agent), which is characterised as an agency relationship (Deegan, 2006).

Under this relationship, managements are motivated by compensation and the provision of wealth in their choice of policies (Watts & Zimmerman, 1986, 1990) and behave in a self-interested way. Since managements view an annual report as a reflection of their managerial performance which is also a source of information that is utilised by interested parties outside the companies (Prakash & Rappaport, 1977), managements may be encouraged to overstate the positive information and understate the negative information. Aerts (2005) also stated that under agency theory the importance of narrative disclosures is "not only as a commodity that can be traded in principal–agent relationships, but as a context-sensitive communication device with symbolic as well as intrinsic substance" (p. 515).

The modern agency relationship could also result in information asymmetry between shareholders and company managements. As mentioned in Chapter One, information asymmetry may arise because outside information users lack sufficient information to make correct predictions.

For most outside information users, the annual report is the main source on which to base decision making, while the narrative disclosures are discretionary, corporate managements can decide the content (what) and way (how) of disclosure. Therefore, the narrative section of the corporate annual report, to some extent, is not just an objective description of corporate performance to shareholders, but also a communication medium to let corporate managements adopt their manipulation strategy (Bowman, 1984).

This conscious and deliberate managerial strategy is called "impression management behaviour" (Bowen, Davis, & Matsumoto, 2005), which is rooted in agency theory. It is explained by agency theory that managements of companies act opportunistically to choose the disclosure style and content that are beneficial to them (Merkl-Davies & Brennan, 2007).

3.3 Impression management behaviour strategy

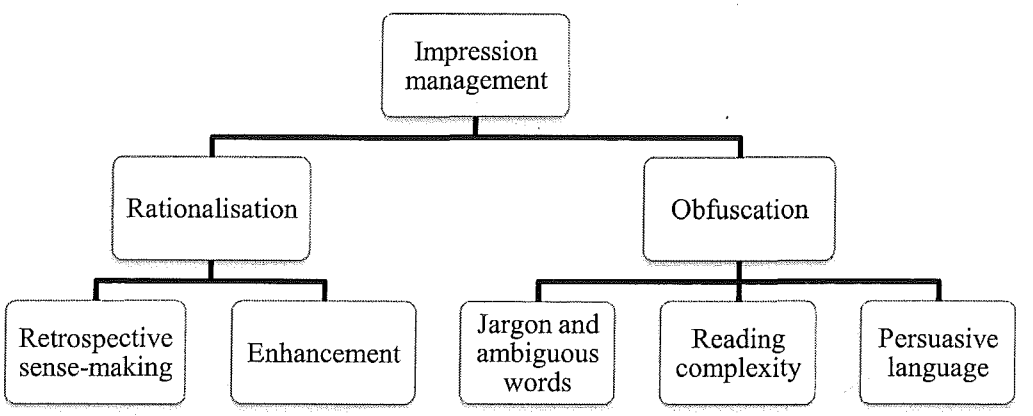
Impression management is a social bias which involves "controlling or manipulating the attributions or impressions" (Tedeschi & Riess, 1981, p. 3). In the accounting disclosure context, it is defined as "control and manipulate the impression conveyed to users of accounting information" (Clatworthy & Jones, 2001, p. 311).

Under these circumstances, managements tend to use narrative disclosure information as a marketing tool to present a self-interested view of corporate performance (Subramanian, Insley, & Blackwell, 1993), and to distort outside information users' evaluations and perceptions of corporate performance (Neu, Warsame, & Pedwell, 1998). For example, managements may enhance or overstate positive information, and

meanwhile legitimise and understate negative information, or even try to mask and hide bad news opportunistically (Courtis, 1998).

These impression management strategy choices are summarised by Smith et al. (2006) as presented in Figure 4. It consists of two techniques: rationalization and obfuscation. Brennan et al. (2009) stated companies are motivated to disclose more information about the financial performance, while disclosing in an unambiguous manner about negative information. This result indicates that companies have a tendency to increase the quantity of disclosures but with lower quality. The remainder of this chapter will discuss this tendency according to two impression management strategies, along with research hypotheses.

Figure 4 Approaches of impression management



3.3.1 Rationalisation

As detailed in Figure 4, rationalisation generally involves two impression management strategy approaches: one is regarded as "retrospective sense-making", also called "attributions" (Merkl-Davies & Brennan, 2007), which attempts to legitimise events and outcomes (Smith et al., 2006); another approach is putting undue emphasis on positive information which is known as enhancement (Smith et al., 2006).

Retrospective sense-making involves interpreting negative actions or performance that have already occurred (Aerts, 2005). It intends to control the feedback of reported

information by giving explanation and legitimacy in order to counteract undesirable consequences. Examples of using this approach are to give excuses, justifications and apologies in the annual report narrative sections (Aerts, 1994).

In terms of the enhancement approach, managements of companies have an incentive to repeat or highlight positive actions or performance for two purposes: on the one hand to enhance the corporate positive image to information users. Ahmed and Courtis (1999) stated, profitable companies tend to disclose more information about their good performance to outside information users. By disclosing more positive information, corporate competitive ability could be advantaged (Singhvi, 1972). On the other hand, uses enhancement to draw information users' attention away from negative information in order to emphasise the positive position. This latter purpose is more commonly adopted by poorly performing companies.

To sum up, managements may adopt both retrospective sense-making and enhancement approaches to either positive or negative performance (Aerts, 2005). Whether the approaches "function in an assertive or in a defensive way depends on their content and its relationship to salient performance characteristics" (Aerts, 2005, p. 515). Moreover, Aerts (2005) found that rationalising positive performance can in turn improve the explanations of reliability in terms of negative performance.

According to the prior literature on impression management, there are seven techniques in this filed (Merkl-Davies & Brennan, 2007). In relation to rationalization, this study chooses the thematic content analysis technique to expand investigation and measurement. Both word variables and theme variables will be used in this study. Thus, the first two hypotheses are:

H_{1a}: Theme-based variables, in the Chairman's Statement, are significantly associated with corporate performance.

H_{1b}: Word-based variables, in the Chairman's Statement, are significantly associated with corporate performance.

As the rationalisation approach states, companies with either good performance or poor performance all tend to provide more information, which means that all will have similar report sizes. There is no significant difference in terms of the disclosure quantity under two performing characteristics. The following hypothesis is therefore developed:

H_{1c}: Report size of the Chairman's Statement is not significantly associated with corporate performance.

Rationalisation is an impression management approach of increasing disclosure quantity, while obfuscation involves reducing disclosure quality to conceal negative information. This approach takes the form of either concealing or distorting the information that is inconsistent with corporate self-concept. The details of this approach will be discussed as follows.

3.3.2 Obfuscation

Obfuscation, also known as self-presentational dissimulation indicates "concealing or disguising events, or trying to minimise their importance" (Smith, et al., 2006, p. 49). This approach involves the manipulation of information for users by increasing the reading complexities of the annual report.

There is an extensive literature that has examined the reading ease level of narrative disclosures over several decades and across many countries (such as US, UK, New Zealand, and Australia). The research demonstrates that the narrative disclosure sections are too difficult for most readers (Smith & Taffler, 1992a, 1992b). Moreover, there is no sign that this tendency had been improved between 1986 and 1991: still 90% of adults found the narrative disclosures are too complex to understand (Courtis, 1995). Three techniques that companies may adopt to increase the reading difficulties are introduced below (Merkl-Davies & Brennan, 2007).

Firstly, some companies prevent readers from gaining an accurate understanding of corporate reality by putting unnecessary jargon in annual report, which is termed "accounting bias" by Aerts (1994). Smith and Taffler (2000) mentioned that managements used "technical accounting terms to obscure the underlying excuses and justifications for negative outcomes and to avoid associated managerial responsibility" (p. 626). Meanwhile, managements may use some ambiguous words to confuse information users. Normally, both jargon and ambiguous words are those "big words" with more than six letters. Thus, this study combines jargon and ambiguous words as "big words".

Secondly, managements of poorly performing companies may use a skilfully crafted writing style to make texts more complex to read and understand, in order to distract readers from gaining a clear understanding of corporate performance (Courtis, 2004).

Thirdly, Merkl-Davies and Brennan (2007) also stated that persuasive language is another technique that managements used to deceive outside information users. Personal pronouns (both 1st personal pronoun and 2nd personal pronoun) and emotional words (both positive and negative) could be used to reflect the use of persuasive language.

Obfuscation is the fundamental theory of syntactic study. This study regards readability as a proxy for obfuscation measurement. Smith and Taffler (1992a) used the FLESCH readability score to assess the quality of disclosure information, and found that there is a positive relationship between the readability of narrative sections in the annual report and corporate performance. This result was named as "obfuscation hypothesis" by Courtis (1998), and stated that the clarity of narrative disclosures in the annual report is positively associated with corporate performance. Based on the prior research and three techniques as mentioned by Merkl-Davies and Brennan (2007), three hypotheses are developed as follows:

H_{2a}: The number of "Big words" (>6 letters) in the Chairman's Statement, is significantly associated with corporate performance.

H_{2b}: The readability level of the Chairman's Statement is significantly associated with corporate performance.

H_{2c}: The use of persuasive language in the Chairman's Statement is significantly associated with corporate performance.

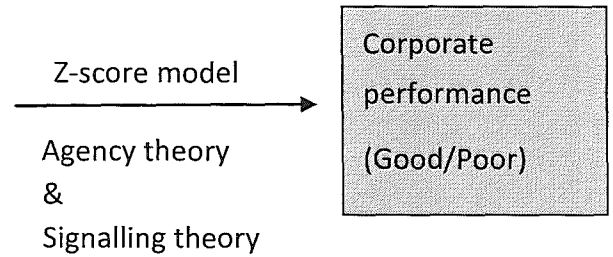
3.4 Research framework

Figure 5 outlines the framework of this research. This research investigates the relationship between narrative disclosures and corporate performance. Agency theory and signalling theory are the two fundamental theories that underpin this research. An impression management strategy guides this research: rationalisation focuses on the quantity respect of narrative information (what to disclose). This research adopts thematic content analysis manually, addresses three hypotheses by three variables (words variable, theme variable, and report size). In terms of obfuscation, it focuses on quality respect of narrative information (how to disclose). Each variable (big words, FLESCH score, an aggregation of personal pronouns and emotional words) links to the three subsets of obfuscation strategy, and addresses three hypotheses respectively.

In summary, this research focuses on investigation of relationship between corporate performance (good/poor) and narrative information (quantity and quality) in the annual report. It is grounded on agency theory and signalling theory, and hypotheses are developed based on impression management strategy. The detailed research design will be discussed in Chapter Four.

Figure 5 Research framework

Narrative information	Impression management	Hypothesis	Approach	Variables
Quantity respect	Rationalisation			
	- Retrospective sense-making	H_{1a}	Thematic content anlaysis (manual)	Word variables
		H_{1b}		Theme variables
	- Enhancement	H_{1c}		Report size
Quality respect	Obfuscation			
	- Jargon and ambiguous words	H_{2a}	Syntactic content analysis (WORD, LIWC)	Big words
	- Reading complexity	H_{2b}		FLESCH score
	- Persuasive language	H_{2c}		Personal pronouns & emotional words



Chapter Four: Research Method

After developing six hypotheses, this chapter explains the overall research methodology adopted in this study. Initially, this study indentifies 64 Australian manufacturing companies based on their 2009 performances, and then builds a predictive classification model based on their 2008 data. In this study, the Chairman's Statement from the annual report is collected as research data to test whether the discretionary narrative disclosures are potentially decision-useful for predicting subsequent corporate performance. Generally, this research collects both quantitative and qualitative secondary data, and adopts both thematic and syntactic content analysis techniques.

As introduced in Chapter One, there are three data units in content analysis: sample unit, context unit, and recording unit. To begin with, this chapter will describe the selected process for each unit, especially the focus on the recording unit, as this unit is collected as an independent variable in this study. After this, the statistical analysis techniques used to measure the dependent variables, are discussed. The final section of this chapter outlines summarisation and evaluation of the research methods.

4.1 Selection of sample unit – annual report

As mentioned in the literature review chapter (Chapter Two), there is no published predictive model directly applied to performance of Australian manufacturing companies. As disclosure regulations vary between different countries, it is necessary to develop a predictive classification model based only on Australian evidence. Four main criteria for company selection of this study are illustrated below.

Firstly, all sample companies are chosen from those companies listed on the Australian Stock Exchange (ASX). The main explanation of this constraint is that listed companies are large enough to provide the most easily accessible and reliable information (Epstein

& Freedman, 1994). Staw, McKechnie, and Puffer (1983) recommended "sample companies should be large enough so that annual reports were readily available" (p. 587). Moreover, as Aerts (2005) confirmed, listed companies tend to offer more explanations which can help with an impression management study.

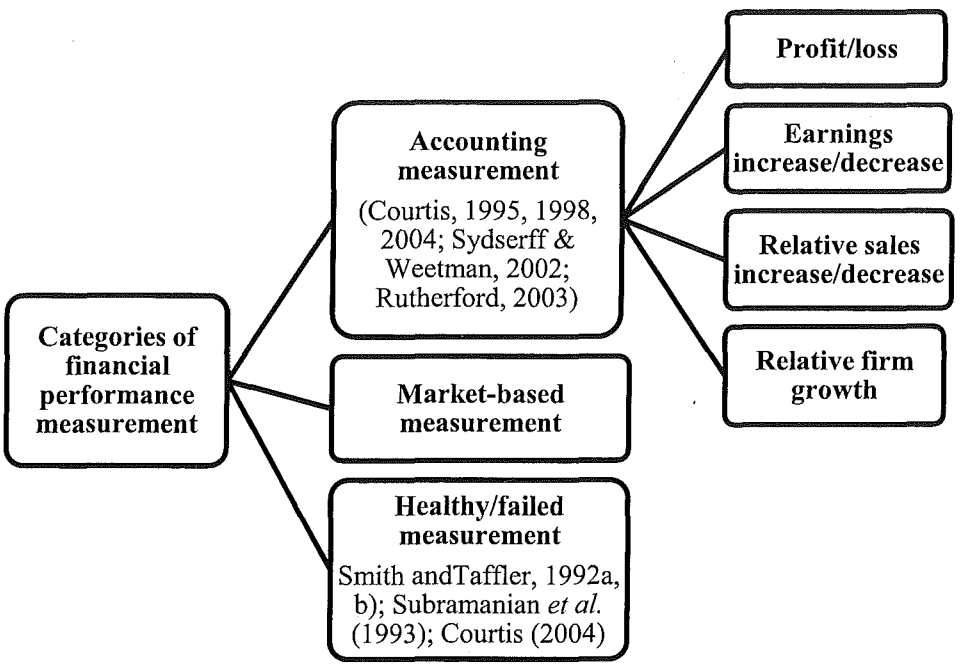
Secondly, the corporate performance of only two fiscal years are taken into consideration, and a single year (2008 fiscal year) of annual report is collected in this study for data analysis purpose. Single year study can eliminate the potential influences of both changes in reporting regulations over time and other economic movements, such as the 2008 Global Financial Crisis. Furthermore, as the end of Australian fiscal year is 30th June, and with a time lag for submitting to the relevant authorities, the 2008 annual report is the most readily available and up to date sampling unit for this study.

Thirdly, the sample for this study is drawn from the Australian manufacturing industry. The Australian and New Zealand Standard Industrial Classification (ANZSIC) defines manufacturing as "the physical or chemical transformation of materials or components into new products, whether the work is performed by machinery or by hand" (ANZSIC, 2010).

The last and also the most vital constraint of this study is the corporate financial performance. Staw et al. (1983) recommended a balanced distribution of high- and low-performing companies should be sought. Thus, this study classifies sample companies into two categories: good performing companies, and poor performing companies. The different levels of performance in these companies are the dependent variables of this study. There are three categories of financial performance measurements that have been used in prior research: accounting measurement, market-based measurement, and healthy/failed measurement (Figure 6). Although most prior researchers adopted a healthy/failed measurement, there is still a lack of accurate measurements for prediction studies of financial performance.

To fill the research gap, this study adopts the accounting measurement by distinguishing companies according to their earnings increase/decrease. It follows the same classification measurement as Staw et al. (1983): the good performing companies are those with an increase of 50 percent or more in regular earnings, and poorly performing companies are those suffering a decrease of at least 50 percent of their earning per share (EPS).

Figure 6 Category of financial performance measures



Based on the four criteria, 64 companies are selected, 29 with good performance and 35 with poor performance. To gain access to corporate narrative disclosures, all 64 sample corporate 2008 annual reports are obtained by downloading from the Morningstar database which covers almost all Australian listed companies.

4.2 Selection of context unit – Chairman's statement

This study focuses on corporate narrative disclosures, and the Chairman's Statement section of annual report is the main context unit, which is for the following four reasons:

- Many researchers have proved that the Chairman's Statement is a reliable and tested medium for narrative study (Smith & Taffler, 1992a, 1992b, 1995; Courtis, 1998; Smith & Taffler, 2000; Clatworthy & Jones, 2001; Sydserff & Weetman, 2002; Courtis, 2004).
- The Chairman's Statement is the first part in most corporate annual reports. It is important because this is the first impression, created by the annual report, on outside information users (Smith & Taffler, 2000), and it provides a general statement that reflects the corporate performance in the current year.
- The Chairman's Statement varies from 300 to 3000 words. In general, the size range for a narrative disclosure is relatively short and suitable for content analysis and narrative study.
- The significance of a Chairman's Statement study is addressed by some researchers. Smith and Taffler (1995) stated that "accounting researchers have largely neglected the content of firms' discretionary unaudited disclosures in the annual report despite the demonstrable utility of the Chairman's Statement to users" (p. 1195).

Based on these reasons, this research uses the Chairman's Statement section as an indication of narrative disclosures. There are many different names describing the Chairman's Statement, such as President's Letter, Letter from the Chairman. For two companies (Waterco Limited, Autodom Limited) whose annual reports do not include a Chairman's Statement, "Chief Executive Officer's Review of Operations" (CEO's Review), and "Managing Directors' Report and Review of Operations" are collected as a context unit for the narrative study respectively. Selected narrative section of each company and the EPS movement with the financial characteristics are listed in Table 2. Among the 64 sample companies in the table, 14 of them have neither a Chairman's Statement nor a CEO's Review. Thus, all variables are count as missing data among these companies.

Table 2 Sample company list

Company name	Section name	2008 EPS	2009 EPS	EPS change	Corporate performance
1. Berklee Limited	~~Director Report~~	0.2	13	6400.00 %	Good
2. Sirtex Medical Limited	Chairman's Report	2.2	26	1082.82%	Good
3. Adelaide Resources Limited	Chairman's Letter	-1.1	4.6	518.18%	Good
4. Quantum Energy Limited	Chairman's Report to Shareholders	-0.8	3	475.00%	Good
5. Codan Limited	Chairman's Report	1.3	7.4	469.23%	Good
6. Mesbon Chinal Nylon Limited	Chairman's Report and Review of Operations	0.4	2	400.00%	Good
7. TMA Group of Companies Limited	~~Director's Report~~	0.06	0.25	316.67%	Good
8. Cellestis Limited	~~Director Report~~	2.3	8.4	265.22%	Good
9. Compumedics Limited	~~Director Report~~	0.5	1.7	240.00%	Good
10. SDI Limited	Chairman and Managing Director's Report	0.9	2.6	188.89%	Good
11. Lemarne Corporation Limited	Chairman & Managing Director's Report	33.1	88.2	166.47%	Good
12. Waterco Limited	*Chief Executive Officer's Review of Operations*	-12.9	8.1	162.79%	Good
13. China West International Holdings LTD	Chairman's Report	1.7	4	135.29%	Good
14. Antaria Limited	Chairman's Report	-2.5	0.8	132.00%	Good
15. Autodom Limited	*Managing Directors' Report and Review of Operations*	-12.4	2.3	118.55%	Good
16. Universal Biosensors, Inc.	Chairman's Letter & CEO's Report	-7.6	0.9	111.84%	Good
17. Bisalloy Steel Group Limited	Chairman and Managing Director's Review	-61.3	-3	95.11%	Good
18. CMI Limited	Chairman's Review	-73.6	-4.4	94.02%	Good
19. Capral Limited	Chairman's Report	-381	-27.5	92.78%	Good
20. Maryborough Sugar Factory	Chairman's Overview	-22.3	-1.7	92.38%	Good

Limited					
21. Phosphagenics Limited	Chairman and CEO's Joint Report	-14.2	-1.2	91.55%	Good
22. Probiomics Limited	~~Director's Report~~	-0.8	-0.1	87.50%	Good
23. NuSep Ltd	~~Director's Report~~	-557	-126	77.38%	Good
24. Frankland River Olive company Limited	~~Director's Report~~	-15.2	-3.8	75.00%	Good
25. AWH Corporation Limited	Chairman's Letter	-1	-0.3	70.00%	Good
26. Austofix Group Limited	~~Director's Report~~	-18.5	-6.3	65.95%	Good
27. Sterling Biofuels International Limited	Letter From Chairman	-19.1	-7.7	59.69%	Good
28. AtCor Medical Holdings Limited	Chairman's Letter to Shareholder	-3.8	-1.7	55.26%	Good
29. USCOM Ltd	Letter from the Chairman	-5.7	-2.8	50.87%	Good
30. Nuplex Industries Limited	Chairman's Report	79.4	37	-53.40%	Poor
31. UnderCoverWear Limited	Chairman's Overview	7.6	3.4	-55.26%	Poor
32. Fisher & Paykel Appliances Holdings Limited	Chairman's Review	15.2	6.7	-55.92%	Poor
33. Anek Tambang (Persero) Tbk (Pt)	~~Director Report~~	1.9	0.8	-57.89%	Poor
34. Incitec Pivot Limited	Chairman's Report	54.7	22.6	-58.68%	Poor
35. Imdex Limited	Chairman's Report	16.4	6.2	-62.20%	Poor
36. Brand New Vintage Limited	~~Director's Report~~	0.3	0.1	-66.67%	Poor
37. Farm Pride Foods Limited	Chairman's and Chief Executive Officer's Report	6.1	2	-67.21%	Poor
38. Maxitrans Industries Limited	Chairman's and Managing Director's Review	9.4	3	-68.09%	Poor
39. Dexion Limited	~~Director's Report~~	11.8	3.2	-72.88%	Poor
40. Style Limited	Message from the Chairman	-7.4	-13.1	-77.03%	Poor
41. Watty Limited	Chairman's Report	13.7	1.8	-86.86%	Poor
42. BlueScope Steel Limited	Chairman's Message	63.9	6	-90.61%	Poor
43. Advanced Braking Technology Ltd	Chairman's Letter	-0.2	-0.4	-100.00%	Poor

44. Coventry Group Limited	Executive Chairman's Report	18.7	-0.5	-102.67%	Poor
45. Ellex Medical Lasers Limited	~~Director's Report~~	6.8	-29.6	-535.29%	Poor
46. ITL Limited	Message from Chairman and Chief Executive Officer	1.1	-0.1	-109.09%	Poor
47. Beyond Sportswear International Limited	Chairman's Report	1.8	-0.2	-111.11%	Poor
48. James Hardie Industries SE	Chairman's Report	45.6	-21.4	-146.93%	Poor
49. TWT Group Limited	Chairman's Report	12.6	-6.2	-149.21%	Poor
50. Buderim Ginger Limited	Chairman's Review	6.5	-4.2	-164.62%	Poor
51. Warrnambool Cheese and Butter Factory	Chairman's Report	55.1	-50.2	-191.11%	Poor
52. Vmoto Limited	~~Director's Report~~	-1.3	-3.9	-200.00%	Poor
53. Ridley Corporation Limited	Chairman's Review	8.6	-10.6	-223.26%	Poor
54. SciGen Limited	Chairman Review	-2.1	-7.8	-271.43%	Poor
55. Oriental Technologies Investment Limited	Chairman's Report	0.3	-0.6	-300.00%	Poor
56. Gale Pacific Limited	Report from the Chairman & the Managing Director and Chief Executive Officer	1.6	-5	-412.50%	Poor
57. Refresh group Limited	Chairman Review	-0.2	-1.1	-450.00%	Poor
58. DataDot Technology Limited	Chairman's Letter	0.4	-1.7	-525.00%	Poor
59. TSV Holdings Limited	Chairman's Report	2.2	-9.6	-536.36%	Poor
60. Advanced Surgical Design & Manufacture Limited	Chairman's Letter	0.5	-2.9	-680.00%	Poor
61. Atlas South Sea Pearl Limited	Chairman's Report	0.6	-6.2	-1133.33%	Poor
62. Lazco Limited	~~Director Report~~	-1.2	-19	-1483.33%	Poor
63. Paperlinx Limited	Chairman's Report	9.1	-132.9	-1560.44%	Poor
64. Garratt's Limited	Chairman's Report	19	4.3	-7730.00%	Poor

4.3 Selection of recording data – independent variables

This study adopts both thematic content analysis and syntactic content analysis to investigate two approaches to impression management. To test each hypothesis, data for 923 independent variables are collected from the Chairman's Statement. Generally, these data can be categorised into seven main groups: word-based variables; theme-based variables; report size; big words; FLESCH readability score; personal pronouns; and emotional words. In this section, each variable group will be introduced based on three data collection approaches: manual, WORD, Linguistic Inquiry and Word Count (LIWC) software. As Table 3 presents, three variables (word-based and theme-based variables, personal pronouns) are collected manually, report size and FLESCH sore are collected by WORD; while big words and emotional words are counted by LIWC software.

Table 3 Variable collection approach

Data collection appraoch		Variables
Manually	Thematic content analysis	word-based variable
	Thematic content analysis	theme-based variable
	Syntactic content analysis	personal pronoun
WORD	Thematic content analysis	report size
	Syntactic content analysis	FLESCH readability score
LIWC	Syntactic content analysis	big words
	Syntactic content analysis	emotional words

Before collecting data, the Chairman's Statement reports all tables, charts, photographs, and forms of address (Dear shareholder), and greeting (Yours sincerely) are deleted, and the resulting text is pasted into a Word document to prepare for future data collection. For several companies (e.g., Autodom Limited and Watty Limited), outside of the main paragraphs, they have additional text in the margin or under photographs. Since these sentences are added to emphasise important narrative information, they also count as separate sentences and are processed for data collection.

4.3.1 Manually collected data

The Chairman's Statement of each company is checked for data collection purpose; both word-based and theme-based recording units are used in measuring the content of narrative. In this stage, data are collected under the guidance of thematic content analysis to evaluate narrative information. As mentioned in Chapter Two, thematic content analysis is a research method that draws inferences from data by systematically identifying characteristics within the data (Jones & Shoemaker, 1994); the recording units have to be categorised, and the frequency of each category is counted before the thematic content results be generated.

As with the performance prediction study conducted by Smith and Taffler (2000), this thesis adopts Houghton's (1988) four-factor cognitive structure as the classification standard of sorting narrative content into themes. The details are listed in Table 4 (Smith & Gunalan, 1996, p. 76). This structure was developed by adding a further dimension to Osgood and his colleagues' (Osgood, Suci, & Tannenbaum, 1957) three dimensions classification, to allow measurement of the connotative meaning.

Table 4 Houghton's (1988) four-factor cognitive structure

Category	**Theme Classification**	**Company Performance**	**Evaluation**
Evaluative	Beneficial Adverse	Positive Negative	Providing details of good news Providing details of bad news
Potency	Tangible intangible	Certainty Vagueness	Degree of certainty about future Vagueness about the past or present
Activity	Dynamic Static	Performance Reluctance	Reference to measures of past performance Reluctance to take action
Manageability	Expected Unexpected	Status Quo External	Emphasis on maintaining the status quo Dependence on external economic factors

After reliable classification, "constructs in content analysis are operationalizing with a coding system, a set of instruments or rules on how to systematically observe and

record content from text" (Neuman, 2006, p. 324). It is the reason why some researchers described content analysis as "textual coding". Both Weber (1985) and Boyatzis (1998) provided basic steps to develop and test the coding scheme. Based on their research, Beattie et al. (2004) summarised the process as follows:

1. Define the recording unit (both word-based and theme-based);
2. Define the categories (based on Houghton's cognitive structure);
3. Test coding of a sample of text;
4. Assess reliability;
5. Revise coding rules;
6. Repeat steps 3-5 until reliability is satisfactory;
7. Code all text; and
8. Assess achieved reliability

Word variables – form oriented

Firstly, this research examines the relationship between individual word and corporate financial performance. Some softwares and specific dictionary are used to collect keywords. In the prior research, Smith and Taffler (2000) adopted a combination of Oxford Concordance Program (Hockey & Martin, 1988) and computer software with SPSS-X (SPSS, 1986) to sort each narrative word in an alphabetic order; while Tennyson et al. (1990) adopted WORD package to investigate the statistical relationship between words and narrative disclosures. The merits of these "software plus dictionary" approaches are that they avoid subjective judgments made by the researcher. The data collection process is more objective if research bias is reduced (Tennyson et al., 1990). However, these computer-based data collection methods cannot distinguish the different thematic meanings of the same word in a different context.

For example, the following three sentences all have the word "high"; however, they differ significantly in meaning. As in the first sentence, "high" indicates an unexpected

and adverse theme; and in second sentence "high" provides a beneficial theme, while the word "high" is just part of a trading name in the third sentence. In thematic content analysis, the word-based variable actually means the "word in context", for this reason the context meaning of each word should be taken into consideration. For acquiring a more accurate data classification, this study collecting both word-based and theme-based data manually in the thematic content analysis procedure. Although the manual collection is time-costly, since the sample is relatively small, this collection design is reasonable and possible to accomplish.

- *"Although total new vehicle sales in this market have declined sharply in recent months, vehicle theft remains unacceptably high" (DatadotData Ltd, 2008).*
- *"Product sales in the second half of 2008 were 27 per cent higher than the first half" (DatadotData Ltd, 2008).*
- *"In High Security Solutions, agreement has been reached with Gopsons, the largest security printer in India, to make DataTraceDNA their exclusive forensic tracer" (DatadotData Ltd, 2008).*

Weber suggested that "'word' was taken to indicate semantically equivalent textual units, including word synonyms, idioms and phrases" (Weber, 1990, p. 22). A customised keyword dictionary is compiled during the word-based data collection procedure, and is set out in the Appendix A. After eliminating the function words that do not affect textual content, the remaining words are allocated to Houghton's (1988) eight categories based on their context meanings. The listings of eight categories compose the customised keyword dictionary, which helps the data classification to be both consistent and organised.

After data collection, the keyword variable can be calculated based on the formula that Smith and Taffler (2000) defined:

$$\text{Word variable} = \frac{\text{Number of common occurrences}}{\text{Total number of words in the narrative}}$$

Theme variables – meaning oriented

Following the same procedure that Smith and Taffler (2000) used in their narrative study, each sentence is grouped to the eight categories of Houghton's (1988) cognitive structure. If "a sentence comprises several separable themes then the theme score unit is subdivided to register the relative importance of those themes in the narrative without weighting" (Smith & Taffler, 2000, p. 632), and each sentence is assigned a theme score of 1. For consistency, only a completed sentence with a full stop is regarded as a sentence in this study, and the groups of phrases linked by semicolons are counted as one sentence.

As cited in Smith and Taffler (2000), the formula of sentence-based thematic content analysis is:

$$\text{Theme variable} = \frac{\text{Sum of theme scores}}{\text{Total number of sentences}}$$

Personal pronouns

Recent research has suggested that the personal pronoun, a form of persuasive language, can be used as an impression management technique to manipulate information users (Merkl-Davies & Brennan, 2007). To test this hypothesis, data for both first personal pronoun and second personal pronoun are collected. LIWC software can only count first personal pronoun including "I", "me", and "my". For more accuracy, this study counts both first pronoun and second pronoun manually. The amount of first pronoun (1st_PRON) is the sum of "I", "me", "we", "our", and "us"; while the total amount of second pronoun (2nd_PRON) is the sum of "you" and "your". All the data, includes seven individual pronouns, first pronoun, second pronoun, and

total pronoun (PRON_TL - the sum of first and second pronoun), are entered into SPSS as independent variables.

4.3.2 WORD collected data

FLESCH Readability score

There are two readability formulas that are widely used as measurements of text readability level, which are presented below (Smith & Taffler, 1992a). The two readability formulas are all based on word length (W), and sentence length (S), while using various different weightings are applied to the component parts. High levels of readability are associated with low LIX scores but high FLESCH scores. Moreover, both formulas are potentially flawed in that their measures are independent of the intended audience.

FLESCH Readability Formula:

FLESCH = 206.385 – 0.846W – 1.015S

Where W = Word length = number of syllables per 100 words;

S = sentence length = total number of words/total number of sentences

LIX Readability Formula:

LIX = S + W

Where S = average number of words per sentence;

W = % of words of seven or more letters

As the FLESCH readability score can be calculated automatically by WORD document, this research chooses FLESCH as the independent variable ("readability") to reflect the level of complexity of the Chairman's Statement.

Report size

In this research, report size is measured by total words in Chairman's Statement which is counted by WORD automatically. It is an independent variable (Report_size) that reflects the quantity of the Chairman's Statement.

4.3.3 LIWC software collected data

Linguistic Inquiry and Word Count (LIWC) software is used in this study to collect some syntactic content analysis data, which includes: "big words" and emotional words.

"Big words"

As mentioned in Chapter Three, poorly performing companies may use unnecessary jargon or ambiguous words to conceal negative performance to outside information users. Since most jargon and ambiguous words are big words which are longer than six letters, this study uses LIWC software to count "big words" as an independent variable (BIG_WORDS) to measure the amount of jargon and potentially ambiguous words.

Emotional words

Emotional words, together with the personal pronouns, are two kinds of indicators to reflect the use of persuasive language. As mentioned before, personal pronouns are counted manually for a more accurate result; while emotional words are counted based on LIWC software. LIWC software can give out the percentage of both positive (EM_positive) and negative (EM_negative) emotional words. By adding the two word-percentages together, a new variable "total emotional words" (Em_TL) is created that is used in this study. Meanwhile, both positive and negative emotional words are also evaluated separately prior to aggregation.

4.4 Statistical analysis

Once these independent variables are coded, inferences must be drawn. This process requires the use of statistical data analysis (measurement models) to form associations for inferential conclusions (Jones & Shoemaker, 1994). SPSS software is used at this stage to help analyse data. It is a comprehensive software package that is used for managing quantitative data and performing statistical analysis. This research uses the z-score for data analysis, and the discriminant function is of the form (Smith & Taffler, 2000, p. 633):

$$Z = d_0 + d_1v_1 + d_2v_2 + d_3v_3 + \cdots$$

where Z is the discriminant score,

$\{v_j\}$ are the variables,

$\{d_j\}$ are the optimal coefficients with d_0,

$\{d_0\}$ the constant term, representing the cut-off criterion between the two groups.

By now, the predictive model has been built, based on Fisher's linear discriminant function, and follows Krippendorff (1980), both validity and reliability are also checked by a co-investigator.

4.5 Conclusions

This study focuses on the association between the Chairman's Statement and corporate performance, to assess the predictive ability of corporate narrative disclosures. The annual report is the sample unit in this study, and the Chairman's Statement is the context unit, all the independent variables are the recoding data. Z-score is adopted in this study for statistical analysis.

Analysis and the selection of thematic content analysis independent variables to be used in this study follow the Smith and Taffler (2000) schema. However, instead of using the Oxford Concordance Program (OCP) to select keywords, this study created a

customised keyword dictionary to manually categorise each word into eight groups (beneficial/adverse, tangible/intangible, dynamic/static, expected/unexpected).

Furthermore, this study adds more variables in the syntactic content analysis respect. All variables used in this study are: word-based thematic analysis data, theme-based thematic analysis data, personal pronouns, FLESCH readability score, report size, big words, and emotional words. Except for the FLESCH readability score and report size, all other variables are expressed as percentages which are weighted by total word number to eliminate the effects of different narrative report size. There are two computer-based softwares used in this study: Linguistic Inquiry and Word Count (LIWC) software and WORD.

This study manually collected word-based variables which could increase the classification accuracy, but also cause the risk of researcher bias. Although inter-temporal coding and independent checks are used, the bias is unavoidable. Further minimising of bias is required in future studies. Moreover, this study involves 64 Australian manufacturing companies which is a small sample size, and does not take company size and type of industry into consideration.

Chapter Five: Results

5.1 Descriptive analysis

All of the correlation coefficients for each word variable with corporate performance are listed in Appendix B following a descending order correlation. Since there are too many variables (923 in total) to report the inter-correlation, only correlations with performance are included.

Total of 28 words were selected due to their significant correlations with corporate performance, and were used as the primary available variables for building a word-based classification model. The correlation of the 28 words with performance and their inter-correlations are listed in Table 5.

5.2 Multivariate models

5.2.1 Variable selection

Starting from the word with highest correlation, each word was sequentially entered into SPSS software to see whether it can increase the classification accuracy. The word variable would be added into the model if it increased the classification accuracy; and it would be eliminated if the word did not make any contribution. Following the same procedure; new variables were added to the exiting model until the classification accuracy could not be increased. During this procedure, seven words were selected to build the word-based predictive model, and the classification accuracy was 90%. The seven selected words and the correlation coefficients between them are listed in Table 6.

Table 5 The correlation coefficient of 28 significant correlated variables

Correlations - Spearman's rho

Variable	Symbol		PC	HIGH_TL	HIGH_BEN	DIV	GW	CONSLD	EBT	LOW_TL	FIN	SV_INTAN	GAIN	RET_SUCC	RET_TAN	OG	VOL	INC_ADV	INC_TL	DIRN	NOT_ACH	WROF F	VAL_DYN	IMP	FRT	COST_TAN	OFFSET	DESIGN	HW	LOW_BEN	ND
performance characteristics	PC	Correlation Coefficient	1.000																												
		Sig. (2-tailed)	.																												
		N	64																												
High (Total)	HIGH_TL	Correlation Coefficient	-.633**	1.000																											
		Sig. (2-tailed)	.000	.																											
		N	50	50																											
High (Beneficial)	HIGH_BEN	Correlation Coefficient	-.624**	.897**	1.000																										
		Sig. (2-tailed)	.000	.000	.																										
		N	50	50	50																										
Dividend	DIV	Correlation Coefficient	-.432**	.380**	.378**	1.000																									
		Sig. (2-tailed)	.002	.006	.007	.																									
		N	50	50	50	50																									
Goodwill	GW	Correlation Coefficient	.391**	-.228	-.228	-.234	1.000																								
		Sig. (2-tailed)	.005	.112	.112	.102	.																								
		N	50	50	50	50	50																								
Consolidate, consolidation	CONSLD	Correlation Coefficient	.366**	-.193	-.185	.042	.028	1.000																							
		Sig. (2-tailed)	.009	.179	.198	.772	.849	.																							
		N	50	50	50	50	50	50																							
Earnings	EBT	Correlation Coefficient	-.362**	.201	.194	.226	-.066	.118	1.000																						
		Sig. (2-tailed)	.010	.161	.176	.116	.649	.414	.																						
		N	50	50	50	50	50	50	50																						
Low, lower (total)	LOW_TL	Correlation Coefficient	-.354*	.307*	.290*	.185	-.176	-.106	.456**	1.000																					
		Sig. (2-tailed)	.012	.030	.041	.252	.224	.467	.001	.																					
		N	50	50	50	50	50	50	50	50																					
Finalise	FIN	Correlation Coefficient	.346*	-.093	-.057	-.284	.114	.124	-.189	-.156	1.000																				
		Sig. (2-tailed)	.014	.666	.692	.064	.431	.392	.189	.282	.																				
		N	50	50	50	50	50	50	50	50	50																				
Service (intangible)	SV_INTAN	Correlation Coefficient	-.342*	.159	.149	.145	-.134	-.022	.356**	.437**	-.118	1.000																			
		Sig. (2-tailed)	.015	.271	.302	.314	.354	.879	.004	.001	.413	.																			
		N	50	50	50	50	50	50	50	50	50	50																			
Gain	GAIN	Correlation Coefficient	-.342*	.182	.156	-.011	-.134	-.162	.024	.128	-.118	-.034	1.000																		
		Sig. (2-tailed)	.015	.205	.281	.938	.354	.262	.867	.377	.413	.815	.																		
		N	50	50	50	50	50	50	50	50	50	50	50																		

Success,successful	SUCC	Correlation Coefficient	.315*	-.073	-.081	-.207	.140	.128	-.086	-.244	-.035	.084	.025	1.000													
		Sig. (2-tailed)	.026	.613	.578	.149	.331	.377	.552	.088	.808	.562	.863	.													
		N	50	50	50	50	50	50	50	50	50	50	50	50													
Return (tangible)	RET_TAN	Correlation Coefficient	-.313*	.023	.003	.094	-.123	-.148	.421**	.217	-.108	.351*	-.008	.040	1.000												
		Sig. (2-tailed)	.027	.874	.985	.518	.397	.305	.002	.130	.453	.012	.954	.780	.												
		N	50	50	50	50	50	50	50	50	50	50	50	50	50												
Ongoing	OG	Correlation Coefficient	-.313*	.394**	.365**	.289*	-.123	-.148	-.013	.113	-.108	.175	.010	.022	.431**	1.000											
		Sig. (2-tailed)	.027	.005	.009	.042	.397	.305	.927	.436	.453	.224	.944	.877	.002	.											
		N	50	50	50	50	50	50	50	50	50	50	50	50	50	50											
Volume, number	VOL	Correlation Coefficient	-.313*	.427**	.432**	.072	-.123	-.148	.227	.234	-.108	.419**	.379**	-.015	.038	.024	1.000										
		Sig. (2-tailed)	.027	.002	.002	.618	.397	.305	.113	.101	.453	.002	.007	.920	.794	.870	.										
		N	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50										
Increase (adverse)	INC_ADV	Correlation Coefficient	-.313*	.266	.245	.147	-.123	.004	.118	.214	-.108	.324*	.209	.045	.188	.415**	.038	1.000									
		Sig. (2-tailed)	.027	.062	.086	.309	.397	.978	.414	.135	.453	.022	.146	.755	.188	.003	.794	.									
		N	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50									
Increase (total)	INC_TL	Correlation Coefficient	-.302*	.306*	.302*	.029	-.095	-.217	.007	.304*	-.274	.247	.265	.015	.170	.270	.295*	.475**	1.000								
		Sig. (2-tailed)	.033	.031	.033	.841	.512	.130	.969	.032	.054	.084	.063	.919	.239	.058	.038	.000	.								
		N	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50								
Direction	DIRN	Correlation Coefficient	.297*	-.012	-.006	-.226	.163	.160	.032	-.132	.205	-.102	-.102	.265	-.093	-.093	-.093	-.093	-.287*	1.000							
		Sig. (2-tailed)	.036	.935	.971	.114	.257	.267	.822	.358	.153	.483	.483	.063	.521	.521	.521	.521	.043	.							
		N	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50							
Not be achieved,to long to achieve	NOT_ACH	Correlation Coefficient	.297*	-.109	-.103	-.226	.196	.170	-.162	-.132	.267	-.102	-.102	-.249	-.093	-.093	-.093	-.093	-.109	-.064	1.000						
		Sig. (2-tailed)	.036	.452	.478	.114	.172	.238	.261	.358	.061	.483	.483	.082	.521	.521	.521	.521	.452	.660	.						
		N	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50						
Writeoff	WROFF	Correlation Coefficient	.297*	-.110	-.104	-.145	.477**	.354*	.212	-.133	-.074	-.102	-.102	.204	-.093	-.093	-.093	-.093	-.264	.305*	-.064	1.000					
		Sig. (2-tailed)	.036	.447	.473	.315	.000	.012	.140	.358	.608	.483	.483	.155	.521	.521	.521	.521	.063	.031	.660	.					
		N	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50					
Evaluation,revaluation	VAL_DYN	Correlation Coefficient	.297*	-.226	-.226	-.115	-.084	.392**	-.162	-.133	.254	-.102	-.102	-.054	-.093	-.093	-.093	-.093	-.161	-.064	.305*	-.064	1.000				
		Sig. (2-tailed)	.036	.114	.114	.427	.562	.005	.261	.358	.075	.483	.483	.709	.521	.521	.521	.521	.263	.660	.031	.660	.				
		N	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50				
Implement	IMP	Correlation Coefficient	-.285*	.165	.162	.126	-.194	-.235	.206	.207	-.172	.246	-.148	-.145	.169	.027	.078	-.121	.021	-.147	-.147	-.147	.065	1.000			
		Sig. (2-tailed)	.044	.251	.262	.382	.176	.101	.152	.150	.232	.086	.303	.314	.242	.852	.589	.403	.887	.307	.307	.307	.654	.			
		N	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50			
Product	PRT	Correlation Coefficient	.286*	-.007	-.002	-.183	.318*	.086	-.053	.033	.198	.086	.019	.221	-.208	-.187	.114	-.122	.203	.064	-.009	-.012	.025	-.065	1.000		
		Sig. (2-tailed)	.044	.962	.990	.202	.026	.550	.717	.818	.167	.553	.895	.123	.147	.194	.431	.399	.157	.660	.948	.934	.861	.653	.		
		N	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50		
Cost (tangible)	COST_TAN	Correlation Coefficient	-.283*	.045	.049	-.037	-.092	-.014	.012	.018	.015	.116	-.003	-.118	.136	.069	.048	.191	.254	-.169	.235	-.169	.060	.101	.017	1.000	
		Sig. (2-tailed)	.046	.758	.734	.799	.526	.922	.935	.899	.917	.424	.984	.413	.348	.634	.740	.183	.075	.241	.101	.241	.680	.487	.908	.	
		N	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	
Offset	OFFSET	Correlation Coefficient	-.283*	.127	.104	.091	-.111	-.134	.146	.101	-.098	.220	.252	-.062	.431**	.463**	.069	.489**	.321*	-.084	-.084	-.084	-.084	.036	-.078	.164	1.000
		Sig. (2-tailed)	.046	.380	.470	.530	.444	.354	.311	.486	.498	.125	.077	.670	.002	.001	.632	.000	.023	.562	.562	.562	.562	.801	.592	.255	.
		N	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50

Design	DESIGN	Correlation Coefficient	-.283*	.158	.148	.062	-.111	.042	.192	.224	-.098	.197	.244	.017	.231	.069	-.123	.239	.234	-.084	-.084	-.084	-.084	.217	.320*	.205	.281*	1.000					
		Sig. (2-tailed)	.046	.272	.311	.669	.444	.772	.182	.101	.498	.171	.088	.909	.107	.633	.397	.095	.102	.562	.562	.562	.562	.130	.024	.152	.048	.					
		N	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50				
Hardwork	HW	Correlation Coefficient	-.283*	.265	.238	.187	-.111	-.124	.208	-.058	-.098	.244	.021	-.076	.047	.257	.098	.047	-.058	-.084	-.084	-.084	-.084	.182	.016	.024	.302*	.077	1.000				
		Sig. (2-tailed)	.046	.062	.096	.193	.444	.354	.148	.690	.498	.088	.885	.598	.748	.071	.554	.748	.627	.562	.562	.562	.562	.205	.913	.871	.023	.596	.				
		N	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50				
Low (beneficial)	LOW_BEN	Correlation Coefficient	-.283*	.216	.198	.132	-.111	.035	.398**	.614**	-.098	.596**	.021	-.128	.047	.069	.111	.231	.102	-.084	-.084	-.084	-.084	.372**	.053	.011	.077	.289*	.077	1.000			
		Sig. (2-tailed)	.046	.132	.168	.361	.444	.809	.004	.000	.498	.000	.885	.376	.748	.633	.442	.107	.480	.562	.562	.562	.562	.008	.713	.941	.595	.042	.596	.			
		N	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50			
No dividend	ND	Correlation Coefficient	-.283*	.086	.091	.124	-.111	.035	.179	.105	-.098	.035	.256	-.151	.054	-.123	.120	.062	.183	-.084	-.084	-.084	-.084	.057	-.077	.154	.085	.297*	.115	.093	1.000		
		Sig. (2-tailed)	.046	.553	.531	.392	.444	.809	.214	.466	.498	.809	.073	.295	.709	.397	.406	.670	.204	.562	.562	.562	.562	.693	.597	.287	.558	.026	.427	.519	.		
		N	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50		

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Table 6 Correlation description of 7 key-word variables

			PC	HIGH_TL	GW	LOW_TL	FIN	GAIN	WROFF	ND
Spearman's rho	performance characteristics	Correlation Coefficient	1.000							
		Sig. (2-tailed)	.							
		N	64							
	HIGH_TL	Correlation Coefficient	-.533**	1.000						
		Sig. (2-tailed)	.000	.						
		N	50	50						
	GW	Correlation Coefficient	.391**	-.228	1.000					
		Sig. (2-tailed)	.005	.112	.					
		N	50	50	50					
	LOW_TL	Correlation Coefficient	-.354*	.307*	-.175	1.000				
		Sig. (2-tailed)	.012	.030	.224	.				
		N	50	50	50	50				
	FIN	Correlation Coefficient	.346*	-.063	.114	-.155	1.000			
		Sig. (2-tailed)	.014	.666	.431	.283	.			
		N	50	50	50	50	50			
	GAIN	Correlation Coefficient	-.342*	.182	-.134	.128	-.118	1.000		
		Sig. (2-tailed)	.015	.205	.354	.377	.413	.		
		N	50	50	50	50	50	50		
	WROFF	Correlation Coefficient	.297*	-.110	.477**	-.133	-.074	-.102	1.000	
		Sig. (2-tailed)	.036	.447	.000	.358	.608	.483	.	
		N	50	50	50	50	50	50	50	
	ND	Correlation Coefficient	-.283*	.086	-.111	.105	-.098	.256	-.084	1.000
		Sig. (2-tailed)	.046	.553	.444	.466	.498	.073	.562	.
		N	50	50	50	50	50	50	50	50

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

As Table 6 shows, the correlation coefficient between "HIGH_TL" and "LOW_TL" is relatively high, which is 0.307. This means the divergence between these two words is insignificant that multicollinearity is potentially a problem when considering whether the word needs to be eliminated. Thus, a future test is required. For this purpose, these two words were tested respectively with other five word variables. The accuracy of the five-words model is 76%, and the classification accuracy with "HIGH_TL" or "LOW_TL" is 86% and 76% respectively. Moreover, the standardised canonical function of each variable was compared under the three scenarios. As Table 7 shows, there is no significant change among the three scenarios. Therefore, both "HIGH_TL" and "LOW_TL" can be included in the model.

Table 7 Three standardized canonical discriminant functions

Standardized Canonical Discriminant Function Coefficients with both "HIGH_TL" and "LOW_TL"

	Function 1
HIGH_TL	.691
GW	-.264
FIN	-.387
GAIN	.538
WROFF	-.439
ND	.262
LOW_TL	.274
EX	.343
MGT_EXP	.735

Standardized Canonical Discriminant Function Coefficients with only "HIGH_TL"

	Function 1
HIGH_TL	.812
GW	-.262
FIN	-.414
GAIN	.513
WROFF	-.446
ND	.242
EX	.331
MGT_EXP	.678

Standardized Canonical Discriminant Function Coefficients with only "LOW_TL"

	Function 1
GW	-.397
FIN	-.417
GAIN	.530
WROFF	-.411
ND	.295
LOW_TL	.639
EX	.238
MGT_EXP	.699

However, either "HIGH_TL" or "LOW_TL" is composed of both "beneficial" aspects and "negative" aspects. "HIGH_Beneficial" includes "high asset" and "high profit"; "HIGH_Adverse" includes "high competitive" and "high production cost". "LOW_Beneficial" contains "low production cost" and "low turnover rate", while "LOW_Adverse" includes "low profit". Under this classification, "HIGH_TL" and

"LOW_TL" are highly associated with each other in the raw data classification process. Since the inclusion of both "HIGH_TL" and "LOW_TL" into the same equation does seem to make a significant difference, and also for more accurate classification, a new variable: "high minus low (High_Low)" is created. The data collecting equation of this new variable is listed as below:

High - Low = HIGH_Beneficial + LOW_Beneficial - HIGH_Adverse - LOW_Adverse

The correlation coefficient between this new variable and performance, along with inter-correlation of other five variables are listed in Table 8.

Table 8 Correlations between six word-variables and performance characteristics

			performance characteristics	HIGH_ LOW	GW	FIN	GAIN	WROFF	ND
Spearman's rho	performance characteristics	Correlation Coefficient	1.000						
		Sig. (2-tailed)	.						
		N	64						
	HIGH_LOW	Correlation Coefficient	-.406**	1.000					
		Sig. (2-tailed)	.003	.					
		N	50	50					
	GW	Correlation Coefficient	.391**	-.172	1.000				
		Sig. (2-tailed)	.005	.231	.				
		N	50	50	50				
	FIN	Correlation Coefficient	.346*	-.023	.114	1.000			
		Sig. (2-tailed)	.014	.876	.431	.			
		N	50	50	50	50			
	GAIN	Correlation Coefficient	-.342*	.090	-.134	-.118	1.000		
		Sig. (2-tailed)	.015	.533	.354	.413	.		
		N	50	50	50	50	50		
	WROFF	Correlation Coefficient	.297*	-.058	.477**	-.074	-.102	1.000	
		Sig. (2-tailed)	.036	.687	.000	.608	.483	.	
		N	50	50	50	50	50	50	
	ND	Correlation Coefficient	-.283*	.050	-.111	-.098	.256	-.084	1.000
		Sig. (2-tailed)	.046	.729	.444	.498	.073	.562	.
		N	50	50	50	50	50	50	50

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

The classification accuracy with the six new variables is 82%. Although the classification accuracy has dropped after substituting the "HIGH_LOW" variable for "HIGH_TL" and "LOW_TL", the new variable makes this model more reasonable. It illustrates the difference between two opposite variables, avoids the overlap, and reduces the error.

Then, the six variables were put into SPSS to make discriminant analysis. Repeat the prior choose process, if the variable could not cause any improvement of the classification accuracy, it would be eliminated. This elimination process is shown in Table 9. One variable ("Goodwill" – GW) was excluded from the study because it did not meet this criterion. Therefore, five variables were chosen to build the classification model. The five words are: high-low (HIGH_LOW), finalise (FIN), gain (GAIN), write-off (WROFF), and no dividend (ND).

Table 9 Final elimination process

HIGH_LOW		Classification Results[a]			
		characteristics	poor	good	Total
Original	Count	poor	19	10	29
		good	4	17	21
	%	poor	65.5	34.5	100.0
		good	19.0	81.0	100.0

a. 72.0% of original grouped cases correctly classified.

HIGH_LOW + GW		Classification Results[a]			
		characteristics	poor	good	Total
Original	Count	poor	26	3	29
		good	15	6	21
	%	poor	89.7	10.3	100.0
		good	71.4	28.6	100.0

a. 64.0% of original grouped cases correctly classified.

HIGH_LOW + FIN		Classification Results[a]			
		characteristics	poor	good	Total
Original	Count	poor	19	10	29
		ime	3	18	21
	%	poor	65.5	34.5	100.0
		good	14.3	85.7	100.0

a. 74.0% of original grouped cases correctly classified.

HIGH_LOW + FIN + GAIN		Classification Results[a]			
		characteristics	poor	good	Total
Original	Count	poor	21	8	29
		good	2	19	21
	%	poor	72.4	27.6	100.0
		good	9.5	90.5	100.0

a. 80.0% of original grouped cases correctly classified.

HIGH_LOW + FIN + GAIN + WROFF		Classification Results[a]			
		characteristics	poor	good	Total
Original	Count	poor	21	8	29
		good	1	20	21
	%	poor	72.4	27.6	100.0
		good	4.8	95.2	100.0

a. 82.0% of original grouped cases correctly classified.

HIGH_LOW + FIN + GAIN + WROFF + ND		Classification Results[a]			
		characteristics	poor	good	Total
Original	Count	poor	22	7	29
		good	1	20	21
	%	poor	75.9	24.1	100.0
		good	4.8	95.2	100.0

a. 84.0% of original grouped cases correctly classified.

5.2.2 Multivariate models

After selecting the five variables, other residual insignificant correlation words were put into SPSS following a descending order correlation. Similar to the above mentioned elimination process, the word would be kept if it made a contribution to classification accuracy, and the word not making a contribution would be eliminated. During this process, another word was selected since only this word improved the classification accuracy from 84% to 86% (Table 10). This word is: "management change_expected (MGT_EXP)", which includes the parses such as "management retire", "management replace", "management transformation", and "management appointment (change in expected respect)".

Table 10 Adding a new variable

HIGH_LOW + FIN + GAIN + WROFF + ND + MGT_EXP		Classification			
		characteristics	poor	good	Total
Original	Count	poor	26	3	29
		good	4	17	21
	%	poor	89.7	10.3	100.0
		good	19.0	81.0	100.0

a. 86.0% of original grouped cases correctly classified.

Thus, the final model has six variables in total. The Classification Function Coefficients (Table 11) are listed below, and the word-based model is formulated using Fisher discriminant analysis as introduced in the research method chapter, Chapter Four. The following model is generated:

$$Z = 0.712 - 922.995(HIGH\_LOW) + 1410.025(FIN) - 1187.951(GAIN) + 1608.653(WROFF) - 983.305(ND) - 450.136(MGT\_EXP)$$

Where Z is the discriminant score,

HIGH_LOW = the difference between beneficial high, beneficial low and adverse high, adverse low;

FIN = the symbol of "finalise";

GAIN = the symbol of "gain";

WROFF = the symbol of "write-off";

ND = the symbol of "no dividend";

MGT_EXP = the symbol of "expected management change".

Table 11 Fisher's linear discriminant function coefficients

	performance		
	poor	good	Cj
HIGH_LO	1045.215	122.220	-922.995
FIN	-35.974	1374.050	1410.025
GAIN	1311.340	123.388	-1187.951
WROFF	-756.966	851.687	1608.653
ND	1053.673	70.367	-983.305
MGT_EXP	663.898	213.761	-450.136
(Constant)	-2.115	-1.403	.712

Fisher's linear discriminant functions

This model can correctly classify 88% of companies (i.e., with seven misclassifications): three Type I error and four Type II errors (Table 12). Although the accuracy of this model is lower than the 98% reported by Smith and Taffler (2000), $\alpha = 0.001$ that is much less than 0.05 and means the model is statistically robust.

Table 12 Final classification result

			Predicted Group		
		performance characteristics	poor	good	Total
Original	Count	poor	26	3	29
		good	4	17	21
	%	poor	89.7	10.3	100.0
		good	19.0	81.0	100.0

a. 86.0% of original grouped cases correctly classified.

5.2.3 Explanatory power

Following the research by Smith and Taffler (2000), Mahanolobis Distance (Mosteller & Wallace, 1963) was used to calculate the explanatory power of each variable. Mahanolobis Distance measures the contribution percentage of each variable. It is a fundamental and important approach in data analysis with multiple measurements (McLachian, 1999). Table 13 illustrates the calculation process of each variable's explanatory power.

Table 13 Calculation of explanatory power

Variable Symbol	Mean			Coefficient		Cj		Explanatory power
	Poor performance	Good performance	Good - Poo	Poor performance	Good performance	Good - Poo	good - poor * Cj	
HIGH_LO	.001346	.000243	-.001103	1045.215	122.220	-922.995	1.018278	32.456%
FIN	.000000	.000367	.000367	-35.974	1374.050	1410.025	.517726	16.502%
GAIN	.000461	.000000	-.000461	1311.340	123.388	-1187.951	.547247	17.443%
WROFF	.000000	.000403	.000403	-756.966	851.687	1608.653	.647561	20.640%
ND	.000180	.000000	-.000180	1053.673	70.367	-983.305	.176703	5.632%
MGT_EXP	.001415	.000905	-.000511	663.898	213.761	-450.136	.229893	7.327%
Sum							3.137407	100.000%

5.3 Test of hypotheses

After building the word-based prediction model, the six hypotheses were tested as follows.

H_{1a}: Theme-based variables, in the Chairman's Statement, are significantly associated with corporate performance.

Based on eight themes only, the classification accuracy is 64% (Table 14), which is not as accurate as the word-based model (86%). Thus, the word-based model is recommended, and it will be the primary focus in the remainder of this thesis. However, the accuracy of 64% is still significant, and can prove that there is an association between theme-based variables and corporate performance characteristics. Also due to $p< 0.05$, the first hypothesis (H_{1a}) cannot be rejected.

Table 14 Classification result of theme-based variable

		performance characteristics	Membership		Total
			poor	good	
Original	Count	poor	23	6	29
		good	12	9	21
	%	poor	79.3	20.7	100.0
		good	57.1	42.9	100.0

a. 64.0% of original grouped cases correctly classified.

H_{1b}: Word-based variables, in the Chairman's Statement, are significantly associated with corporate performance.

As discussed above, the word-based classification model can successfully classify 86% of companies between good and poor performance, also since $p< 0.05$, this hypothesis (H_{1b}) cannot be rejected. This study is consistent with the research of Smith and Taffler (2000) which proved that word-based variables in the Chairman's Statement are significantly associated with corporate performance, and that these words can be used as indicators of performance classification and prediction.

H_{1c}: Report size of the Chairman's Statement is not significantly associated with corporate performance.

Report size is measured by total word number of the Chairman's Statement. As Table 15 shows, the correlation coefficient with corporate performance is insignificant

(r=-0.212). Moreover, since p=0.139, which is higher than 0.05, H_{1c} is accepted. Thus, there is no significant association between report size and corporate performance.

H_{2a}: The number of "Big words" (>6 letters), in the Chairman's Statement, is significantly associated with corporate performance.

In this study, the number of "big words" is used as a proxy for jargon and as an indicator of corporate obfuscation practises. Its correlation coefficient with corporate performance characteristic is not significant (r=0.031). Moreover, since p=0.831, which is higher than 0.05, H_{2a} is rejected. Thus, there is no significant association between "big words" and corporate performance.

H_{2b}: The readability level of the Chairman's Statement is significantly associated with corporate performance.

Readability level reflects the complexity of narrative disclosures, which is used as another indicator for corporate implement of obfuscation practices. As Table 15 shows, the correlation coefficient with corporate performance is not significant (r=0.098). Moreover, since p=0.497, which is higher than 0.05, H_{2b} is rejected. Thus, there is no significant association between readability and corporate performance. Although this result conflicts with the finding of Smith and Taffler (1992a), it is consistent with Courtis (1995) "From the limited sample studied, no apparent relationship exists between corporate profitability and enhanced annual report readability" (p. 11).

Table 15 Effective correlations between three variables and corporate performance

Spearman's rho		performance characteristic	Report_size	BIG_WORDS	readability
performance characteristics	Correlation Coefficient	1.000			
	Sig. (2-tailed)	.			
	N	64			
Report_size	Correlation Coefficient	-.212	1.000		
	Sig. (2-tailed)	.139	.		
	N	50	50		
BIG_WORDS	Correlation Coefficient	.031	.249	1.000	
	Sig. (2-tailed)	.831	.081	.	
	N	50	50	50	
readability	Correlation Coefficient	.098	.071	-.321*	1.000
	Sig. (2-tailed)	.497	.624	.023	.
	N	50	50	50	50

*. Correlation is significant at the 0.05 level (2-tailed).

H_{2c}: The use of persuasive language in the Chairman's Statement is significantly associated with corporate performance.

Personal pronouns and emotional words reflect the use of persuasive language, which are aggregated to be the third indicator for corporate implement of obfuscation practice. As listed in Table 16, their correlation coefficients with performance are all not significant (r_{PRON_TL}=-0.044, r_{EM_TL}=-.0.146). Moreover, p_{PRON_TL}=0.764, p_{EM_TL}=0.312, which are both higher than 0.05. Therefore, H_{2c} is rejected. Persuasive language, including personal pronouns and emotional words, are not significant associated with corporate performance.

Table 16 Correlations between performance characteristics and three obfuscation indicators

Spearman's rho		performance characteristics	PRON_TL	I	me	we	us	our	1st_PRON	you	your	2nd_PRON	EM_TL	EM_positive	EM_negative
performance characteristics	Correlation Coefficient	1.000													
	Sig. (2-tailed)	.													
	N	64													
PRON_TL	Correlation Coefficient	-.044	1.000												
	Sig. (2-tailed)	.764	.												
	N	50	50												
I	Correlation Coefficient	-.181	.541**	1.000											
	Sig. (2-tailed)	.210	.000	.											
	N	50	50	50											
me	Correlation Coefficient	-.110	.012	.032	1.000										
	Sig. (2-tailed)	.445	.935	.824	.										
	N	50	50	50	50										
we	Correlation Coefficient	.018	.879**	.336*	-.046	1.000									
	Sig. (2-tailed)	.900	.000	.017	.749	.									
	N	50	50	50	50	50									
us	Correlation Coefficient	-.025	.404**	.101	-.180	.448**	1.000								
	Sig. (2-tailed)	.864	.004	.484	.210	.001	.								
	N	50	50	50	50	50	50								
our	Correlation Coefficient	-.044	.880**	.334*	-.131	.698**	.252	1.000							
	Sig. (2-tailed)	.763	.000	.018	.364	.000	.078	.							

	N	50	50	50	50	50	50	50							
1st_PRON	Correlation Coefficient	-.039	.993**	.519**	-.017	.892**	.424**	.889**	1.000						
	Sig. (2-tailed)	.786	.000	.000	.906	.000	.002	.000	.						
	N	50	50	50	50	50	50	50	50						
you	Correlation Coefficient	.181	.240	.220	.176	.201	.023	.040	.188	1.000					
	Sig. (2-tailed)	.209	.093	.125	.222	.163	.874	.781	.191	.					
	N	50	50	50	50	50	50	50	50	50					
your	Correlation Coefficient	.186	.395**	.273	.211	.310*	-.061	.206	.328*	.316*	1.000				
	Sig. (2-tailed)	.195	.004	.055	.142	.028	.673	.150	.020	.026	.				
	N	50	50	50	50	50	50	50	50	50	50				
2nd_PRON	Correlation Coefficient	.123	.411**	.327*	.219	.305*	-.036	.170	.332*	.626**	.898**	1.000			
	Sig. (2-tailed)	.396	.003	.021	.127	.031	.802	.239	.019	.000	.000	.			
	N	50	50	50	50	50	50	50	50	50	50	50			
EM_TL	Correlation Coefficient	-.146	-.073	-.104	-.025	.050	-.184	-.079	-.082	-.287*	.118	.007	1.000		
	Sig. (2-tailed)	.312	.612	.472	.865	.731	.201	.587	.571	.043	.415	.962	.		
	N	50	50	50	50	50	50	50	50	50	50	50	50		
EM_positive	Correlation Coefficient	-.204	.023	.063	.021	.052	-.124	-.008	.001	-.231	.227	.101	.881**	1.000	
	Sig. (2-tailed)	.156	.875	.664	.883	.720	.390	.955	.995	.107	.112	.486	.000	.	
	N	50	50	50	50	50	50	50	50	50	50	50	50	50	
EM_negative	Correlation Coefficient	.040	-.084	-.341*	-.146	.042	-.057	.012	-.060	.013	-.262	-.167	.225	-.185	1.000
	Sig. (2-tailed)	.785	.563	.015	.312	.770	.697	.934	.677	.930	.066	.246	.116	.198	.
	N	50	50	50	50	50	50	50	50	50	50	50	50	50	50

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

5.4 Reliability and validity

Many researchers have pointed out that the reliability and validity of content analysis are critical and debatable. This study adopts the following approaches to improve data classification reliability and result validity.

5.4.1 Reliability

Firstly, a customised keyword dictionary was compiled to ensure that the word classification is stable. Secondly, to reduce cognitive classification error, a whole classification recheck was conducted after the draft data collection had been completed. This check ensures that the classification process is reproducible, and also contributes to keyword dictionary reliability. Thirdly, an independent check from another researcher was conducted to confirm the reliability of classification.

5.4.2 Validity

The results of this study (H_{1a} and H_{1b}) are consistent with prior research (Ingram & Frazier, 1983; Tennyson, et al., 1990; Smith & Taffler, 2000; Aerts, 2005; Rutherford 2005) that both word-based and theme-based narrative disclosures have the predictive ability of corporate performance.

Moreover, six variables in word-based classification model are reasonable for predicting the corporate performance (Table 17). Both "HIGH_TL" and "LOW_TL" are prominent variables associated with corporate performance, since "LOW_TL" is dominated by the "HIGH_TL" variable, the combined variable "HIGH_LOW" is positive related with corporate performance. The variable "ND (no dividend)" is also chosen by Smith and Taffler (2000) as "NOMDIV", which is the sum of "no dividend" and "nominal dividend". Variable "GAIN" is negatively associated with corporate performance, which indicates poorly performing companies may tend to provide more narrative disclosures about "gain". This is consistent with the research result of

Merkl-Davies and Brennan (2007). On the contrary, both "FIN" and "WROFF" variables are positively related with corporate performance. Variable "ND" is a variable that was also incorporated by Smith and Taffler (2000). In their study, they found that the variable "NOMDIV (no dividend + nominal dividend)" contributed the highest explanatory power to the classification model (26.7%). Variable "MGT_EXP" is significant negatively correlated with corporate performance. This can be explained by a change of management often being related to fluctuating corporate financial situation which is a negative signal.

Table 17 Details of six variables

Variable symbol	Keywords	Explanatory power %
HIGH_LOW	high, higher, highest - low, lower, lowes	32.456%
WROFF	writeoff	20.640%
GAIN	gain	17.443%
FIN	finalise	16.502%
MGT_EXP	expected management change	7.327%
ND	no dividend	5.632%
Total		100%

H_{1c} confirms that there is no significant difference between report size and corporate performance. In terms of the other three performance related hypotheses (H_{2a}: big words; H_{2b}: readability level; H_{2c}: persuasive language), they are rejected by this study. The result validity may be influenced by limited sample size; however, the findings are largely consistent with those of prior research. Smith and Taffler (1992b), Courtis (1995), Clatworthy and Jones (2001), and Rutherford (2003) all concluded that the readability level is not related to corporate performance.

Chapter Six: Conclusions

6.1 Discussion

The purpose of this study is to explore the relationship between narrative disclosures and corporate performance. It only focuses on discretionary narrative disclosures, especially the Chairman's Statement in the corporate annual report. There are several reasons why the study is significant. To start with, this kind of narrative disclosures contains incremental information which assists outside information users to make better decisions. However, compared with outside information users and regulatory authorities, corporate management acquires more information and has the initiative of information discourse. Under this information asymmetry, the managements may take advantage of their superior information position to choose the disclose content and the disclose approach. Therefore, it is essential for both outside information users and regulatory authorities to understand "what" and "how" narrative information is disclosed by corporate management.

There is an extensive literature in this research field. Content analysis is a predominant study approach that has been used by many researchers. This paper concentrates on discretionary narrative disclosures by studying both thematic content analysis approach and syntactic content analysis approach.

Signalling theory and agency theory are two underpinning theories in this research field. Furthermore, two branches of impression management strategy, rationalisation and obfuscation are involved in this study. Rationalisation focuses on "what" is disclosed by companies, which is a quantity-oriented approach; while obfuscation focuses on "how" information is disclosed by companies, which is a quality-oriented approach.

This is a parallel study of Smith and Taffler (2000) which they examined whether the discretionary narrative disclosures have the ability to measure corporate financial risk of bankruptcy. Similar with their research, this study explores the predictive ability of discretionary narrative disclosures (the Chairman's Statement) in terms of distinguish good performers from poor performers. Six related hypotheses have been developed based on prior research and the theoretical framework. Based on the sample of 64 Australian listed manufacturing companies, this study develops a six-words classification model and finds that both theme-based variable (meaning oriented) and word-based variable (form oriented) in the Chairman's Statement are significantly correlated with corporate performance. It confirms the findings of the research conducted by Smith and Taffler (2000). Moreover, this study expands prior research by adding syntactic content analysis variables to test their correlations with corporate performance. The result indicates that other selected variables include report size, big words, readability level, and persuasive language do not have a significant relationship with corporate performance. This study adopts both a manual coding approach and used computer-based softwares (LIWC and SPSS) to collect data; independent checks and reproducing checks are processed to improve research reliability and validity.

6.2 Study values

This study is based on the Smith and Taffler (2000) framework which focused on UK manufacturing companies and found that there is a significant association between narrative disclosures (the Chairman's Statement) and corporate performance (healthy/failed). Since there is no predictive model directly applied to Australian companies, and most research focuses on the predictive ability of financial ratios instead of narrative disclosures, this study makes its contribution to filling this research gap by employs the most current discretionary narrative disclosures (the Chairman's Statement) of 64 Australian manufacturing companies.

Importantly, this study does not adopt the traditional healthy/failed delineation between companies (e.g.: Smith & Taffler, 2000). Instead, all of the companies in this study are surviving in 2009, and are grouped into "good performers" and "bad performers" based on the Staw et al. (1983) methodology.

Moreover, this parallel study also distinguishes itself by combines prior thematic content analysis research ("what to disclose") with syntactic content analysis research ("how to disclose"). It focuses on whether there is a relationship between corporate performance and disclosures in the respects of both "what" and "how" disclosure messages be conveyed to convince readers. For this purpose, both thematic content analysis and syntactic content analysis are adopted in this study. Meanwhile, this study develops a new classification model which is developed with a proven accuracy of 86%. This is an extremely high classification accuracy given that it is considering companies which are good/poor rather than healthy/failed (e.g.: Smith & Taffler, 2000).

6.3 Limitations

As with other empirical studies, there are some potential limitations in this study, and the generalizability of this research result into other areas needs to be evaluated in further research.

Firstly, many researchers have criticized the content analysis approach because "content analysis is partly an art and depends on the judgment and interpretation of the investigator" (Weber, 1990, p. 62). Thus researcher bias is unavoidable and exists in coding and data selection stages. Although the manual coding approach, computer software, independent check, and reproducing check have substantially overcome the problem of subjective impact and enhanced the reliability of outcomes, there is still an absence of an objective methodology.

Secondly, this research only selects a small group of sample which is 64 Australian manufacturing companies (29 good performers and 35 poor performers). The small sample will limit the reliability of the findings.

Thirdly, this research only focuses on manufacturing companies without industries comparison, and it is not sure whether the result can be applied to other Australian industries.

Lastly, due to time limitations, the classification model has not been tested after its development.

6.4 Further study

An extension of this study will be to repeat the narrative analysis conducted for 2008, in 2009. Then the classification model (built on 2008 data) can be tested to determine the extent to which it is a successful predictive model for 2009.

Moreover, the prediction study lacks an accurate measurement to distinguish between good and poor performance companies. The sample size has been limited by being confined to: (a) Australian manufacturing companies, and (b) to groups determined by the Staw et al. (1983) metric. Future study might address a larger company base, and use alternative metrics for distinguishing between "good" and "poor" performances.

This study focuses on corporate narrative disclosures, and does not consider the possible relationship between corporate performance and graphs, pictures, and other pictorial information. Furthermore, this study did not take firm size, type of industry, or fiscal year into consideration. Thus, further study in these areas is recommended.

References

Adelberg, A. H., & Razek, J. R. (1984). The cloze procedure: A methodology for determining the understandability of accounting textbooks. *Accounting Review, 59*(1), 109-122.

Aerts, W. (1994). On the use of accounting logic as an explanatory category in narrative accounting disclosures. *Accounting, Organizations and Society, 19*(4/5), 337-353.

Aerts, W. (2005). Picking up the pieces: Impression management in the retrospective attributional framing of accounting outcomes. *Accounting, Organizations and Society*, 30, 493-517.

Ahmed, K., & Courtis, J. (1999). Associations between corporate characteristics and disclosure levels in annual reports: A meta-analysis. *The British Accounting Review, 31*(1), 35-61.

Altman, E. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of finance, 23*(4), 589-609.

American Institute of Certified Public Accountants (AICPA). (1994). *Improving business reporting - A customer focus: Meeting the information needs of investors and creditors, comprehensive report of the special committee on financial reporting* (the Jenkins report). New York: American Institute of Certified Public Accountants.

Andersen, A. (2000). Spice up the story: A survey of narrative reporting in annual reports. *Arthur Andersen*, London.

Australian and New Zealand Standard Industrial Classification (ANZSIC). (2010). *Themes - Manufacturing Statistics.* Retrieved May 24, 2010 from http://www.abs.gov.au/websitedbs/c311215.nsf/20564c23f3183fdaca25672100813ef1/24d3a88344f5182aca256d1d0004e737!OpenDocument

Beattie, V., McInnes, B., & Fearnley, S. (2004). A methodology for analysing and evaluating narratives in annual reports: A comprehensive descriptive profile and metrics for disclosure quality attributes. *The Accounting Forum, 28*(3), 205-236.

Botosan, C. A. (1997). Disclosure level and the cost of equity capital. *The Accounting Review, 72*(3), 323-349.

Bowen, R., Davis, A., & Matsumoto, D (2005). Emphasis on pro forma versus GAAP earnings in quarterly press releases: Determinants, SEC intervention and market reactions. *The Accounting Review, 80*(4), 1011-1038.

Bowman, E. H. (1984). Content analysis of annual reports for corporate strategy and risk. *Interfaces, 14*(1), 61-71.

Boyatzis, R. (1998). *Transforming qualitative information: Thematic analysis and code development.* London: Sage Publications, Inc.

Brennan, N. M., Guillamon-Saorin, E., & Pierce, A. (2009). Impression management: Developing and illustrating a scheme of analysis for narrative disclosures – a methodological note. *Accounting, Auditing and Accountability Journal, 22*(5), 789-832.

Bryan, S.H. (1997). Incremental information content of required disclosures contained in management discussion and analysis. *The Accounting Review, 72*(2), 285-301.

Clarke, F., & Dean, G. (2007). *Indecent disclosure: Gilding the corporate lily.* Melbourne: Cambridge University Press.

Clatworthy, M., & Jones, M. J. (2001). The effect of thematic structure on the variability of annual report readability. *Accounting, Auditing and Accountability Journal, 14*(3), 311-326.

Coakes, S. J., Steed, L., & Ong, C. (2010). *SPSS: Analysis without anguish: version 17.0 for windows (CD)* . Milton Queensland: John Wiley & Sons Ltd.

Courtis, J. K. (1995). Readability of annual reports: Western versus Asian evidence. *Accounting, Auditing and Accountability Journal, 8*(2), 4-17.

Courtis, J. K. (2004). Corporate report obfuscation: Artefact or phenomenon? *The British Accounting Review, 36*(3), 291-312.

Courtis, J. K. (1998). Annual report readability variability: test of the obfuscation hypothesis. *Accounting, Auditing, and Accountability Journal, 11*(4), 459-471.

Dame, R.F. (1979). The abundance, diversity, and biomass of macrobenthos on North Inlet, South Carolina, intertidal oyster reefs. *Proceedings of the National Academy of Science, 69*, 6–10.

DatadotData Technology Limited. (2008). *2008 Annual report.* Retrieved August 24, 2010 from Mergent Online database.

de Beaugrande, R., & Dressler, W. (1981). *Introduction to Text Linguistics.* New York: Longman.

Deegan, C. (2006). *Financial Accounting Theory* (2nd ed.). Sydney: McGraw Hill Book Company.

Department of Trade and Industry (DTI). (2004). *Draft Regulation on the Operating and Financial Review and Directors' Report: A consultative document.* London: Department of Trade and Industry.

Dixon, K., Coy, D., & Tower, G. (1991). External reporting by New Zealand Universities 1985–1989: Improving accountability. *Financial Accountability and Management, 7*(3), 159-178.

Donoher, W., Reed, R., & Storrud-Barnes, S. (2007). Incentive alignment, control, and the issue of misleading financial disclosures. *Journal of Management, 33*(4), 547-569.

Epstein, M. J., & Freedman, M. (1994). Social disclosure and the individual investor. *Accounting, Auditing and Accountability Journal, 7*(4), 94-109.

Financial Accounting Standards Board 〔FASB〕. (1978). *Concepts Statement No. 1: Objectives of Financial Reporting by Business Enterprises.* Stamford: FASB.

Hackston, D., & Milne, M. (1996). Some determinants of social and environmental disclosures in New Zealand companies. *Accounting, Auditing and Accountability Journal, 9*(1), 77-108.

Healy, P., & Palepu, K. (2001). Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical disclosure literature. *Journal of Accounting and Economics, 31*(1-3), 405-440.

Hockey, S. & Martin, J. (1988). *Oxford Concordance Program: User's Manual 2.*0. Oxford University Computing Service.

Houghton, K.A. (1988). The measurement of meaning in accounting: A critical analysis of the principal evidence. *Accounting, Organizations and Society, 13*(3), 263-80.

Hrasky, S. (2008). *A multimodal Goffmanian analysis of impression management behaviour in the annual reports of listed Australian companies.* Unpublished PhD dissertation, University of Tasmania, Tasmania, Australia.

Ingram, R. W., & Frazier, K. B. (1983). Narrative disclosures in annual reports. *Journal of Business Research, 11*(1), 49-60.

Jones, M. J., Smith, M. & Whale, J. (2010). Readability and understandability revisited. *Financial Reporting and Communication Conference.* Cardiff, UK.

Jones, M., & Shoemaker, P. (1994). Accounting narratives: A review of empirical studies of content and readability. *Journal of Accounting Literature, 13*(1), 142-184.

Jones, M.J. (1988). A longitudinal study of the readability of the chairman's narratives in the corporate reports of a UK company. *Accounting and Business Research, (18)*, 297-305.

Krippendorff, K. (1980). *Content analysis: An introduction to its methodology*. Beverly Hills: Sage Publications.

Lang, M., & Lundholm, R. (1993). Cross-sectional determinants of analyst ratings of corporate disclosures. *Journal of Accounting Research, 31*(2), 246-271.

Marston, C. L., & Shrives, P. J. (1991). The use of disclosure indices in accounting research: A review article. *British Accounting Review, 23*(3), 195-210.

McLachlan, G. (1999). Mahalanobis distance. *Resonance, 4*(6), 20-26.

Meek, G., Roberts, C., & Gray, S. (1995). Factors influencing voluntary annual report disclosures by US, UK and continental European multinational corporations. *Journal of International Business Studies, 26*(3), 555-572.

Merkl-Davies, D., & Brennan, N. (2007). Discretionary disclosure strategies in corporate narratives: Incremental information or impression management? *Journal of Accounting Literature, 26*, 116-194.

Mosteller, F., & Wallace, D. (1963). Inference in an authorship problem. *Journal of the American Statistical Association, 58*(302), 275-309.

Neu, D., Warsame, H. A., & Pedwell, K. A. (1998). Managing public impressions: Environmental disclosures in annual reports. *Accounting, Organizations and Society, 23*(3), 265-282.

Neuman, W. (2006). *Social Research Methods: Qualitative and Quantitative Approaches* (6th ed.). Boston: Pearson Education.

Osgood, L.E., Suci, G.J., & Tannenbaum, P.H. (1957). *The Measurement of Meaning*. Urbana: University of Illinois Press.

Parker, L. (2005). Corporate governance crisis down under: Post-Enron accounting education and research inertia. *European Accounting Review, 14*(2), 383-394.

Pennebaker, J. W. (2002). What our words can say about us: Toward a broader language psychology. *Psychological science agenda, 15*(1), 8–9.

Pennebaker, J. W., Mehl, M. R., & Niederhoffer, K. G. (2003). Psychological aspects of natural language use: Our words, our selves. *Annual review of psychology, 54*, 547-578.

Pennebaker, J.W., Booth, R.J., & Francis, M.E. (2007). *Linguistic Inquiry and Word Count Version 1.08 (CD)*. Austin, Texas: liwc.net.

Prakash, P., & Rappaport, A. (1977). Information inductance and its significance for accounting. *Accounting, Organizations and Society, 2*(1), 29-38.

Robb, S., Single, L., & Zarzeski, L. (2001). Nonfinancial disclosures across Anglo-American countries. *Journal of International Accounting, Auditing and Taxation, 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*(1), 17-31.

Roseberry, R. L. (1995). A texture index: Measuring texture in discourse. *International Journal of Applied Linguistics, 5*(2), 205-223.

Ross, S. A. (1977). The determination of financial structure: The incentive-signalling approach. *The Bell Journal of Economics, 8*(1), 23-40.

Rutherford, B. A. (2003). Obfuscation, textual complexity and the role of regulated narrative accounting disclosure in corporate governance. *Journal of Management and Governance, 7,* 187-210.

Rutherford, B. A. (2005). Genre analysis of corporate annual report narratives: A corpus linguistics-based approach. *Journal of Business Communication, 42*(4), 349-378.

Singhvi, S. (1972). Corporate Management's Inclination to Disclose Financial Information. *Financial Analysts Journal, 28*(4), 66-73.

Smith, M., & Gunalan, S. (1996). The identification of recovery candidates among financially distressed companies. *Accounting and Performance, 2*(2), 69-91.

Smith, M., & Taffler, R.J. (1992a). The Chairman's Statement and corporate financial performance. *Accounting and Finance, 32*(2), 75-90.

Smith, M., & Taffler, R.J. (1992b). Readability and understandability: Different measures of the textual complexity of accounting narrative. *Accounting, Auditing, and Accountability Journal, 5*(4), 84-98.

Smith, M., & Taffler, R.J. (1995). The incremental effect of narrative accounting information in corporate annual reports. *Journal of Business Finance and Accounting, 22*(8), 1195-1210.

Smith, M., & Taffler, R.J. (2000). The Chairman's Statement: A content analysis of discretionary narrative disclosures. *Accounting, Auditing, and Accountability Journal, 13*(5), 624-46.

Smith, M., Jamil, A., Johari, Y. C., & Ahmar, S. A. (2006). The Chairman's Statement in Malaysian companies: A test of the obfuscation hypothesis. *Asian Review of Accounting, 14*(1/2), 49-65.

SPSS Inc. (1986). *SPSS-X: User's Guide, 2nd ed (CD)*. Maidenhead: McGraw Hill.

Staw, B. M., McKechnie, P.I., & Puffer, S. M. (1983). The justification of organizational performance. *Administrative Science Quarterly, 28*(4), 582-600.

Subramanian, R., Insley, R. G., & Blackwell, R. D. (1993). Performance and readability: A comparison of annual reports of profitable and unprofitable corporations. *Journal of Business Communication, 30*(1), 49-61.

Sydserff, R., & Weetman, P. (1999). A texture index for evaluating accounting narratives: An alternative to readability formulas. *Accounting, Auditing and Accountability Journal*, 12(4), 459-488.

Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology, 29*(1), 24-54.

Tedeschi, J., & Riess, M. (1981). *Impression management theory and social psychological research.* New York: Academic Press.

Tennyson, B.M., Ingram, R.W., & Dugan, M.T. (1990). Assessing the information content of narrative disclosures in explaining bankruptcy. *Journal of Business Finance and Accounting, 17*(3), 390-410.

Verrecchia, R. (2001). Essays on disclosure. *Journal of Accounting and Economics, 32*, 97-180.

Watts, R. L., & Zimmerman, J. L. (1986). *Positive theory of accounting.* Englewood Cliffs, NY: Prentice-Hall.

Watts, R. L., & Zimmerman, J. L. (1990). Positive accounting theory: A ten year perspective. *Accounting Review, 65*(1), 131-156.

Weber, R. P. (1990). *Basic Content Analysis* (2nd ed.). California, Newbury Park: Sage Publications.

Appendix A: Customised keyword dictionary

Evaluative		Potency		Activity		Manageability	
Beneficial	**Adverse**	**Tangible**	**Intangible**	**Dynamic**	**Static**	**Expected**	**Unexpected**
able	abnormal conditions	accountability	legacy issues	able, ability	cash reserves	acquisition/acquire	(mgt) departure
ably (chaired)	absence	action	ability	accept offer	cash retention	aims (v)	abnormal conditions
acceleration	accretive	Indicators	accountability	access	closure/disposal	anticipate	alternative uses
acceptance	adverse impact	adjustments	achievement	accommodate	contractions	assume	augur
achieved, achievement	affected yield	adviser	advice	achieve	feedstock	be looking to	bad weather(drought)
actively(engage)	amortization/depreciation	amortization charge	agreement	acquired, acquisition	funding	believe	BSE
acumen	bad weather (frost)	assessment	aim	activity	held up	bode	business condition
adaptations	badly affected	asset	ambitions	address (v)	impact	budget	challenging,challenge,
add	bank debt	benefits (tax)	announced,	adjust	maintain(ing)	certainty	change
address (weaknesses)	bank Notes payable	bids	approach	adopted	no significant changes	climate change	global economy changes
adequately	bankers support	board	aspirations	advance,advancing	partners sought	climate.	claim
advanced technology	bankrupting	brand	attention	advise	remain	consistent,(ly)	climatic conditions
advances	below the target	break even	availability	aim	retain the funding	continue, continuous	competitor
advancing	borrowings	broker coverage	awareness	align	retained earnings	contract	complications
advantage	BSE	budget	belief	allows	returns	development, develop	constant currency
aggressive(pursue/market)	cancellation	business targets	business condition	application	share price	distributor appointment	consumer preference
ahead (budget, plan)	cannot afford to	capital demand	business culture	appointment	shift	envisage	contingent upon
aligned to, align	can't recovered	carrying value	business line	arrangement	stabilized	estimate	court action
ameliorated	cash outflow	cash equivalent	business model	ask (to approve)	static	expansion, expand	financial/capital crisis
appreciation	challenging, challenges	cash flow	capability	assess, assessing,	stayed the same	expect, expectation	currency moves

appropriate (measure)	close subsidiary	cash outflow	challenging, challenge	attain	stock, inventory	extend, extension	damage claims
aspiration	complex	cash position	change	aware	storage	forecast	demographic changes
asset	constrain	charge	circumstances	believes	unchanged	foresee	difficult market condition
assist	contraction	client	claim	building, built	watershed	foreseeable	disease
assure (shareholders)	cost	commodities	code	calculate	withdrawal of support	future (plan)	Industry downturn
attract_v	cost increase	compensation	commitment	capacity		future growth	earthquakes
attractive (markets, price)	cost pressures	competitor	committee meeting	capital raising/capitalize		goal	economic backdrop
available	costly	concession	communication	capitalise		growing range	economic factors
awarded (title)	counteract	cost, cost base	competency.competence	capture		improvement	economic slowdown
believe	cut	dealer	competition,competitiveness	cash flow		in process	economic times
benefit_n (tax benefits)	damaged (severely)	demand (DEPENDS)	concept	cash generating activity, cash generation		increase (size)	economy, economic
benefited_v	debate	deposit	concern	close		indicates	employment rates
best (efforts)	debt	depreciation	confidence	combine		intended, intention	environmental impacts
better	decline	directors	consequence	commenced, commencing		investment	evolution
bolstered	Deferred	distributor	consideration	commend (v)		look for	exchange rates
boost	delay	diversify	constitution	comment on		look forward	export market
bright future	depreciation	dividend	consumption	commissioned		mgt appointment	external (affect)
broaden (presence)	depress	drivers of growth	contingent upon	commitment,committed		move	external (environment)
build its future	depressed (market condition)	earning	contract	completing		objective	external (factor)
buoyant (conditions)	destroy	EBIT	contribution	completion, complete		ongoing	external forces
capable	deteriorating, deterioration	EPS, earnings	corporate governance	concentrated		outlook	failed product/acquisit
capital_n	devaluation	equipment	counsel	concluded		oversee	farming subsidies
capital_v (raising)	difficult	equity capital	court action	consequence		perceived	financial condition
capitalize	difficult (condition)	equity, equity capital	coverage	considered, take consideration, consider,		plan	fire
cash equivalent	difficult (market)	exchange rate	dealership	consolidation		predict	fluctuating prices

cash flow	difficult (period)	expenditure	decision	consummated	prefeasibility	foreign economy
cash generating activities	difficult economic	exploration potential	declaration	contain	preform	foreign exchange
cash reserves	difficulty	face value	demand	contributed	programme, program	further improvements
clear (objective)	dire	figure	demand	control	progress	further investigation
commendable	disadvantage	financial contribution	demonstration	control/take over	project	global (acceleration)
committed, commitment	disappointment, disappoint	financial position	device	convert	promise	global demand
compensation = salary	discontinued operations	financial ratios	direction	cope_with	propose	economic changes
competent	dislocation	franchise group	discipline	corporate governance	prospect, prospective	industry participants
competitive (advantage, price), competitiveness	dispute	gearing ratio/gearing level = leverage	discussion	creating, create,	prospectus	global market
completed	disruption, disrupt	goodwill	economic benefits	deal with	relist	global partner
completion	distracted	governance	effect	decide, make decision,	directorreplace	global supply alliance
comprehensive (search)	divestment	guarantee	effectiveness, efficiency	declare	restructure	globally
confidence, confident	down	hedging	effort	dedicated	retires, retirement	government
confirmed	downturn (in industry)	infrastructure	element	defence, defended	salesmen appointment	government initiatives
congratulate	dreadful financial performance	installments	ensure	define	scheduled	financial stimulus package
conscious	drop	intake	enthusiasm	deliver, delivering,	scheme	ice storm
considerable	economic backdrop	interest	estimate	demonstrates	strategic priorities	in the future
consolidate, consolidation	economic slowdown	interested parties	evolution	determine	strategy	inflationary pressure
contributed, contribution	encountered	inventory	executives	discovered	target	insurance
control over / take over	eroded	investment	expectation	discussion	mgt change/transform	insurance cover
cost reduction/saving/less	eventful	investors	experience	distribution	under control	interest rates
cost savings	exacerbated	item	expertise	diversification, diversify	upgrading, upgrade	international alignment
cost-effective (technologies)	expenditure	joint venture (parties)	financial mgt	divestment	will be	international companies
counteract pressure	expensive	lead times	finding	do business		international currency

create value	extraordinary	Legal/consult cost/fee	focus	driven	internationally
creative	extremely	locations	focus (n)	eliminating	judgment
credentials	fire	logistics (facility/assets)	force	embarked	lack of financial control
curtailed (expense)	flat	loss	fountation	emergence	lawsuits
dedication, dedicate	global financial crisis	manager	framework	emphasis on	legal case
definitive	hampered	manufacturer	fundamentals	enable	Legal/consult cost/fee
deliberate	hard time	margin	goal	engage	legislative certainty and policy clarity
deliberations	high lead prices	market share	government initiatives	enter into agreement	litigation
delight	high level of uncertainty	material flow	growth area	establish, establishing, establishment,	look forward
deposit	high redundancy costs	minority interest	guidance	evaluation	major difficulties
develop grant	hinder	momentum	help	examine	market (condition)
development	hostile takeover bid	monetary impact	honour	exchange	market (economics)
diligence	hurdles	office	hope	execution	market
dividends	ill-time	option	human capital	execution, execute	market changes
drop frequency of incident	illusion	order	idea	expansion	market positioning
dynamic (company)	imbalance	ordinary share	impact (n)	export	market potential
EBIT, earning,	impairment	output	importance	exposure	market segments
effective	inappropriate	overheads	improvements	finalized, finalise,	marketplace depress
efforts (intensify efforts)	increase costs	owner	incentive	finance (v)	medical community
enabled	increased (competitive)	papers	indication	focus/focused on/focusing	movement
encouraging, encouraged	increased overheads	parameters	influence	fund	no upturn in industry
endeavour	inefficient	partner	initiative	generated, generating,	opportunities
endorsed	instability	patent	insight	governs	outside
endured (shareholder)	instalments	payable	instability	Hedge	policy matters
enhance the value	irregular	Payload	intellectual property	hedging	population

enlarged business	irritation	payment	investor interest	held	potential
enormous (market)	legal cost	payout	joint venture(relation)	identified	potential buyers
ensure	less	payout ratio	judgment	implant	primarily aimed at
enthusiasm, enthusiastic	liabilities	performance	knowledge	implement,implementation	process, process issues
environmental friendly	limited, limit	persent	leadership	improve	promises
equity	liquidated	personnel	licensing, licence,	increased	public awareness
equity capital	little value for shareholder	physicians	manufacturing base	incurred	rainfall
exceed (budget, profit target)	loan	plantation	market position	initiate	rampant
excellent, excellently	loosen.loosened	platform	method	initiative funding	resignations, resigns
exceptional (return, team)	losses	poised to take advantage	motivation	innovative, innovate, innovation	risk area
exciting (future/technology)	loss-making	portion	movements	install (base)	sentiment
expansion, expand	low, lower, (harvest)	practice	need	installation	settlement
experience	low, lower, (price)	preference share	network	integration, integrate,	severe winter
extend, extension	low, lower, (profit)	presence	objective	intended to	snow storms
extensive experience	massive	presence	occasion	investment, invest,	social
fall (cost, gearing)	negative	price	offer	issue	strikes
fast, faster	no (offers)	proceeds	operation	joined	tariffs
favourable	no upturn	processer	opportunity	launch	technical problems
firmly	no/nominal dividend	producer	outcomes	leverage	trading (condition)
flexibility	nor	product	ownership	lift	trading (environment)
focused(organisation/plan)	not be achieved	product lines	partnership	liquidating	transition
fulfill	not grow	production output	patience	loan	uncertain, uncertainly,
full (strength)	not offer any synergies	production rates	pattern	lodged	unclear (impact)
funding, funds	not out of the woods	productivity	plan	made contract	uncontrollable risks
get funding (Capral - 25)	not possible	profit / net profit	policy	make money	under appeal

good	not satisfactory	profit targets	potential	make offer	unsettled
great (opportunities)	not sound	profitability	prefeasibility	make progress	unstable financial
great (year)	not sufficient	property (assets)	pressure	manage	volatile
great breadth of experience	oversubscribed	publiçations	principles	manufacture	volatility (commodity)
great, greater, grateful	pain	range	priority	market (v)	volatility (currency)
growth, grow(acceptance)	payable	receipts	procedure	marketed (v)	volatility (market)
hard work	poor (financial performance)	Redeemable Floating Notes	process	measured	warranty
harvest	poor (planning)	remuneration	product range	meet demand	warranty claim
healthy	pressure	reseller	production tonnage/output	merger	weather
high level of demand	problems	resolution	program, programme	modified	weather-induced
high quality	profit down	resource	progress	monitored, monitor	world demand
high, highest, higher	protracted	result	progression	name change	worldwide market
highlight	rampant	retail stores	project	negotiate	
honours	recoupment	retain the funding	prominence	new chairman/MD	
ideally	reduce, reduction (debt)	retained earning	promise	new initiative	
implement strategy	regretted	retire debt	propose	new instrument	
Impressed/impressive capacity	re-imbursement	return	prospectus	new markets	
improvement, improve	rejection	revenue/sales revenue	qualifications	new OTI	
imputation credit	removal	Right Issue	quality	new sales	
in line	require better information	sales level	questions	new structure	
income	restriction	sales representative	recession	new technology	
increase output,	retire debt	service centres	regulatory	obeservation	
increase presence	retraction	share price	relationship	obsolete	
increased sales	retrained	shareholder	report	objective	
infers income	sadness	shareholder value	reputation	obtained	

inflow	sentiment	shareholders' funds	requirement	offer (v)
innovative, innovation	severely impacted	shareholders' loans	response	offsetting
insight	shortfall	shipment	responsibility	open
insignificant debt	shortfall in revenue	sourced	retrained	operate, operates
inspirational	shrinking	staff (engineer)	review	outsource
inspiring	significant time and cost	stations	rewarding	overview
integration	skewed	steady gains	safe	participation
interest	slow process	stock figure	sales cycle	partnered
invaluable	slowdown	stores	sales model	pay, paid,
investment	slower	strategic interest	scheme	performe
large sales	sobering (message)	subsidiary	service offering	plan
larger, largest	soft (market)	sufficiency	serviced, service	positioned
leader, lead the way, leadership	suffer	supplies (n)	signal	pre-development
leadership (position)	suspension	system (product)	situation	productive, production, produce
lean / leaner organization	take time to	target customer group	skill	program
less cash usage	tax payment	tax payment / taxation	standards	progress
leverage	technical problem	tax rebate	status	project
listed, listing,	threaten	tax refund	stewardship	promote
logistics assets	tightening, tight	team, team member,	strategic initiatives	prove
longer term gain	too long to achieve	third party	strategic roadmap	provide,provision
long-term value	tough	traffic	strategy	publish
low cost	trim margins	training centre	strenuous	purchase
low gearing	turbulence	valuation multiples	structure	pursue
loyal	turmoil	value (enterprise)	success	raise (money)
make money	unacceptably High	value recognition	supply	rationalisation

make progress	unavailable	volume	support	reach (goal)
margin	uncertainty, uncertain,	warranty claim	sustainability	reached
market acceptance	uncontrollable risks	work capital	system	realized
market penetration	under	workforce	targeted treatment	reassess
maximize return	unfavourable		task	recapitalising
meaningful (relationship)	unfortunate		technology	recommend
meet (all challenge)	unprofitable		tenure	record
meet current demand/need	unsatisfactory/no longer		testimony	recoup
met expectation	unsettled		tradition	recruitment, recruiting
minimal effect	unsustainable		transaction	re-election
minimize (adverse effect)	unwelcome		trend	re-establish
minimize (debt)	volatility		trial	re-establishing
mitigate adverse effects	weak		trust	re-exam
modest	withdraw business		uncertainty	refinancing
momentum	write off (tax benefits)		uncontrollable risks	refinancing
motivated team	write off (goodwill)		value chain	reflect
move_forward	yet achieved		viability	refocused
moving forward	yet to be resolved/unsolved		volatility	refreshment
nil borrowings			wish	registration
no interest cost				release
number 1				release
offset rising costs				relocate
on time				relocated
on track				removed
ongoing (support)				removed / removal

opportunities, opportunity	renewal
optimising	repayment
optimism	replace
orders	report
ordinary share	represents
outstanding (performance	Research and Development
overcame, overcome (bottle neck)	reshape
pacific and focused plan	re-shape
passionately	responded
pay tribute	responsibility
phased out the restriction	restore
pleased, pleasure,	restore to profitability
positive (contribution)	restructuring, restricted
positive milestones	result
precise	resume
preference share	return profitability
premium	revaluation
proceed	revert
production in full	review
productivity	roll out
profit improvement	sale, sell
profitable, profitability,	sales initiative
profits, net profit	scheduled
profound	search
progress	secured

progressed	seek
proper	send
property assets, property	serve, serving
prosper	service,
protect	set out
proud (history,	set up
prudent measure	settlement
prudent, prudently,	shipping
pursuing funding	shows
qualifications	stimulate
qualified person	strategic priorities
raising, arising, arise,(adj) (capital)	strategy initiative
raising, arise,(adj) (profit)	study
rapidly	supplies
reach (goal)	support
reach acceptable sales levels	survive
realistic (forecast)	sustain
receipt	take an action
record (profit)	take approach
recoup	take consideration
recovered, recovery,	targeted
recruitment	terminate
reduce (claims)	test
reduce (debt)	train
reduce (lead times)	transaction

reduce (pollution)	transferred
reduce (risk)	transformation
reduce (tax)	trials
reduce (the perceived risk)	underpin
reduce (time)	undertaken, undertake, undertook
reduce bank debt	upgrading, upgrade,
reduce personnel / labour	uptake
reduce concentration	used
reduce waste	voting
reduce/reduction cost	widen (the range)
reduced interest rates	wish
reduced net borrowings	withdraw
reduction of expense	working with
reduction of waste usage	write down
refined	write off
reinvigorate	yield
reinvigorating	
relist	
remarkable	
remarkable demand	
remunerate	
resilient	
respected	
restore to profitability	
result returned	

retain the funding
retained earnings
return to profitability
returns
revenue/sales revenue
reverse the loss
reward
rich with a wealth of ideas
and opportunities.
ride out these difficult
right fundamation
rise, raised capital
robust
safely, safe
sale, sales, sold
sales growth
satisfactory (result)
satisfactory (return)
secured
self sustaining
share
share distribution
shareholder support
shareholder value
shareholders' funds

shareholders' loans
sharper (focus)
skills, skilled
smoothly
solid (foundation)
sound, soundness
sound, soundness
soundness
stable
staff fully employed
steady gains
stimulate growth
streamlined
strides
strong commitment
strong demands
strong growth
strong partner
strong performance
strong presence
strong process
strong result
strong sale
strong track record
strong, stronger,

strengthened, strengthen, strength,
substantial, substantially
success
successfully, success, successful
sufficient
sufficient cash
suitable
superior
support better
support customers
support employees
support growth
support ongoing
support overwhelming
surpassed
surplus
sustain well
sustained growth
take advantage of
take ownership
talented
tax rebate
tax refund
tightening (supply)

traction
transparent
triumph
turnover
unanimously recommends
under control
unique
unlock value
unprecedented level of efficiency
untiring
unwavering
up (profit)
upgrading, upgrade,
upside
upturn
useful
Value Creating
valued,valuable
shareholder
vastly
viable
watershed
wealth
well ahead of plan
well experienced

well know
well matched
well performance
well placed
well positioned
well served
well-defined strategy
well-shaped
wider range of technologies
winning new business
wise counsel
wise guidance
within (budget)
withstand force/pressure
won awards
world class

Appendix B: Correlation coefficients between all the 923 variables and corporate performance

	Variable name	Coefficient correlation with performance		Variable name	Coefficient correlation with performance
1.	high.highest.higher_TOTAL	-.533**	2.	#beneficial#_standard.quality.price	-.524**
3.	dividend	-.432**	4.	goodwill	.391**
5.	consolidate.consolidation	.365**	6.	earning.EBIT.EBT	-.362**
7.	low.lower_TOTAL	-.354*	8.	finalized.finalise	.346*
9.	gain	-.342*	10.	service#intangible	-.342*
11.	successfully.success.successful	.315*	12.	increase#adverse#_cost.pressure	-.313*
13.	ongoing	-.313*	14.	returns#tangible	-.313*
15.	volume.number	-.313*	16.	increase_TOTAL	-.302*
17.	writeoff_goodwill.taxbenefits	.297*	18.	not_be_achieved.to_long_to_achieve.yet_achieved	.297*
19.	direction	.297*	20.	revaluation.evaluation.	.297*
21.	implement.implementation	-.286*	22.	product	.286*
23.	cost_base.cost#tangible	-.283*	24.	low.lower#beneficial#_gearing.cost	-.283*
25.	hard_work	-.283*	26.	design	-.283*
27.	no_nominal_dividend	-.283*	28.	offset_costs	-.283*
29.	increase#beneficial#_output.presence.sales	-.276	30.	growing.growth.grow	-.274
31.	execution.execute.executive	.267	32.	result	-.266
33.	effect.effectiveness.efficiency#intangible	-.264	34.	source.resource.	.263
35.	improvement.improved_competitive.efficiency.output.position	-.258	36.	funding.funds#dynamic	.256
37.	technology	.255	38.	nor.not.yet (total)	.254
39.	goal	.251	40.	recovered.recovery.	.251
41.	approach#intangible	-.251	42.	ratios	-.251
43.	unfavourable	-.251	44.	steady#static	-.251
45.	listed.listing	-.251	46.	tightening.tight	-.251
47.	management	.250	48.	new (total)	-.249
49.	advanced. advancing	.249	50.	Expected	.249
51.	demand	-.246	52.	consistent.consistently	-.244
53.	belief.believe	.242	54.	focuse.focusing.refocus#dynamic	.241
55.	advisors	.240	56.	examine.reexame	.240
57.	aggressively	.240	58.	foresee.foreseeable	.240
59.	ahead	.240	60.	group	.240
61.	aligned_to.align	.240	62.	impairment	.240
63.	attention	.240	64.	manager	.240
65.	bank_guaranteed_notes_payabl	.240	66.	tax.taxation#beneficial#rebate.refund	.240

Variable name	Coefficient correlation with performance	Variable name	Coefficient correlation with performance
e.bankers_debt.support			
67. cash_inflow	.240	68. not_satisfactory	.240
69. cash_position	.240	70. regulatory	.240
71. certainty	.240	72. reimbursement	.240
73. course	.240	74. survive	.240
75. divestment	.240	76. manufacturer	.240
77. release	.240	78. stabilized	.240
79. satisfactory_return.result	.240	80. value_down	-.240
81. expected.expectation.expects	.237	82. contract	-.235
83. report.reporting	-.230	84. low.lower (adverse)	-.230
85. negative	-.226	86. continue.continuing.continuation	.222
87. future	.221	88. commitment	-.221
89. accretive	-.215	90. remunerate.remuneration	-.215
91. environmental#unexpected	-.215	92. represents	-.215
93. finance.financial.refinance	-.215	94. scheme	-.215
95. #adverse#	-.215	96. train	-.215
97. impressed.impressive_capacity	-.215	98. transferred.transformation	-.215
99. momentum	-.215	100. yield	-.215
101. payout	-.215	102. Principles	-.215
103. presence	-.215	104. rates	-.215
105. performance	-.214	106. interest#tangible	.212
107. Report_size	-.212	108. market#unexpected	-.211
109. seek	.206	110. export	.206
111. exceed#beneficial	.206	112. proceed	.206
113. profitable.profitability.profit	-.205	114. agreement	.205
115. difficult period_hard.time.year	.204	116. EM_positive	-.204
117. benefits#tangible	-.201	118. acceleration.accelerate	.200
119. capitalise.recapitalise_raising.demand	.200	120. acceptance.acceptable	.200
121. economy.economic	-.200	122. application	.200
123. exciting.excited#beneficial	.200	124. credentials	.200
125. look_for.forward	.199	126. upgrading.upgrade	.199
127. gearing_ratio/level___leverage	-.197	128. strong.stronger.strongth.strengthen#beneficial	-.195
129. position#tangible	-.196	130. industry	-.195
131. deteriorating.deterioration	.194	132. take_consideration.consider#dynamic	.193
133. global.globally.oversea.	.191	134. dedication.dedicate	-.189
135. reach_goal.acceptable	.188	136. remain.remainder.remaining	-.187
137. weak.weakening	-.187	138. your	.186
139. retain.retaining_funding.earnings.business	.184	140. expenditure.expense.fee	.183
141. solid	-.182	142. you	.181

Variable name	Coefficient correlation with performance	Variable name	Coefficient correlation with performance
143. I	-.181	144. EPS	-.177
145. contribute.contribution.	-.175	146. competency.competence#intangible	-.174
147. prospect.prospective.prospecte	.174	148. congratulate	-.174
149. issue#tangible	-.174	150. constrain	-.174
151. review	-.174	152. corporate governance	-.174
153. able	.174	154. convert.conversion	-.174
155. accountability	-.174	156. cut	-.174
157. action	-.174	158. determine	-.174
159. bids	-.174	160. received	-.174
161. buoyant	-.174	162. reduce_borrowing	-.174
163. centres	-.174	164. reduce_pollution.	-.174
165. charge	-.174	166. reduce_time.leadtime	-.174
167. eliminating	-.174	168. reinvigorate.reinvigorating	-.174
169. faster	-.174	170. robust_performance	-.174
171. flexibility	-.174	172. unsettle	-.174
173. imputation_credit	-.174	174. skilled#beneficial	-.174
175. income	-.174	176. smoothly#beneficial	-.174
177. insight	-.174	178. streamlined	-.174
179. issue#dynamic	-.174	180. supplier	-.174
181. legal_case	-.174	182. transparent_governance	-.174
183. lift	-.174	184. turnover	-.174
185. loyal	-.174	186. unfortunate	-.174
187. minimize_debt	-.174	188. warranty	-.174
189. network	-.174	190. wise#beneficial	-.174
191. new_facility	-.174	192. ownership	-.174
193. new initiative	-.174	194. realized	-.174
195. open	-.174	196. facility	-.172
197. support#beneficial	.172	198. abnormal	.168
199. accept#dynamic	.168	200. arrangement	.168
201. accommodate	.168	202. assess.assessing.	.168
203. adjust.make_adjustment(dynamic)	.168	204. augur	.168
205. allows	.168	206. be_looking_to_do	.168
207. take_approach#dynamic	.168	208. legislative	.168
209. bolstered	.168	210. counsel	.168
211. bottleneck	.168	212. coverage#tangible	.168
213. business_culture	.168	214. definitive	.168
215. capture	.168	216. deferred	.168
217. cash_ourflow.cash_usage	.168	218. delight	.168
219. cash_retention	.168	220. destroy	.168
221. circumstances	.168	222. dire	.168

Variable name	Coefficient correlation with performance	Variable name	Coefficient correlation with performance
223. close_subsidiary	.168	224. discovered	.168
225. commend#dynamic	.168	226. disease	.168
227. comment_on	.168	228. division	.168
229. commissioned	.168	230. do_business	.168
231. committee_meeting	.168	232. dreadful_financial_performance	.168
233. concession	.168	234. drop_frequency_ of_ incident	.168
235. concluded	.168	236. earthquakes	.168
237. construction	.168	238. efficacy	.168
239. consummated	.168	240. element	.168
241. contingent_upon	.168	242. emergence	.168
243. discontinued	.168	244. endorsed	.168
245. lack_control.uncontrollable	.168	246. endured	.168
247. evolution	.168	248. enormous_market	.168
249. exacerbated	.168	250. eventful	.168
251. expertise	.168	252. material flow	.168
253. exploration	.168	254. motivated#beneficial	.168
255. feasibility.prefeasibility	.168	256. name_change	.168
257. franchise_group	.168	258. new_equipment	.168
259. full_strength	.168	260. no_interest_cost	.168
261. fulfill	.168	262. no_upturn	.168
263. fundamentals	.168	264. not_possible	.168
265. hampered	.168	266. not_sustainable.unsustainable	.168
267. healthy	.168	268. not_sound	.168
269. hedge.hedging	.168	270. obtained	.168
271. held	.168	272. occasion	.168
273. help	.168	274. outsource.outsourcing	.168
275. hope	.168	276. owner.	.168
277. hurdles	.168	278. spacific#beneficial	.168
279. idea	.168	280. parameters	.168
281. ideally	.168	282. partnered	.168
283. identified	.168	284. patience	.168
285. individual	.168	286. pattern	.168
287. insurance	.168	288. perceived	.168
289. interested_parties	.168	290. percent	.168
291. irritating	.168	292. personnel	.168
293. judgment	.168	294. plantation	.168
295. liabilities	.168	296. make_money	.168
297. platform	.168	298. retraction	.168
299. practice	.168	300. revert	.168
301. premium	.168	302. rich	.168

Variable name	Coefficient correlation with performance	Variable name	Coefficient correlation with performance
303. Product range	.168	304. right#beneficial	.168
305. producer	.168	306. roll_out	.168
307. propose	.168	308. service#dynamic	.168
309. prosper	.168	310. sobering_message	.168
311. publications	.168	312. unstable	.168
313. purchaser	.168	314. staff_fully_employed	.168
315. raising.arising.arise#adverse	.168	316. status	.168
317. rampant	.168	318. stay_the_same	.168
319. rationalisation	.168	320. stewardship	.168
321. recession	.168	322. stores	.168
323. recoup.recoupment	.168	324. strides	.168
325. reduce_personnel.staff	.168	326. study	.168
327. relentless	.168	328. sufficient	.168
329. Research_and_Design	.168	330. task	.168
331. respected	.168	332. testimony	.168
333. response.respond.	.168	334. threaten	.168
335. restore_to_profitability	.168	336. transaction	.168
337. resume	.168	338. transition	.168
339. turbulence	.168	340. trim_margins	.168
341. under_appeal	.168	342. will_be	.168
343. unique	.168	344. without_redundancies	.168
345. untiring	.168	346. write_down	.168
347. watershed	.168	348. unlock_value	.168
349. wealth.	.168	350. BENEFICIAL	-.167
351. snow storms	.168	352. decision	.165
353. fund#tangible	-.166	354. activity.activities.	.161
355. TOTAL_fall.shortfall.reduce.decrease.decline	-.163	356. cost#beneficial#_savings.reduce.decrease.reduction.less.cost_effective	.161
357. new product	-.161	358. uncertainty.uncertain,	-.160
359. ADVERSE	-.161	360. cash (total)	.159
361. management#expected#replace. transform.change.appoint.retire	-.159	362. competitive.competitiveness.competitively.	-.157
363. worldwide	-.159	364. guidance.guide	-.154
365. ensure	-.157	366. benefit.benefited	-.153
367. achieved#dynamic.beneficial	.156	368. capacity#tangible	-.153
369. reduce_profit	-.149	370. safety.safe	-.148
371. Tangible	-.149	372. pay.paid.repaid	-.145
373. EM_TL	-.146	374. effective.efficient#beneficial	.145
375. margin	.145	376. structure.restructuring.restructure	-.144
377. move.movement	-.145	378. underpin	-.143

Variable name	Coefficient correlation with performance	Variable name	Coefficient correlation with performance
379. instal.instalment.installation	-.143	380. undertaken.undertake.undertook	-.143
381. relocate	-.143	382. investment.reinvestment#tangible	-.141
383. achievment#intangible	-.141	384. board.committee	.138
385. pleased.pleasure.pleasingly	-.139	386. equipment	-.138
387. weather_induced.climate.	-.138	388. leverage	-.138
389. maximize	-.138	390. establish.establishment.reestablish	.138
391. perform#dynamic	-.138	392. deposit	.133
393. reflect.reflection	.133	394. directors	.133
395. stable.	.133	396. investor	.133
397. withstand#beneficial.dynamic	.133	398. joint_venture （relationship）	.133
399. assessment	.133	400. patent	.133
401. crisis.financialcrisis.capital_crisis.global_financial_crisis	.133	402. move#beneficial#_forward.forward_strategy	.133
403. concentrated	.133	404. target#dynamic	.132
405. promote	.133	406. standards	-.132
407. talented	.132	408. positioned#dynamic	-.128
409. leader.leading.lead_the_way.leadership	-.129	410. attract.attractive	.127
411. change	-.127	412. suffer	.126
413. limited.limit	.126	414. world_class.first	.126
415. manufacturing.manufacture	.126	416. operation#intangible	.124
417. development.developing.predevelopement#intangible	.124	418. control.under_control.take_over.take_ownership	-.123
419. bad_weather (total)	-.123	420. 2nd_PRON	.123
421. investigation.invest.reinvest#dynamic	-.123	422. advantage.take_advantage_of.take_full_advantage	-.122
423. ably	-.122	424. assure	-.122
425. acumen	-.122	426. avaliability	-.122
427. adaptations	-.122	428. backdrop	-.122
429. additional	-.122	430. base	-.122
431. disadvantage	-.122	432. big	-.122
433. advice.	-.122	434. BSE	-.122
435. affected	-.122	436. calculate	-.122
437. ambitions	-.122	438. closure.disposal	-.122
439. ameliorated	-.122	440. containment	-.122
441. inappropriate	-.122	442. cope_with	-.122
443. aspirations	-.122	444. counteract	-.122
445. combine	-.122	446. court_action	-.122
447. commendable#beneficial	-.122	448. creative	-.122
449. commodities	-.122	450. damaged	-.122
451. communications	-.122	452. deal_with	-.122

Variable name	Coefficient correlation with performance	Variable name	Coefficient correlation with performance
453. compensation	-.122	454. dealership	-.122
455. complex	-.122	456. debate	-.122
457. comprehensive	-.122	458. define	-.122
459. conscious_extremely	-.122	460. defence.defended	-.122
461. constitution	-.122	462. held_up	-.122
463. consumption	-.122	464. deliberate.deliberation	-.122
465. demographic	-.122	466. envisage	-.122
467. difficulty (adj)	-.122	468. eroded	-.122
469. discipline	-.122	470. escalate	-.122
471. dislocation	-.122	472. exposure	-.122
473. distracted	-.122	474. extensive_experience	-.122
475. double	-.122	476. extraordinary	-.122
477. inefficient	-.122	478. factors	-.122
479. embarked	-.122	480. firmly	-.122
481. engage	-.122	482. foundation	-.122
483. gratitude	-.122	484. framework	-.122
485. honour	-.122	486. ill_time	-.122
487. hostile_takeover_bid	-.122	488. illusion	-.122
489. human_capital	-.122	490. imbalance	-.122
491. incentive	-.122	492. implant	-.122
493. inflationary.inflation	-.122	494. in_line	-.122
495. influence	-.122	496. market_place	-.122
497. inspirational.inspiring	-.122	498. matched	-.122
499. intake	-.122	500. meaningful	-.122
501. intellectual_property	-.122	502. minimize.minimal#beneficial#	-.122
503. irregular	-.122	504. minimize_adverse_effect	-.122
505. item	-.122	506. modified	-.122
507. lack	-.122	508. new_distribution	-.122
509. lawsuits	-.122	510. new_system	-.122
511. lead_times	-.122	512. new instrument	-.122
513. lean.leaner_organization	-.122	514. new sales networks	-.122
515. liquidating.liquidated	-.122	516. new_share	-.122
517. locations	-.122	518. new structure	-.122
519. loosen.loosened	-.122	520. new technology	-.122
521. new_manufacturing	-.122	522. nil_borrowings	-.122
523. new_vehicle	-.122	524. no_shrink	-.122
525. new_model	-.122	526. nor	-.122
527. new OTI	-.122	528. yet_to_be_resolved.unsolved	-.122
529. non_cash_item	-.122	530. population	-.122
531. cannot (total)	-.122	532. portion	-.122

Variable name	Coefficient correlation with performance	Variable name	Coefficient correlation with performance
533. cannot_recover	-.122	534. preform	-.122
535. number_1	-.122	536. prepare	-.122
537. obeservation	-.122	538. premise	-.122
539. obsolete	-.122	540. prevent	-.122
541. office	-.122	542. priority	-.122
543. order	-.122	544. procedure	-.122
545. output	-.122	546. production_rate	-.122
547. outside	-.122	548. prominence	-.122
549. pain	-.122	550. proper	-.122
551. pay_tribute	-.122	552. protect	-.122
553. payable	-.122	554. protracted	-.122
555. payload	-.122	556. publish	-.122
557. plan#dynamic	-.122	558. range	-.122
559. realistic_forecast	-.122	560. profit_down	-.122
561. receipt	-.122	562. reduce_concentration	-.122
563. reconciliation	-.122	564. reduce_claim	-.122
565. recruitment.recruiting	-.122	566. reduce_tax	-.122
567. reduce_workforce	-.122	568. reduce_loss	-.122
569. reduce_harvet	-.122	570. reduce_water	-.122
571. refined	-.122	572. reduce_consumption	-.122
573. refreshment	-.122	574. shrinking	-.122
575. regretted	-.122	576. skewed	-.122
577. relist	-.122	578. social	-.122
579. removed.removal	-.122	580. soft_market	-.122
581. renewal	-.122	582. sourced	-.122
583. reputation	-.122	584. Static	-.122
585. reseller	-.122	586. instability	-.122
587. resilient	-.122	588. stimulate#beneficial	-.122
589. restate	-.122	590. strenuous	-.122
591. restricte.restriction	-.122	592. sufficiency	-.122
593. retrained	-.122	594. superior	-.122
595. sales_cycle	-.122	596. supply_alliance	-.122
597. sadness	-.122	598. surplus	-.122
599. scheduled	-.122	600. take_action	-.122
601. send	-.122	602. tariffs	-.122
603. set_out	-.122	604. tenure	-.122
605. shows	-.122	606. terminate	-.122
607. surpassed	-.122	608. tradition	-.122
609. value chain	-.122	610. traffic	-.122
611. vastly	-.122	612. transilation	-.122

Variable name	Coefficient correlation with performance	Variable name	Coefficient correlation with performance
613. versatile	-.122	614. trend	-.122
615. climate_change	-.122	616. triumph	-.122
617. ice storm	-.122	618. tumble	-.122
619. severe winter	-.122	620. unacceptable	-.122
621. withdraw#adverse	-.122	622. unanimously_recommends	-.122
623. unprofitable	-.122	624. unclear	-.122
625. little_value_for_shareholder	-.122	626. under	-.122
627. unwelcome_and_opportunistic_bid	-.122	628. progress.progression#beneficial.intangbile	.121
629. partner	.121	630. cancellation	.119
631. search	.119	632. consequence	.119
633. invaluable	.119	634. focused#beneficial	.119
635. badly.worse.worst	.119	636. option	-.119
637. prudent.prudently	-.119	638. good	-.115
639. outcomes	.116	640. larger.largest	-.115
641. deliver.delivering,	-.115	642. need	-.110
643. skills#intangible	.111	644. up_profit	-.110
645. me	-.110	646. conditions	-.109
647. declare.declaration	-.110	648. loss	-.107
649. cost#adverse#costly_pressure.rising.increase.high	-.107	650. encouraging.encouraged	.106
651. cash.cash_flow.equivalent	.105	652. committed	-.104
653. adjustment#intangible	-.104	654. emphasis	-.104
655. below#adverse	-.104	656. wider.widen#beneficial	-.104
657. brand	-.104	658. price	-.102
659. excellent.excellently	-.102	660. share#tangible	.099
661. revenue.sales_revenue.#beneficial.tangible	-.099	662. more.most	-.098
663. adopted.adoption	-.098	664. pursue.pursuing_funding	-.098
665. amortization.depereciation	-.098	666. readability	.098
667. currency	-.098	668. impact	.098
669. joint_venture (parties)	-.098	670. positive	.098
671. debt	-.095	672. well#beneficial	-.095
673. disappointment.disappointing	-.095	674. stock.storage.inventory.feedstock	-.093
675. serve.serving	-.095	676. tax.taxation#adverse#_payment	-.093
677. address	-.093	678. tough	-.093
679. depress	-.093	680. licensing.licence	-.093
681. in/make_process.processing	-.092	682. appreciation	.092
683. diligence.diligent	.092	684. Unexpected	.088
685. Beneficial	-.091	686. produce	.088
687. policy.political	-.086	688. staff.engineer	.084

Variable name	Coefficient correlation with performance	Variable name	Coefficient correlation with performance
689. awarded.won_award.reward	.084	690. add	.082
691. difficulty (n)	-.082	692. generated.generating.regenerate	.082
693. fall.falling_price.gearing	.082	694. completed.completion.complete	-.081
695. project	-.081	696. quality	.079
697. ability	.079	698. product lines	.077
699. plan#intangible.expected	-.078	700. production	.077
701. place.replace#dynamic	.077	702. initiate#dynamic	-.075
703. International_TOTAL	.076	704. milestones	-.075
705. shareholder	-.075	706. better.best	-.074
707. system	.075	708. provide.provision	.072
709. reduce_revenue	-.073	710. aware.awareness	-.070
711. acquisition.acquiring.acquired	-.070	712. commenced.commencing#dynamic	.068
713. develop#dynamic	-.069	714. anticipate	-.067
715. cost_TOTAL	-.068	716. Adverse	-.066
717. favourable	.067	718. no.non.nil. (total)	-.062
719. distribution.distributor	.062	720. reduce_debt	.060
721. maintain.maintaining,	.061	722. resolution.solution.	.060
723. reduction	-.060	724. rights#tangbile	.060
725. responsibility	-.060	726. drop	.060
727. overcame.overcome.ride_out	.060	728. program.programme	.059
729. partnership	.060	730. building.built	.057
731. experienced.experience	-.058	732. relationship	.057
733. announced	.057	734. pressure	-.055
735. strategy	-.056	736. focus#intangible	-.055
737. meet_all_challenge.demand.ne ed.expectation	.055	738. management#unexpected#transform .change.resignation.resign.departure	-.055
739. payment.repayment	.054	740. rainfall.	-.051
741. diversification.diversify	.054	742. adequately	-.051
743. exchange_rate	.052	744. capable#beneficial	-.051
745. Static	.051	746. grant.guarantee	-.051
747. record_profit	-.051	748. interest_rates	-.051
749. further	-.051	no_significant_changes.unchanged	-.051
750. market_segment	-.051	751. joint_venture.TOTAL	-.050
752. government.financial_stimulus_package	-.050	753. enlarged.expansion.expand.extend.extension	.048
754. integration.integrate	-.049	755. rapidly	.048
756. foreign_exchange.	.048	757. model	-.046
758. indicate.indication.indicator	.048	759. progressed.progress#dyanmic	-.045
760. Dynamic	-.045	761. participation	-.044
762. break_even	-.044	763. workforce	-.044
764. claim	-.044	765. require.requirement_information	-.044

Variable name	Coefficient correlation with performance	Variable name	Coefficient correlation with performance
766. discussion	-.044	767. supply	-.044
768. enthusiasm.enthusiastic_acceptance	-.044	769. productive.productivity.production_in_full.	-.044
770. settlement.	-.044	771. enhance_the_value	.044
772. slowdown#adverse	-.044	773. PRON_TL	-.044
774. our	-.044	775. overheads	.042
776. control (total)	-.042	777. 1st_PRON	-.039
778. EM_negative	.040	779. aim	-.038
780. Intangible	-.038	781. broad.broaden#beneficial	-.038
782. announcement	-.038	783. participation	-.044
784. participation	-.044	785. delay	-.038
786. competitor	-.038	787. dispute	-.038
788. concern	-.038	789. figure	-.038
790. constant_currency	-.038	791. penetration	-.038
792. knowledge	-.038	793. problems	-.038
794. monitored.monitor	-.038	795. property_assets.property	-.038
796. new business	-.038	797. sales_representative	-.038
798. qualifications.qualified_person	-.038	799. shape.reshape	-.038
800. recommend	-.038	801. slow.slower#adverse	-.038
802. trial	-.038	803. advise	.037
804. external#unexpected	-.037	805. enter_into_contract.agreement	.037
806. actively	.037	807. cash_generation	.037
808. appropriate	.037	809. clear#beneficial	.037
810. assist	.037	811. business_line	.037
812. assume	.037	813. competent#beneficial	.037
814. boost	.037	815. excess	.037
816. contraction	.037	817. fluctuating	.037
818. dealer	.037	819. force	.037
820. decide.	.037	821. harvest	.037
822. disruption.disrupt	.037	823. no_offers	.037
824. legal/consulting costs/legal/fee	.037	825. no_longer_satisfact.unsatisfactory	.037
826. logistics	.037	827. infrastructure	.037
828. market#dyamic	.037	829. not_sufficient.insufficient	.037
830. merger.merge	.037	831. on_time	.037
832. mitigate	.037	833. on_track	.037
834. new trial sites	.037	835. risk	.037
836. oversee	.037	837. severely_impacted	.037
838. precise	.037	839. reduce_expense	.037
840. predict	.037	841. reduce_risk	.037
842. processer	.037	843. reduce_intrestrate	.037
844. production_tonnage.output	.037	845. viable.viability	.037

Variable name	Coefficient correlation with performance	Variable name	Coefficient correlation with performance
846. suitable.suited	.037	847. winning#beneficial	.037
848. take_time	.037	849. within_budget	.037
850. used.useful	.037	851. capability#intangible	-.033
852. client.customer	-.036	853. volatile.volatility	.031
854. offer	-.032	855. initiative#intangible	-.031
856. BIG_WORDS	.031	857. consumer_preferences	.029
858. Bode	.029	859. device	.029
860. bright	.029	861. endeavour	.029
862. cash_reserves	.029	863. estimate	.029
864. concept	.029	865. exceptional_return.team	.029
866. confirmed	.029	867. prove	.029
868. information	.029	869. notes#tangible	.029
870. intended.intention	.029	871. remarktable	.029
872. less	.029	873. shipping.shipment	.029
874. promise	.029	875. turmoil	.029
876. signal.sign	.029	877. upturn.upward	.029
878. situation	.029	879. effort	.029
880. suspension.suspended	.029	881. confidence.confident	-.028
882. difficulty (total)	.028	883. great.greater.greatful	.027
884. asset	-.027	885. highlight	.026
886. budget	.026	887. us	-.025
888. downturn.downward	.026	889. optimising.optimism	-.024
890. demonstrates.demonstration	-.024	891. enable	-.024
892. adverse	-.024	893. target#intangible	.021
894. team.teammember	.022	895. negotiate	-.019
896. loan	-.019	897. we	.018
898. opportunities.opportunity	.014	899. tax.taxation_TOTAL	-.013
900. outstanding_performance	-.013	901. operated.operates.operating	-.013
902. value.carrying/face_value.share holder_fund#tangible.beneficial	.015	903. international.internationally_company.currency.presence.alignment	.012
904. challenging.challenges	-.012	905. poor#adverse	-.008
906. forecast	.010	907. purchase	-.008
908. manage	-.008	909. return#beneficial	-.008
910. outlook	-.008	911. employee	-.007
912. substantial.substantially	-.008	913. reduce_cost	-.007
914. creating.create	-.007	915. value (enterprise).valued.valuable	.003
916. raising.arising.arise#beneficial#	-.005	917. proud_history	-.003
918. subsidiary	.003	919. sustain.sustainability	-.003
920. sound.soundness	.003	921. equity.equity_capital	-.002
922. capital#tangible	.002	923. sale.sales.sold#dynamic	.002