English for Specific Purposes 80 (2025) 126–139

Contents lists available at ScienceDirect

English for Specific Purposes

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

Frame-based semantic patterns in business discourse: A case study

Yubin Qian [a], Hans C. Boas [b], Qun Zheng [a,*]

[a] Department of Foreign Languages, University of Chinese Academy of Sciences, No. 19A, Yuquan Road, Shijingshan District, Beijing, China

[b] Department of Germanic Studies, The University of Texas at Austin, 2505 University Ave, C3300, Austin, TX, USA

A R T I C L E I N F O

Article history:

Keywords:
Semantic frame
Semantic dimension
Semantic pattern
FrameNet
Business English

A B S T R A C T

Semantic analysis is a context-based activity that requires a detailed level of language analysis. This paper focuses on the application of the methodology underlying the FrameNet database in the process of decoding word meanings in business discourse. By focusing on the Management Discussion and Analysis text from the 3M Corporation as a case study, we identified four major semantic dimensions: “Financial Metrics”, “Operation and Business”, “Time and Duration”, and “Legal and Ethical Considerations”. In each of these dimensions, we found dominant semantic patterns closely related to those top frames that highlight essential aspects of corporate communication from financial reports to ethical statements. This article demonstrates how FrameNet can serve as a useful tool to kindle the unexposed linguistic features of business texts and support linguistic analysis of business discourse, which in turn can enhance the teaching of Business English by deepening students’ understanding of complex business vocabulary and context.

1. Introduction

Semantic pattern refers to recurring structures or templates in language that convey specific meanings or concepts (Stubbs, 2001). These recurring patterns reveal the mechanisms by which meaning is interpreted (Goddard, 2012), improve natural language processing systems (Shawar & Atwell, 2005), and explain the cognitive mechanisms involved in language use (Doumen, Beuls, & Van Eecke, 2023). Typically, they are represented by a sequential combination of high-level meaning elements. For example, in sentences “Sales increased by 10 % in the last quarter” and “Weighted-average diluted shares declined 0.1 percent in 2017”, meaning elements “sales” and “weighted-average diluted shares” are metrics for corporate earnings, thus can be abstracted to the [Earnings] high-level meaning element; “increase” and “decline” reflect [Change_direction]; “10 %” and “0.1 percent” are [Change_amount], while “last quarter” and “2017” denote time [Unit]. We propose that the recurrence of these meaning elements shared by both sentences form a systematic pattern “[Earnings][Change_direction][Change_amount][Unit]”, which reveals measurable characteristics of an entity over time. Identifying these elements to abstract a semantic pattern requires an integrated approach that combines both quantitative and qualitative methods, as detailed in the methodology section below.

* Corresponding author.
E-mail addresses: qianyubin@ucas.ac.cn (Y. Qian), hcb@mail.utexas.edu (H.C. Boas), zhengqun@ucas.ac.cn (Q. Zheng).

https://doi.org/10.1016/j.esp.2025.08.002
0889-4906/

In the business context, the rhetorical and pragmatic functions of corporate disclosure texts have been widely studied. For instance, rhetorical move analyses have uncovered typical discursive strategies used to greet participants, review financial performance, and manage discussion in earnings calls (Crawford-Camiciottoli, 2010). Thematic analyses have also highlighted key recurring content areas such as financial results, risk factors, and strategic positioning (Nelson, 2006; Ren & Lu, 2021). Sentiment-based or keyword-driven studies, such as those reviewed by Beattie (2014), have further illuminated the tone and evaluative stance adopted by firms in financial narratives.

Our study shares common ground with these prior efforts in recognizing the central role of dimensions such as "Financial Metrics" and "Operations and Business", which align closely with the communicative purposes of Management Discussion and Analysis (MD&A), namely, to report, analyze, assess, and forecast business performance (Ren & Lu, 2021). In fact, the semantic patterns identified in our data often mirror rhetorical moves found in existing literature, such as "Description of operating results", "Presentation of liquidity and capital resources", and "Description of contractual obligations" (Qian, 2020, p. 429). This convergence suggests that semantic and rhetorical structures can jointly illuminate the functional architecture of business texts.

However, our study also differs from existing research in several key respects. First, it applies a FrameNet-based approach to semantic analysis, grounding the study in a systematic lexicographic framework that links meaning to grammatical form and discourse structure (Fillmore, Johnson, & Petruck, 2003). Second, it redefines FrameNet frames in light of business usage, creating a tailored analytical lens that retains theoretical rigor of semantic frames while adapting to domain-specific communicative functions (see Boas, Ruppenhofer, & Baker, 2024). Third, instead of focusing on isolated linguistic features (e.g., keywords, sentence polarity), our study identifies and analyzes semantic dimensions and frame-based patterns, which are structured groupings of meaning elements that reflect the deeper conceptual logic of MD&A discourse.

For fields such as Business English and ESBP, this kind of semantic approach offers a valuable complement to existing pedagogical models. As noted by Shin (2020) and Trinder (2013), there is a growing interest in instruction that is both genre-sensitive and cognitively informed. Traditional approaches often emphasize rhetorical moves or fixed functional expressions, focusing primarily on surface-level realizations. In contrast, a frame-based semantic analysis reveals deeper conceptual structures that organize business discourse. By identifying interpretable, reusable semantic patterns grounded in real-world texts, this method supports more flexible and transferable language instruction by providing semantic templates such as [Earnings][Change][Explanation] that help learners understand and organize professional discourse beyond fixed expressions like "profits rose slightly".

The MD&A discourse, in particular, offers a valuable test case. It is one of the most linguistically rich and strategically important sections of financial reporting (Li, 2010). Its informational and persuasive nature reflects the full spectrum of corporate narrative strategies from quantitative reporting to qualitative positioning, with carefully chosen language that subtly conveys the company's values and mission (Qian & Wu, 2025). By focusing on semantic patterns in MD&As, we not only gain insight into corporate communication practices but also generate pedagogical tools, such as semantic templates and frame-based instructional tasks, for helping students learn to write within this high-stakes professional genre.

Given this motivation, our study aims to investigate the semantic patterns that characterize MD&As, organizing these patterns into broader thematic categories known as semantic dimensions. Using the FrameNet lexicographic database for English (Fillmore et al., 2003) as the primary analytical tool, the analysis focuses on the 3M Corporation's MD&A, chosen randomly from a representative corpus. The research questions guiding our study are as follows.

(1) How do semantic dimensions manifest themselves in the MD&A?
(2) What are the predominant semantic patterns associated with each identified semantic dimension?
(3) How do these semantic patterns align with the linguistic and pragmatic characteristics specific to business communication?

2. FrameNet and semantic patterns

Drawing on the principles of Frame Semantics (Fillmore, 1982; Fillmore & Baker, 2015), FrameNet (https://framenet.icsi.berkeley.edu) was initially developed as a lexicographic resource to structure the lexicon of English. While not originally conceptualized for natural language processing, FrameNet has since been adapted and applied to various NLP tasks, such as semantic role labeling (Gildea & Jurafsky, 2002) and information extraction (Mohit & Narayanan, 2003).

In contrast to other semantic approaches like WordNet or Natural Semantic Metalanguage, which primarily focus on word sense relations or universal semantic primitives, FrameNet captures meaning in terms of semantic frames, usually defined as "a script-like conceptual structure that describes a particular type of situation, object, or event along with its participants and props" (Ruppenhofer, Ellsworth, Schwarzer-Petruck, Johnson, & Scheffczyk, 2016, p. 7). Each frame in FrameNet is associated with a set of lexical units (LUs), which are the words and phrases that evoke the frame, and high-level meaning elements, or frame elements (FEs), which are situation-semantic roles that participants and props play in the frame (Boas, 2020). This detailed lexico-semantic and valence information contained in FrameNet facilitates more accurate and contextually appropriate language understanding and processing across diverse domains and applications (Boas, Ruppenhofer, & Baker, 2024).

To illustrate how frames are structured in FrameNet, consider the "Cooking_creation" frame. This frame describes a situation in which a [Cook] creates a [Produced_food] using [Ingredients]. The frame definition in FrameNet provides a detailed explanation of the frame and its associated FEs:

> "The [Cook], [Produced_food] and [Ingredients] are all present, and the [Ingredients] are typically transformed, either by being heated or mixed together (or both). The [Heating_instrument] and/or [Container] may also be specified." (FrameNet, 2024)

In the above definition of FEs, each FE is defined in terms of its semantic role within the frame and its syntactic realization in the context of the LUs that evoke the frame. For example, the FE [Cook] is the person who produces the FE [Produced_food] from the FE [Ingredients], and it is typically realized as the subject of LUs that evoke the frame, such as "bake", "cook", "fry", and "grill" (FrameNet, 2024).

The workflow underlying FrameNet involves a multi-step process highlighting frame development, LU identification, and annotation. First, lexicographers and linguists define the semantic frames based on their understanding of the conceptual structure of the domain. This process involves identifying the key events, situations, and entities in the domain and defining the semantic roles that participants and props play in each frame (Fillmore et al., 2003). Once the frames have been defined, the next step is to identify the LUs that evoke each frame. This process involves searching for words and phrases that are semantically related to the frame and that exhibit similar syntactic behavior. For each LU, lexicographers provide a definition, specify its part of speech, and map it to the relevant frame(s) it evokes (Ruppenhofer et al., 2016). Finally, annotators tag example sentences with the LUs and their corresponding FEs, creating a rich corpus of annotated data. The annotation process involves identifying the syntactic constituents that realize each FE in the context of the LU and labeling them with the appropriate FE tags (Baker, Fillmore, & Lowe, 1998).

This annotated data serves as a resource for studying the syntactic and semantic patterns associated with each frame and LU. To illustrate the end result of the FrameNet workflow, consider a LU such as "bake", from the "Cooking_creation" frame. In FrameNet, the entry for "bake" includes its definition ("to cook food by dry heat without direct exposure to a flame, typically in an oven"), its part of speech (verb), and the frame it evokes ("Cooking_creation").[1] The lexical entry also includes several manually annotated example sentences, such as "[$_{\text{Cook}}$She] BAKED$^{\text{Target}}$ [$_{\text{Produced_food}}$a cake] [$_{\text{Ingredients}}$from scratch]". In this sentence, the FE [Cook] is realized by the subject "She", the [Produced_food] is realized by the object "a cake", and the [Ingredients] is realized by the prepositional phrase "from scratch". By studying these annotated examples, it becomes feasible to identify the common semantic pattern "[Cook][Produced_food][Ingredients]" which is associated with the "bake" LU and the "Cooking_creation" frame more broadly.

Several studies have explored the application of FrameNet data to computational inquiries. Gildea and Jurafsky (2002), for example, pioneered the use of FrameNet for semantic role labeling, developing a statistical model to automatically identify the semantic roles associated with each predicate. Their work demonstrates the feasibility of using FrameNet for this task and lays the groundwork for subsequent research. In the domain of information extraction, Mohit and Narayanan (2003) used FrameNet to improve the accuracy of named entity recognition and relation extraction. By leveraging the semantic information encoded in FrameNet frames and FEs, they were able to develop a more robust and semantically-informed approach to these tasks. Other researchers have focused on extending and adapting FrameNet to new languages and domains. For example, the Spanish FrameNet project (Subirats & Sato, 2004) and the Japanese FrameNet project (Ohara et al., 2004) have developed FrameNet-like resources for these languages, enabling cross-linguistic studies of semantic patterns and supporting the development of multilingual NLP applications.

To date, FrameNet has not been extensively applied to the business domain, particularly in the analysis of corporate disclosures. However, related semantic annotation frameworks demonstrate that this computational approach offers greater flexibility and automation than traditional customized schemes. For example, Kloptchenko et al. (2004) used a custom-built semantic annotation scheme to analyze the content and style of quarterly earnings press releases, finding that companies strategically use language to convey a positive image and influence investor perceptions. Similarly, Kogan, Levin, Routledge, Sagi, and Smith (2009) developed a semantic representation scheme for analyzing the MD&A section of corporate 10-K filings. Their approach involved manually annotating a sample of MD&A texts with a set of predefined semantic tags, such as "revenue", "cost", and "risk". They then used these annotations to train a machine learning model to automatically identify these semantic categories in new MD&A texts, enabling a large-scale analysis of the semantic content of these disclosures.

While the aforementioned studies demonstrate the utility of semantic annotation using ad hoc, task-specific schemes, FrameNet offers a linguistically grounded, theory-driven alternative that enables richer, more structured representations of meaning. Unlike earlier schemes that rely on fixed, surface-level tags, FrameNet organizes meaning around semantic frames, each of which captures a type of situation along with its associated participant roles (Fillmore et al., 2003). For example, whereas Kogan et al.'s (2009) tags such as "revenue" or "cost" label semantic fields directly, FrameNet allows us to analyze these entities within broader "Commerce_buy" or "Earnings_and_losses" frames, which specify the roles of Buyer, Seller, Goods, and Money, or the relationships among Profits and Losses. This allows for cross-sentential and discourse-level pattern recognition that traditional tag-based systems may miss. Moreover, FrameNet supports polysemy and contextual

[1] See https://framenet.icsi.berkeley.edu/fnReports/data/lu/lu4896.xml?mode=lexentry.

disambiguation through its frame-to-frame relations and lexical unit mappings, enabling greater consistency in annotation across large and heterogeneous corpora (Ruppenhofer et al., 2016). This proves especially valuable in business texts where the same term (e.g., "risk") may instantiate different frames as to whether it refers to financial exposure, legal liability, or strategic uncertainty. In this study, we leverage FrameNet's cross-domain applicability and internal coherence to trace how corporate actors construct strategic narratives. Its ability to reveal the underlying lexico-grammatical patterns that constitute business discourse offers deeper insights than surface tag approaches. These insights are valuable not only for linguistic analysis but also for stakeholders such as investors, researchers, and educators who seek to decode the rhetorical and semantic strategies embedded in corporate messages.

3. Methods

3.1. Data

The primary data source for our study is the MD&A section from the 3M Corporation, which is a narrative component of annual business reports. MD&A is a section that provides a comprehensive overview of a company's financial performance, operational strategies, risk factors, and future outlook. It is often deemed the most frequently accessed section of annual reports (Humpherys, Moffitt, Burns, Burgoon, & Felix, 2011) and has been shown to have enhanced predictive capacity regarding business performance compared to external reports (Feldman, Govindaraj, Livnat, & Segal, 2010). These characteristics make MD&A an ideal choice for analyzing linguistic features, thematic categorizations, and semantic components within the context of corporate financial discourse (Tailab & Burak, 2021).

The decision to focus on the MD&A of 3M Corporation was informed by both methodological and practical considerations. As a multinational conglomerate operating across industrial, healthcare, and consumer sectors, 3M produces highly standardized, comprehensive, and rhetorically rich MD&A texts. Its disclosure practices closely follow U.S. SEC guidelines, making its reports genre-typical and structurally representative of U.S. corporate financial narratives (Basoglu & White, 2015). This makes 3M a strong prototypical case for fine-grained semantic analysis.

To ensure methodological rigor and minimize observer bias in case selection, we randomly sampled the 3M's MD&A from a larger curated corpus of approximately 2 million words drawn from 118 Fortune Global 500 U S. corporations (Qian, 2023). This sampling approach provides some protection against idiosyncratic bias and supports the validity of using 3M as a focal case. However, we acknowledge the limitations in generalizability. The findings of this case study are not intended to exhaustively characterize all MD&A texts across industries or firms. Rather, they serve as an initial demonstration of how FrameNet can be applied to corporate reporting texts to identify semantic structures and strategic framing. Future research will extend this approach to a broader set of MD&As across sectors to validate and refine the proposed semantic dimensions and patterns.

3.2. Research procedures

Within a broader methodological framework, our methodology consists of a comprehensive sequence of activities. It commences with data preprocessing and progresses through FrameNet integration, semantic dimension recognition, ultimately culminating in semantic pattern analysis. The process is facilitated by Python packages, specifically *nltk*, *pandas*, *string*, and *re*.[2] Here is a detailed explanation of these stages.

The first stage is to subject the MD&A dataset to a carefully planned series of preprocessing measures, thereby ensuring its suitability for in-depth analysis. This preliminary phase involves a multifaceted approach including the cleansing of textual data through the removal of redundant blanks, punctuation marks, and line breaks, thereby isolating the core textual content. Subsequently, a tokenization procedure is applied to segment the text into individual word strings. For example, the sentence "3M's total sales increased 2.5 percent in 2019 to $32.1 billion." would be tokenized into *3M, s, total, sales, increased, 2.5, percent, in, 2019, to, $32.1, billion* after cleanse. Furthermore, this phase includes the elimination of stop words (e.g., *s, in, to*), ensuring that only semantically meaningful words remain, and employing lemmatization techniques to standardize words to their root forms (e.g., *sales* to *sale*).

The subsequent step, termed "FrameNet integration", involved linking the MD&A text to relevant semantic frames from the FrameNet database, thereby generating a series of frame annotations. It is worth noting that FrameNet, originally developed based on the British National Corpus (BNC), is primarily tailored to general-domain English. To enhance its applicability to business discourse and to establish a standardized repertoire of frames along with their contextual interpretations (Ruppenhofer et al., 2016), we reinterpreted the original frame definitions[3] in light of their usage within the MD&A text. This reinterpretation followed a structured procedure: (1) The top 10 % most frequently occurring frames were identified to discern prototypical and core semantic patterns in the text (Fillmore & Atkins, 1992, pp. 75–102); (2) For each frame, a random sample of 30 sentences was selected. This sample size was chosen to balance representativeness with analytical manageability (Bhatia, 2004; McEnery & Hardie, 2011). Each sentence was then manually examined to determine

[2] The repository of Python packages is available at https://pypi.org/.
[3] See https://framenet.icsi.berkeley.edu/fnReports/data/frameIndex.xml.

Table 1
Semantic dimensions and their top frames.

Dimension	Frame (counts)	Definition
1. Financial metrics	• Earnings_and_losses (200)	• The company's financial performance, indicating its profits and losses over a specific period.
	• Money (122)	• The currency or capital used in financial transactions and operations.
	• Expensiveness (103)	• The cost structure of the company, indicating the level of expenses incurred in its operations.
	• Capital_stock (85)	• The total value of shares issued by a company.
	• Aggregate (80)	• The total or combined value of various financial elements.
2. Operations and business	• Businesses (166)	• The various aspects of a company's operations, including its products, services, and organizational structure.
	• Operating_a_system (107)	• The operational processes and systems used to conduct business activities efficiently.
	• Expertise (117)	• The specialized knowledge or skills possessed by the employees or leadership.
	• Collaboration (116)	• The cooperation and partnerships that the company engages in to achieve its business objectives.
	• Building_subparts (89)	• The components or sections of physical structures, such as buildings or facilities.
3. Time and duration	• Calendric_unit (389)	• Units of time, such as days, months, or years, used in the context of reporting and analysis.
	• Frequency (361)	• Represent how often certain events or activities occur within the company or its operations.
	• Relative_time (145)	• Involve the temporal relationship between events or activities, often used for comparing timelines.
	• Duration_description (93)	• Provide details about the length or time span of specific events or processes.
	• Time_vector (83)	• Represent a sequence of time-related data points, often used for analyzing trends.
4. Legal and ethical considerations	• Documents (181)	• Any document that has a legal status or conventional social significance.
	• Morality_evaluation (180)	• Include written or recorded materials of legal significance.
	• Statement (165)	• Provide information, make claims, or express legal positions.
	• Social_interaction_ evaluation (135)	• Relate to the assessment of social interactions within the company.
	• Judgment_communication (132)	• Suggest the exchange of opinions or assessments, often related to legal or ethical matters.

the frame's context-specific meaning within the business discourse domain; (3) frame definitions were then refined to reflect usage patterns and communicative functions pertinent to the MD&A discourse (e.g., risk disclosure, and performance evaluation). A detailed account of the redefinition procedures is provided in Appendix 2. Through this process, FrameNet was treated not as a static taxonomy, but as a flexible semantic framework: while its internal coherence was preserved, its application was recalibrated to accommodate the rhetorical and functional specificities of business reporting.

However, we recognize that using a general-domain resource like FrameNet comes with limitations. For example, some frames, especially those tied to complex or highly specialized financial concepts, such as "options" or "hedging", may not align neatly with the meanings these terms carry in corporate reporting. Although our reinterpretation protocol helped adjust definitions to fit the business context, it cannot fully address areas where FrameNet lacks domain coverage, such as derivatives, regulatory compliance, or taxation. This limitation is reflected in the prominence of more general financial frames (e.g., Earnings_and_losses, Money) in our analysis and in our decision to emphasize qualitative interpretation over precise automated tagging. Even so, our focus on frequently occurring frames and recurring semantic structures allows us to capture the dominant communicative patterns in MD&A discourse, making the findings both representative and relevant to the genre.

After adapting the frame definitions, we grouped frames into higher-order semantic dimensions. Semantic dimensions here refer to broad thematic categories that take in related frames and semantic elements. The identification of these dimensions is an iterative process that involves a close examination of frame definitions, their roles, and their relationships to one another (Fillmore & Baker, 2015). For example, frames such as "Earnings_and_losses", "Money", and "Expensiveness" could be grouped under the semantic dimension of "Financial Metrics", as they collectively contribute to the communication of financial performance and related aspects in the MD&A. Given the business context where the MD&A is situated, three criteria are set to determine the semantic dimensions: (1) the frequency of frames, i.e., the most frequent frames were prioritized; (2) the conceptual coherence of frames, i.e., the frames within each dimension share conceptual similarities and are related to a common theme or aspect of business reporting; and (3) communicative purpose, i.e., the dimensions are identified based on their relevance to the key aspects of MD&A's communicative functions, such as presenting financial performance, operational status, temporal references, and legal/ethical considerations (Qian, 2020). The grouping of frames was independently conducted twice by the first author, with a two-month interval between the sessions. A high percentage agreement rate of 0.91 was achieved, suggesting substantial consistency between the two rounds of coding (Lin & Evans, 2012). Any discrepancies were then resolved through discussion with the second author, an expert in Frame Semantics.

Once the semantic dimensions have been identified, the analytical procedure continues to the fourth stage, i.e., "semantic pattern analysis" by scrutinizing the recurrent patterns within each dimension. Semantic patterns, in this context, refer to the recurrent configurations of FEs and their associated lexical units within the identified frames, constructing business meanings. To identify patterns, we selected 30 sentences at random from each primary frame within each dimension. Each sentence was manually coded by the same author twice with a two-month interval, who labeled the relevant FEs based on the redefined frames and the LUs that associate with each frame. The inter-coder agreement across the two rounds was 0.83, indicating a satisfactory level of consistency. For example, in the sentence "3M continued to invest in research and development, with expenditures of $1.9 billion, or 5.9 percent of sales", the LU "invest" evokes the "Businesses" frame, with "3M" as the [Business] FE, "research and development" as the [Descriptor], "expenditures" and "sales" as the [Earnings], and "$1.9 billion" and "5.9 percent" as the [Change_amount]. By analyzing the recurrent configurations of these FEs and their associated lexical units across multiple sentences, we can identify semantic patterns that are meaningful within the context of business communication. For instance, a semantic pattern found in this example could be: [Business][Descriptor][Earnings][Change_amount]. This pattern captures the recurring structure of a company investing a specific amount in a particular area, which is a common theme in business discourse related to financial strategies and resource allocation. This focused approach ensures that the extracted patterns are not only relevant to the identified semantic dimensions but also aligned with the specific communicative goals of business discourse, such as reporting financial performance, discussing operational strategies, and highlighting key investments or initiatives.

4. Semantic dimensions

Based on the procedures discussed in the previous section, we identified four fundamental semantic dimensions in our corpus: Financial Metrics, Operation and Business, Time and Duration, and Legal and Ethical Considerations. To ensure the replicability and falsifiability of our classification, we tabulated semantic frames appeared in each dimension together with a definition assigned to each frame. Table 1 presents a summary of these four dimensions and the top five frames within each dimension.

We now turn to a discussion of the four semantic dimensions in Table 1. The **"Financial Metrics" dimension** includes frames such as "Earnings_and_losses", "Money", "Expensiveness", "Capital_stock", and "Aggregate", which are all related to the communication of financial information. The "Earnings_and_losses" frame, for example, defined as "the company's financial performance, indicating its profits and losses over a specific period", captures the essence of how successful a corporate business is towards gains and expenses. For example, in the sentence drawn from the dataset "For the fourth quarter of 2017, net income attributable to 3M was $523 million, or $0.85 per diluted share, compared to $1.155 billion", the LUs "net income" evokes the "Earnings_and_losses" frame, providing a concrete instance of how this frame is used to communicate financial metrics.

The **"Operations and Business" dimension** includes frames such as "Businesses", "Operating_a_system", "Expertise", "Collaboration", and "Building_subparts". This dimension sheds light on how an organization conducts its day-to-day operations, resource management, and interactions with external entities. While the "Financial Metrics" dimension revolves around financial data and performance indicators, the "Operations and Business" dimension looks into the operational underpinnings of a company's activities. It explores organizational structure, process management, expertise utilization, partnerships, and physical infrastructure maintenance. The "Businesses" frame, for example, is defined as "the various aspects of a company's operations, including its products, services, and organizational structure". An example of this frame in action can be found in the sentence "From a business segment perspective, 3M achieved both total sales growth and organic local-currency sales growth in all five business segments." Here, the LU "segment" is an instance that evokes the "Businesses" frame, highlighting the company's financial performance from the perspective of its individual business segments.

The **"Time and Duration" dimension** pertains to temporal aspects within a company's activities and operations. It contains frames such as "Calendric_unit", "Frequency", "Relative_time", "Duration_description", and "Time_vector". This dimension focuses on capturing and analyzing various temporal factors, supporting data-driven decision-making, trend identification, and performance evaluation over time. These aspects are crucial for effective management and strategic planning within an organization, constituting central elements of MD&As, which aim to provide a comprehensive narrative of the company's financial and operational performance across time. For example, in the sentence "For the full year 2017, net income attributable to 3M was $4.858 billion, or $7.93 per diluted share, compared to $5.050 billion, or $8.16 per diluted share, for the full year 2016, a decrease of 2.8 percent on a per diluted share basis", the LUs "full year 2017" and "2016" evoke the "Calendric_unit" frame, which is defined as "units of time, such as days, months, or years, used in the context of reporting and analysis". By using LUs evoking this frame, the company provides a clear temporal context for the reported share decline.

Finally, the **"Legal and Ethical Considerations" dimension** involves frames associated with the evaluation, communication, and documentation of ethical and legal aspects within an organization. Notable frames include "Documents", "Morality_evaluation", "Statement", "Social_interaction_evaluation", and "Judgment_communication". This dimension stands out in the MD&A as it reflects the company's commitment to legal compliance, ethical conduct, and responsible business practices. It addresses potential risks, concerns related to reputation, and the expectations of investors and stakeholders. Effectively addressing this dimension is instrumental in building trust and confidence in the company's management and operations, a critical element for its long-term success. For example, in the sentence "The Company does not have a required minimum cash pension contribution obligation for its U.S. plans in 2018", the word "obligation" evokes the "Statement" frame, which is defined as "providing information, make claims, or express legal positions". By using this frame, the company emphasizes a fact about 3M's pension funding requirements for the upcoming year.

In summary, these four semantic dimensions in the MD&A are interconnected and stand out for their contributions to the narrative of 3M's business performance: Financial Metrics quantify results, Operation and Business detail the activities that generate them, Time and Duration provide context, and Legal and Ethical Considerations ensure adherence to ethical standards and legal requirements. Collectively, these four dimensions provide a comprehensive picture of the company's past, present, and future, helping stakeholders make informed decisions and assess its health and sustainability. Besides, by first establishing the broader dimensions and their associated frames, we set the stage for a more detailed analysis of the semantic patterns that emerge within these dimensions. This progression from dimensions to patterns allows for a comprehensive understanding of the semantic structure and communicative strategies employed in the MD&A.

5. Semantic patterns in top frames across dimensions

5.1. "Earnings_and_losses" in financial metrics

The "Earnings_and_losses" frame exhibits the highest frequency in the Financial Metrics dimension. As previously defined in Table 1, this frame describes a company's financial performance by indicating its profits and losses over a specific period. This frame is crucial in the context of MD&A, as it provides investors and stakeholders with essential information about the company's financial health. As illustrated in Figure 1, noteworthy tokens within this frame include "income" (occurring 88 times), "net" (occurring 67 times), "earnings" (occurring 24 times), "result" (occurring 13 times), and "loss" (occurring 6 times). This observation confirms that MD&A furnishes comprehensive insights into the financial standing of the company, with a particular emphasis on income generation, net earnings, and the potential for losses (Feldman et al., 2010). Stakeholders taking notice of this frame can anticipate in-depth information pertaining to the company's financial performance, involving aspects such as profitability, financial results, and the implications of potential losses on its overall fiscal well-being (Bochkay & Levine, 2019).

To identify the key semantic pattern representing the "Earnings_and_losses" frame, we analyzed the patterns and highlighted the most frequent ones as follows.

(1) **Semantic pattern**: [Unit][Earnings][Change_amount][Unit][Change_direction][Change_amount]
Example sentence: For [$_{\text{Unit}}$the fourth quarter of 2017], net [$_{\text{Earnings}}$INCOME$^{\text{Target}}$] attributable to 3M was [$_{\text{Change\_amount}}$\$523 million, or \$0.85 per diluted share, compared to \$1.155 billion, or \$1.88 per diluted share], in [$_{\text{Unit}}$the fourth quarter of 2016], [$_{\text{Change_direction}}$a decrease] [$_{\text{Change_amount}}$of 54.8 percent on a per diluted share].

The [Earnings] and [Change_amount] elements in pattern (1) are core to communicating the "Earnings_and_losses" frame, conveying the company's bottom-line financial performance. The optional [Unit] element offers the time information into the change. While additional frames like "Calendric_unit" evoked by "the fourth quarter of 2017" and "the fourth quarter of 2016", and "Comparison" evoked by comparing financial metrics between two time periods are also present in this sentence, in the dimension category of "Financial Metrics", they serve to contextualize rather than center the analysis. Here, the "Earnings_and_losses" frame takes precedence as it directly addresses the company's business performance, aligning with the semantic dimension's focus on the communication of financial information.

There are other less frequent semantic patterns which provide a structured approach to conveying changes in financial performance, focusing on the cause-and-effect relationship between earnings metrics and the factors influencing them. Consider the following example.

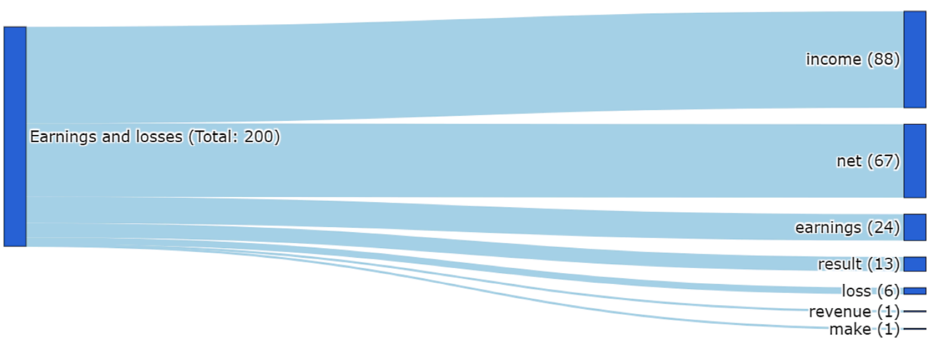

Figure 1. Lexical units in "Earnings_and_losses".

(2) **Semantic pattern**: [Earnings][Unit][Explanation]
Example sentence: Organic [$_{\text{Earnings}}$GROWTH$^{\text{Target}}$] in [$_{\text{Unit}}$2017] includes [$_{\text{Explanation}}$benefits from higher organic local-currency sales, raw material cost decreases from sourcing cost reduction projects].

In pattern (2), it establishes a link between earnings ("organic growth") of a particular time ("in 2017") and the underlying reasons (e.g., "benefits", "higher sales", "cost decreases"). This semantic structure allows for a more comprehensive understanding of the company's financial performance by identifying the key drivers behind changes in earnings (Loughran & McDonald, 2016). By providing this context, the pattern enables stakeholders to assess the sustainability and potential future impact of these factors on the company's earnings.

(3) **Semantic pattern**: [Geographic_area][Earnings][Change_direction][Change_amount][Explanation]
Example sentence: [$_{\text{Geographic_area}}$In Japan], total [$_{\text{Earnings}}$SALES$^{\text{Target}}$] [$_{\text{Change_direction}}$increased] [$_{\text{Change_amount}}$6 percent], [$_{\text{Explanation}}$as organic local-currency sales growth of 7 percent was partially offset by foreign currency translation impacts].

Pattern (3) offers a more granular view of earnings performance by focusing on specific geographic areas. It includes the direction and quantitative change in earnings and the reasons behind the changes. By disaggregating financial information by geographic area, this semantic pattern provides insights into the company's performance in different markets, allowing for a more comprehensive analysis of its earnings. Furthermore, the inclusion of foreign currency translation impacts highlights the role of external factors in influencing earnings, which is crucial for understanding the company's true underlying performance.

These alternative patterns effectively convey the "Earnings_and_losses" frame by highlighting fluctuations in financial performance metrics, albeit with variations in explicitly filled semantic roles and syntactic structures.

5.2. "Businesses" in operations and business

"Businesses" is the most frequently occurring frame in the dimension of "Operation and Business", which refers to the "various aspects of a company's operations, including its products, services, and organizational structure" (see Table 1 above). As indicated in Figure 2, the key tokens associated with this frame include "business" (occurring 99 times), and "per" (occurring 42 times). This observation suggests that the MD&A offers extensive insights into the company's operations, its core business activities, and the key strategies employed to drive growth and success. Stakeholders inquiring into this frame can anticipate a comprehensive exploration of the company's operational landscape, which contains its diverse businesses, the strategies employed to enhance efficiency and profitability, and the utilization of performance ratios for assessing business performance.

The most frequent semantic pattern identified for the "Businesses" frame takes the following form.

(4) **Semantic pattern**: [Business][Earnings][Change_direction][Change_amount][Explanation]
Example sentence: In [$_{\text{Business}}$INDUSTRIAL$^{\text{Target}}$], total [$_{\text{Earnings}}$sales] [$_{\text{Change_direction}}$increased] [$_{\text{Change_amount}}$6.9 percent, or 3.9 percent] on an organic local currency basis, [$_{\text{Explanation}}$with organic sales growth in abrasives, automotive and aerospace solutions, industrial adhesives and tapes, automotive aftermarket, and separation and purification].

The [Business] and [Earnings] elements are core to communicating the businesses frame, as they specify the performance of a particular business segment. The [Change_direction], [Change_amount], and [Explanation] elements provide

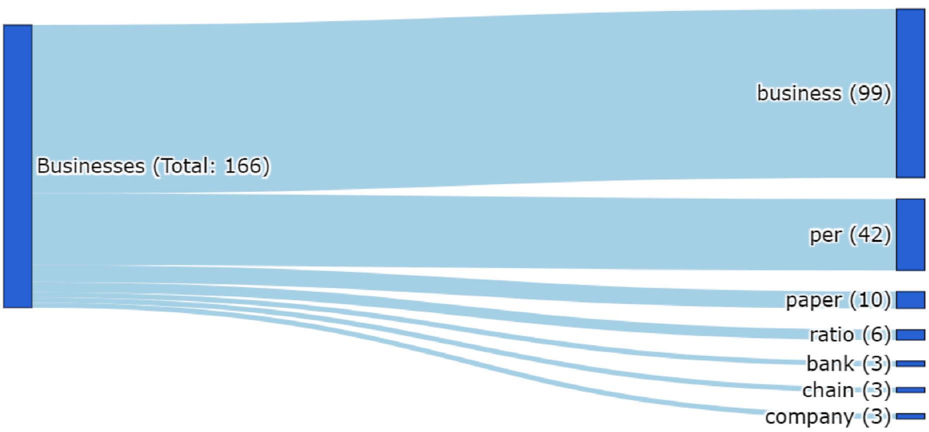

Figure 2. Lexical units in "Businesses".

additional context to understand the segment's performance. While this example sentence also contains elements related to the "Earnings_and_losses" frame, such as the change in total sales, the primary focus is on the performance of the industrial business segment, which is useful for understanding the company's overall financial health and identifying growth areas or potential challenges.There are other less noticeable semantic patterns which highlight the importance of business segments in analyzing a company's performance and strategic decisions. Consider the following example.

(5) **Semantic pattern**: [Earnings][Unit][Business]
Example sentence: The following discusses total year [$_{\text{Earnings}}$results] for [$_{\text{Unit}}$2017] compared to [$_{\text{Unit}}$2016 and 2016] compared to [$_{\text{Unit}}$2015], for each [$_{\text{Business}}$BUSINESS$^{\text{Target}}$] segment.

Pattern (5) allows for a more detailed analysis of a company's performance by disaggregating financial information by business segment. By providing this level of granularity, the pattern enables stakeholders to identify the strengths and weaknesses of each business segment, as well as potential opportunities for growth or areas of concern (Bens & Monahan, 2004). Moreover, this pattern highlights the importance of trend analysis in evaluating a company's performance over time, as it compares financial metrics across multiple time periods ("2017 compared to 2016" and "2016 to 2015") for each business segment.

(6) **Semantic pattern**: [Unit][Business][Product][Descriptor]
Example sentence: In [$_{\text{Unit}}$December 2017], [$_{\text{Business}}$3M$^{\text{Target}}$] announced it had reached an agreement to sell substantially all of [$_{\text{Product}}$its Communications Markets Division], [$_{\text{Descriptor}}$ which consists of optical fiber and copper passive connectivity solutions, structured cabling solutions, and telecommunications system integration services].

Pattern (6) provides a detailed account of 3M's strategic business move in December 2017, highlighting the sale of its Communications Markets Division. By identifying both the business entity involved and the specific division being sold, the sentence outlines a critical business decision that could significantly affect the company's market positioning. The inclusion of time reference adds a sense of urgency and relevance to the decision, suggesting that it is part of a broader strategic shift within the company. The four elements involved in this pattern together contribute to readers' understanding of 3M by offering a comprehensive glimpse into the company's strategic business decision and organizational focus.

5.3. "Calendric_unit" in time and duration

"Calendric_unit", which means "the units of time, such as days, months, or years" (Table 1 above), emerges as the most frequent frame within the "Time and Duration" dimension, signifying a substantial emphasis on time-related references and intervals. As evident from the data presented in Figure 3, the key tokens associated with this frame include "year" (occurring 192 times), "quarter" (occurring 83 times), "December" (occurring 36 times), "end" (occurring 15 times), "October" (occurring 11 times), and various other temporal references. This observation indicates that the MD&A thoroughly addresses temporal aspects, such as yearly and quarterly performance evaluations, with a specific focus on the month of "December" and other significant milestones like the "end" of reporting periods. Stakeholders exploring this frame can anticipate a detailed examination of the company's performance over different time horizons, including annual and quarterly reviews, allowing for a comprehensive understanding of its temporal dynamics.The most frequently observed semantic pattern for this frame takes the following form.

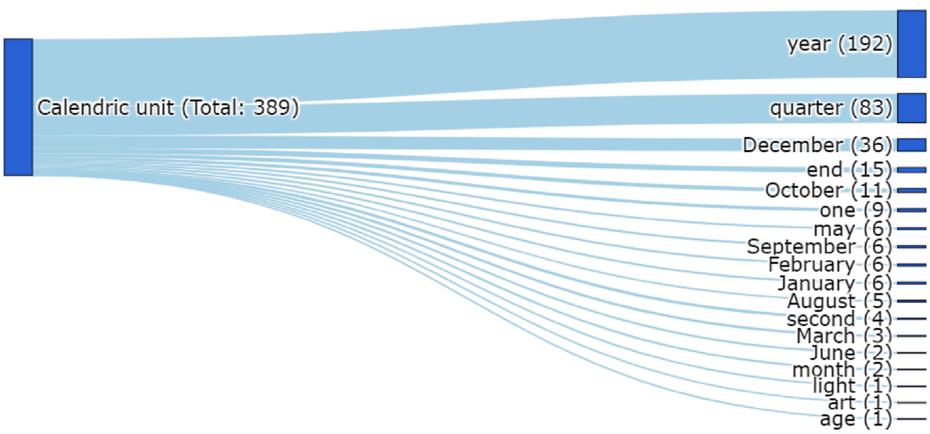

Figure 3. Lexical units in "Calendric_unit".

(7) **Semantic pattern**: [Earnings][Change_direction][Unit][Explanation]
Example sentence: [$_{\text{Earnings}}$Cost of sales as a percent of sales] [$_{\text{Change_direction}}$increased] during [$_{\text{Unit}}2017^{\text{Target}}$] due to [$_{\text{Explanation}}$incremental strategic investments in productivity, portfolio actions and footprint optimization, foreign currency effects and higher defined benefit pension expense].

In pattern (7), the inclusion of the [Unit] element is crucial for conveying the “Calendric_unit” frame, providing temporal context for the discussed financial metric. The [Earnings], [Change_direction], and [Explanation] elements contribute additional insights for interpreting the company’s performance within the specified timeframe. It is notable that the “Earnings_and_losses” frame is also present in this pattern, as evidenced by the discussion on changes in the cost of sales. The overlap between the two frames is conventional in business reporting discourse, as mapping corporate financial metrics to temporal dynamics and making a comparison is a common strategy in MD&As to evaluate a corporate business success (Qian & Sun, 2021). However, compared to “Earnings_and_losses”, the “Calendric_unit” attaches more importance to the dynamic and temporal aspects within a company’s activities rather than solely focusing on financial information. In addition to displaying the company’s performance achieved in a static date, it also discerns trends or seasonal fluctuations by comparing the performance across different reporting periods.

The rest of the semantic patterns share similarities in their contributions to the understanding of the “Calendric_unit” semantic frame by providing a structured approach to convey information about a company’s financial performance within specific units of time. They highlight the importance of temporal context in analyzing financial metrics and comparing them across different reporting periods. Consider the following example.

(8) **Semantic pattern**: [Earnings][Change_amount][Unit]
Example sentence: [$_{\text{Earnings}}$Operating income margins] were [$_{\text{Change_amount}}$22.4 percent compared to 21.4 percent] in [$_{\text{Unit}}$2015$^{\text{Target}}$].

In pattern (8), the sentence focuses on presenting an earnings metric (“operating income margins”) as a percentage for a specific period (”2015”). This structure allows for a clear and concise representation of the company’s performance within a given calendric unit, enable stakeholders to assess the company’s financial health and compare its performance against industry benchmarks or its own historical data.

(9) **Semantic pattern**: [Earnings][Unit][Change_amount][Change_direction][Change_amount] [Unit]
Example sentence: The [$_{\text{Earnings}}$effective tax rate] for the [$_{\text{Unit}}$fourth quarter 2017$^{\text{Target}}$] was [$_{\text{Change_amount}}$68.6 percent], an [$_{\text{Change_direction}}$increase] of [$_{\text{Change_amount}}$40.4 percentage points] versus [$_{\text{Unit}}$2016].

Pattern (9) builds upon pattern (8) by including the direction of change in the earnings metric compared to a previous period. This additional information enhances the understanding of the company’s performance by highlighting the magnitude and direction of change within the specified calendric units.

5.4. “Documents” in legal and ethical considerations

The “Documents” frame is the most frequent within the dimension of “Legal and Ethical Considerations”. It means “any document that has legal status or conventional social significance” (Table 1 above), highlighting the corporate commitment to conducting its operations with a robust ethical foundation and in compliance with legal standards. The lexical units evoking this frame highlight the importance of accurate reporting, adherence to rates and standards, and the aspiration to maintain high ethical and legal standards in various aspects of business operations. Key lexical units include, as indicated in Figure 4, “business” (72 occurrences), “item” (30 occurrences), “enact” (18 occurrences), and other terms hinting at

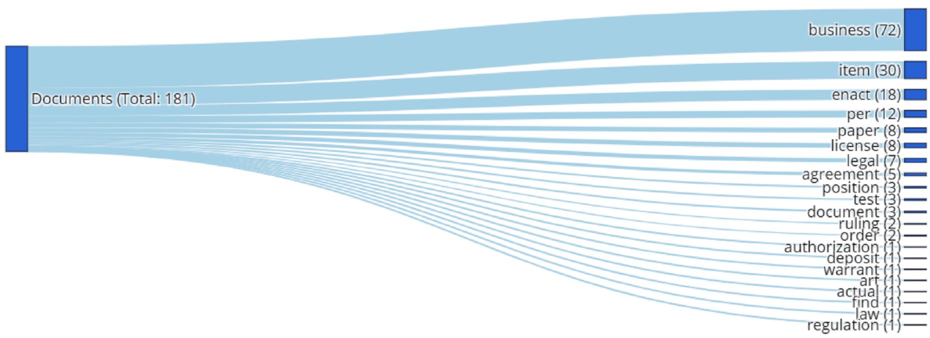

Figure 4. Lexical units evoking the “Documents” frame.

legislative factors, such as “license”, “legal”, “agreement”, and “ruling”. Within the frame of “Document”, the most frequent semantic pattern is the following.

(10) **Semantic pattern**: [Earnings][Unit][Obligation]
Example sentence: The [$_{\text{Earnings}}$tax rate] can vary from [$_{\text{Unit}}$quarter to quarter] due to discrete [$_{\text{obligation}}$ITEMS$^{\text{Target}}$], such as the settlement of income tax [$_{\text{Obligation}}$audits], changes in tax [$_{\text{Obligation}}$laws].

The [Obligation] element is crucial to conveying the “Documents” frame, as it provides the underlying tax laws that influence the changes in the financial metric. The [Earnings] and [Unit] elements offer additional information to help interpret the company’s performance and the impact of the specified reasons. In this pattern, the settlement of income tax audits and changes in tax laws, imply an evaluation of the company’s financial practices from a legal and ethical perspective.

Settling income tax audits suggests that the company’s tax practices have been scrutinized and potentially adjusted to comply with legal requirements. Changes in tax laws indicate that the company must adapt its financial strategies to align with new regulations, which may have moral and ethical implications. By highlighting these reasons, the sentence implies that the company’s financial performance is subject to legal and ethical considerations, and that changes in the tax rate are not solely driven by the company’s internal decisions but also by external factors that may have moral and ethical dimensions.

In addition to the main semantic pattern, there are other patterns that highlight the importance of the legal and regulatory context in shaping the content and presentation of financial information in the MD&A. Consider the following example.

(11) **Semantic pattern**: [Earnings][Unit][Obligation][Issuer]
Example sentence: This [$_{\text{Earnings}}$provisional amount] is subject to adjustment during the measurement [$_{\text{Unit}}$period of up to one year following the December 2017] enactment of the [$_{\text{Obligation}}$TCJA$^{\text{Target}}$], as provided by recent [$_{\text{Issuer}}$SEC guidance].”

Pattern (11) builds upon pattern (10) by including an issuer SEC associated with the provisional amount. This additional information highlights the regulatory changes and their potential impact on a company’s financial reporting. By specifying the issuer, this pattern emphasizes the need for companies to adapt their financial reporting practices in response to evolving legal and regulatory requirements.

(12) **Semantic pattern**: [Earnings][Obligation][Business]
Example sentence: The [$_{\text{Earnings}}$expense] is primarily related to the [$_{\text{Obligation}}$TCJA$^{\text{Target}}$]’s transition tax on previously unremitted earnings of non-The United States [$_{\text{Business}}$subsidiaries].

Pattern (12) differs from pattern (11) by shifting the focus from regulatory authority and timing ([Issuer] and [Unit]) to organizational structure and affected entities ([Business]). While pattern (11) centers on a provisional financial figure influenced by regulatory guidance within a defined temporal framework, pattern (12) highlights the organizational impact of regulatory obligations, specifically how a legal requirement (“the TCJA’s transition tax”) financially affects a company’s subsidiaries. The [Business] frame element is essential in this context because it identifies the specific business units or subsidiaries that are subject to the financial obligation. This situates the financial metric ([Earnings]) within a more concrete business context, helping readers understand which part of the organization is affected and why.

6. Discussion and conclusion

Our investigation has revealed four major semantic dimensions that permeate 3M’s MD&A: “Financial Metrics”, “Operation and Business”, “Time and Duration”, and “Legal and Ethical Considerations”. Within these dimensions, we unveiled a spectrum of high-frequency frames, each of which was redefined to reflect its use in the MD&A context. These redefined frames form the foundation for analyzing the semantic patterns that represent critical facets of corporate communication, ranging from financial disclosures to ethical positioning, thereby offering insight into the narrative construction of corporate performance.

MD&A discourse often evokes multiple frames within a single sentence, resulting in semantic overlap and multidimensionality. To manage this complexity, we adopted a top-down analysis strategy: rather than starting from individual lexical items, we began with semantically and communicatively coherent frame groupings (i.e., dimensions) and proceeded to extract the most frequent semantic patterns within each. This approach ensures consistency and interpretive coherence, anchoring pattern identification in the most salient semantic categories across the corpus.

Unlike bottom-up approaches such as semantic coding (Kloptchenko et al., 2004), bag-of-words analysis (Beattie, 2014), or rhetorical move analysis (Qian, 2020), our method is grounded in Frame Semantics, which systematically links meaning to syntactic structure. By redefining frame meanings based on real-life usage in the MD&A, we generated business-relevant semantic structures that reflect the functional architecture of corporate communication. This allows us to uncover replicable patterns such as [Unit][Business][Product][Descriptor], revealing how financial narratives are built at the semantic level.

Our findings reaffirm themes emphasized in prior research. Dimensions such as “Financial Metrics” and “Operations and Business” are closely aligned with the central purposes of MD&A texts, namely reporting, evaluating, and forecasting (Ren & Lu, 2021). Moreover, the semantic patterns identified align with rhetorical moves such as “Description of operating results”, “Presentation of liquidity and capital resources”, and “Description of contractual obligations” (Qian, 2020), highlighting the convergence between structural and semantic approaches in corporate discourse analysis.

These results also offer practical implications for Business English and ESBP instruction. Currently, genre-based business writing instruction often focuses on rhetorical moves or functional expressions, such as stating a company’s goals or presenting market risks. While valuable, this approach tends to emphasize what is said rather than how meaning is structured. Our semantic pattern analysis offers a complementary resource that focuses on the conceptual architecture behind the language. For example, consider a standard MD&A writing task where students are asked to describe quarterly financial results. Traditional instruction might teach formulaic expressions like “net profit increased” or “sales decreased slightly”. However, based on our findings, instruction can be reframed using semantic templates, such as [Earnings][Unit][Explanation], which students can populate not only with key metrics but also with explanatory content such as reasoning, contextual factors, or narrative framing, as illustrated in Example 2.

Furthermore, these semantic frames can be used to design genre-based diagnostic tasks. For example, learners might be given sentences from authentic MD&As and asked to identify the frame elements, explain their function, and generate new sentences using the same semantic configuration. This approach promotes both linguistic awareness and communicative fluency in financial reporting contexts. By offering reusable, interpretable semantic structures grounded in real-world corporate texts, this study contributes to the development of pedagogical models that are both authentic and transferable (Boas, 2025), qualities that are essential for effective ESBP teaching.

Our study also responds to broader discussions in genre pedagogy and ESP research, particularly the call for balancing genre acquisition with genre awareness (Johns, 2002, 2008). While much genre-based instruction in ESP still prioritizes rhetorical moves or fixed templates (Hyland, 2008; Swales, 1990), recent scholarship has emphasized the need for approaches that help learners develop critical flexibility to adapt genre knowledge across contexts (Flowerdew, 2023; Paltridge, 2014). By illuminating how corporate meaning is constructed through semantic dimensions and frame-based patterns, our study complements traditional genre models by adding a replicable, conceptual layer that supports both structural understanding and rhetorical adaptability. In this sense, it echoes Johns’ vision of genre awareness as a bridge between explicit instruction and context-sensitive language use, strengthening the pedagogical foundation for ESBP teaching.

A common concern, however, is that FrameNet’s general-domain origin may limit its ability to capture the semantic complexity of business language. We acknowledge this and took deliberate steps to address it. Rather than using FrameNet as a fixed taxonomy, we treated it as adaptable scaffolding. We retained the tagging architecture but redefined the top most frequent frames based on their function and usage in MD&A, following a structured set of procedures (Appendix 2). This strategy ensured that our semantic framework was aligned with the genre-specific purposes of business reporting, thereby enhancing both authenticity and domain relevance. Crucially, this adaptation does not undermine authenticity but preserves it. Our semantic categories are based entirely on a real-world MD&A text and reflect the communicative strategies that actual business professionals use. FrameNet simply provides the backbone for organizing those strategies into analyzable semantic structures.

Finally, we differentiate our method from prior work by combining theoretical grounding, corpus authenticity, and analytic rigor. While previous studies often used pre-defined tags, content themes, or rhetorical schemas, we offer a lexical-semantic framework (Frame Semantics) that allows generalization, a protocol validated through inter-rater reliability (percent agreement of 0.91 for frame grouping, and 0.83 for pattern annotation), and a balance between automated tagging and manual frame redefinition, documented in Appendix 2.

Future research may build on this by developing a dedicated domain-specific FrameNet for business reporting, akin to Bio FrameNet or Kicktionary (cf. Schmidt, 2009). Such an effort could systematize the reinterpretation process, expand the coverage of business-relevant frames, and support automatic tagging with greater genre specificity. Beyond computational enhancements, further pedagogical research could also test how frame-based tasks promote genre awareness in classroom settings, responding directly to Johns’ (2002, 2008) call for instructional models that connect explicit genre knowledge with adaptable, reflective language use. For instance, pilot studies could investigate how semantic pattern exercises help students internalize MD&A discourse conventions and transfer this awareness to new business contexts.

CRediT authorship contribution statement

Yubin Qian: Writing – original draft, Funding acquisition. **Hans C. Boas:** Writing – review & editing. **Qun Zheng:** Writing – review & editing.

Funding

The study is financially supported by the China Scholarship Council (grant number 202304910148) and the Beijing Social Science Foundation (grant number 21YYC013).

Data availability

Data will be made available on request.

Appendix. 1Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.esp.2025.08.002.

References

Baker, C. F., Fillmore, C. J., & Lowe, J. B. (1998). The Berkeley framenet project. *Paper Presented at the the 17th International Conference on Computational Linguistics, Montreal, Cananda*.

Basoglu, K. A., & White, C. E. (2015). Inline XBRL versus XBRL for SEC reporting. *Journal of Emerging Technologies in Accounting, 12*(1), 189–199.

Beattie, V. (2014). Accounting narratives and the narrative turn in accounting research: issues, theory, methodology, methods and a research framework. *The British Accounting Review, 46*(2), 111–134.

Bens, D. A., & Monahan, S. J. (2004). Disclosure quality and the excess value of diversification. *Journal of Accounting Research, 42*(4), 691–730.

Bhatia, V. K. (2004). *Worlds of Written Discourse: A Genre-based View*. London: Continuum.

Boas, H. C. (2020). A roadmap towards determining the universal status of semantic frames. In R. Enghels, B. Defrancq, & M. Jansegers (Eds.), *New approaches to contrastive linguistics: Empirical and methodological challenges* (pp. 21–52). Berlin/Boston: De: Gruyter.

Boas, H. C. (2025). Zur Konzeption eines ingenieurwissenschaftliches Online-Lernerwörterbuchs. G-FOL plus Gingko. In C. Fandrych, A. Portmann, L. Schirrmeister, & F. Wallner (Eds.), *Weichgeglüht und luftvergütet": Potenziale eines ingenieurwissenschaftlichen Korpus für Forschung und Vermittlung* (pp. 151–177).

Boas, H. C., Ruppenhofer, J., & Baker, C. (2024). FrameNet at 25. *International Journal of Lexicography, 37*(3), 263–284.

Bochkay, K., & Levine, C. B. (2019). Using MD&A to improve earnings forecasts. *Journal of Accounting, Auditing and Finance, 34*(3), 458–482.

Crawford-Camiciottoli, B. (2010). Earnings calls: Exploring an emerging financial reporting genre. *Discourse & Communication, 4*(4), 343–359.

Doumen, J., Beuls, K., & Van Eecke, P. (2023). Modelling language acquisition through syntactico-semantic pattern finding. In A. Vlachos, & I. Augenstein (Eds.), *Findings of the Association for Computational Linguistics: EACL 2023* (pp. 1347–1357). Association for Computational Linguistics. Dubrovnik.

Feldman, R., Govindaraj, S., Livnat, J., & Segal, B. (2010). Management's tone change, post earnings announcement drift and accruals. *Review of Accounting Studies, 15*(4), 915–953.

Fillmore, C. J. (1982). Frame semantics. In T. L. S.o. Korea (Ed.), *Linguistics in the Morning Calm: Selected Papers from SICOL-1981* (pp. 111–137). Seoul: Hanshin Pub. Co.

Fillmore, C. J., & Atkins, B. T. (1992). *Toward a Frame-based Lexicon: the Semantics of RISK and its Neighbors Frames, Fields, and Contrasts: New Essays in Semantic and Lexical Organization*. New York: Routledge.

Fillmore, C. J., & Baker, C. F. (2015). A frames approach to semantic analysis. In B. Heine, & H. Narrog (Eds.), *The Oxford Handbook of Linguistic Analysis* (pp. 313–340). Oxford: Oxford University Press.

Fillmore, C. J., Johnson, C. R., & Petruck, M. R. (2003). Background to FrameNet. *International Journal of Lexicography, 16*(3), 235–250.

Flowerdew, J. (2023). Commentary on Flowerdew, J., & Wan, A.(2010). The linguistic and the contextual in applied genre analysis: the case of the company audit report. *English for Specific Purposes, 29*(2), 78–93. English for Specific Purposes, 71, 57-59.

FrameNet. (2024). The "Cooking creation" frame. https://framenet.icsi.berkeley.edu/frameIndex.

Gildea, D., & Jurafsky, D. (2002). Automatic labeling of semantic roles. *Computational Linguistics, 28*(3), 245–288.

Goddard, C. (2012). Semantic primes, semantic molecules, semantic templates: key concepts in the NSM approach to lexical typology. *Linguistics, 50*(3), 711–743.

Humpherys, S. L., Moffitt, K. C., Burns, M. B., Burgoon, J. K., & Felix, W. F. (2011). Identification of fraudulent financial statements using linguistic credibility analysis. *Decision Support Systems, 50*(3), 585–594.

Hyland, K. (2008). Genre and academic writing in the disciplines. *Language Teaching, 41*(4), 543–562.

Johns, A. M. (2002). *Genre in the Classroom: Multiple Perspectives*. Mahwah: Lawrence Erlbaum Associates.

Johns, A. M. (2008). Genre awareness for the novice academic student: an ongoing quest. *Language Teaching, 41*(2), 237–252.

Kloptchenko, A., Eklund, T., Karlsson, J., Back, B., Vanharanta, H., & Visa, A. (2004). Combining data and text mining techniques for analysing financial reports. *Intelligent Systems in Accounting, Finance and Management: International Journal, 12*(1), 29–41.

Kogan, S., Levin, D., Routledge, B. R., Sagi, J. S., & Smith, N. A. (2009). *Predicting Risk from Financial Reports with Regression. Paper Presented at the the 2009 Annual Conference of the North American Chapter of the*. Boulder, Colorado: Association for Computational Linguistics.

Li, F. (2010). The determinants and information content of the forward-looking statements in corporate filings-a Naïve Bayesian machine learning approach. *Journal of Accounting Research, 48*(5), 1049–1102.

Lin, L., & Evans, S. (2012). Structural patterns in empirical research articles: a cross-disciplinary study. *English for Specific Purposes, 31*(3), 150–160.

Loughran, T., & McDonald, B. (2016). Textual analysis in accounting and finance: a survey. *Journal of Accounting Research, 54*(4), 1187–1230.

McEnery, T., & Hardie, A. (2011). *Corpus Linguistics: Method, Theory and Practice*. Cambridge: Cambridge University Press.

Mohit, B., & Narayanan, S. (2003). Semantic extraction with wide-coverage lexical resources. *Paper Presented at the the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology, Edmonton, Canada*.

Nelson, M. (2006). Semantic associations in business English: a corpus-based analysis. *English for Specific Purposes, 25*(2), 217–234.

Ohara, K. H., Fujii, S., Ohori, T., Suzuki, R., Saito, H., & Ishizaki, S. (2004). The Japanese FrameNet project: an introduction. *Paper Presented at the the 4th International Conference on Language Resources and Evaluation, Lisbon, Portugal*.

Paltridge, B. (2014). Genre and second-language academic writing. *Language Teaching, 47*(3), 303–318.

Qian, Y. (2020). A critical genre analysis of MD&A discourse in corporate annual reports. *Discourse & Communication, 14*(4), 424–437.

Qian, Y. (2023). Mining culture from professional discourse: a lexicon-based hybrid method. *Language Resources and Evaluation, 57*, 945–961.

Qian, Y., & Sun, Y. (2021). The correlation between annual reports' narratives and business performance: A retrospective analysis. *SAGE Open, 11*(3), 1–14.

Qian, Y., & Wu, N. (2025). Cultural interdiscursivity in managing corporate discourse: A corpus analysis. *Journal of Psycholinguistic Research, 54*(1), 1–15.

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

Ruppenhofer, J., Ellsworth, M., Schwarzer-Petruck, M., Johnson, C. R., & Scheffczyk, J. (2016). FrameNet II: Extended theory and practice. https://framenet2.icsi.berkeley.edu/docs/r1.7/book.pdf.

Schmidt, T. (2009). The kicktionary - A multilingual lexical resource of football language. In H. C. Boas (Ed.), *Multilingual Framenets in Computational Lexicography: Methods and Applications* (pp. 101–134). Berlin/Boston: de Gruyter.

Shawar, B. A., & Atwell, E. S. (2005). Using corpora in machine-learning chatbot systems. *International Journal of Corpus Linguistics, 10*(4), 489–516.

Shin, Y. K. (2020). Evaluative prosody and semantic preference: extending the analysis of recurrent multiword sequences. *English for Specific Purposes, 59*, 45–58.

Stubbs, M. (2001). *Words and Phrases: Corpus Studies of Lexical Semantics*. Massachusetts: Blackwell Publishers.

Subirats, C., & Sato, H. (2004). Spanish FrameNet and FrameSQL. *International Conference on Language Resources and Evaluation, Lisbon, Portuga*. Paper presented at the The 4th.
Swales, J. M. (1990). *Genre Analysis: English in Academic and Research Settings*. Cambridge: Cambridge University Press.
Tailab, M. M., & Burak, M. J. (2021). Examining the effect of linguistic style in an MD&A on stock market reaction. *International Journal of Business Communication, 58*(3), 430–458.
Trinder, R. (2013). Business students' beliefs about language learning in a university context. *English for Specific Purposes, 32*(1), 1–11.

Yubin Qian is an Associate Professor in the Department of Foreign Languages at the University of Chinese Academy of Sciences. His research interests include discourse analysis, intercultural communication, and corpus linguistics, with a particular focus on business communication. Dr. Qian has published in journals such as *Discourse & Communication, International Review of Applied Linguistics in Language Teaching,* and *Language Resources and Evaluation*. He also contributes to the academic community as an editor for several distinguished journals.

Hans C. Boas holds the Raymond Dickson, Alton C. Allen, and Dillon Anderson Centennial Professorship in the Department of Germanic Studies and the Department of Linguistics at the University of Texas at Austin. Professor Boas has made significant contributions to the linguistic field through influential publications on frame semantics, computational lexicography, and the relationship between language, cognition, and culture. He is also a leading figure in several high-impact research projects advancing the field.

Qun Zheng is a Professor in the Foreign Languages Department at the University of Chinese Academy of Sciences, where she has been dedicated to the field of linguistics for over two decades. Her research has been published in journals such as *System* and *Foreign Language World*, and she has made substantial contributions to a range of research projects. Professor Zheng's work bridges theoretical linguistics and its practical applications, with particular emphasis on language teaching and discourse analysis.