Systematic Literature Review and Bibliometric Analysis of Grammatical Violations in Financial Reports

Paulo Henrique Amaral Rody
https://orcid.org/0000-0001-9786-8049

José Elias Feres de Almeida
https://orcid.org/0000-0001-5220-0598

Abstract

Objective: This study presents a systematic literature review and bibliometric analysis of grammatical violations in financial statements and their effects on the capital market. Similar to previous studies, the conceptual framework concerns the violation of grammatical norms associated with the quality of financial statements.

Method: The method included two alternative and complementary paths: i) a systematic review in which the grammatical violation proxy is presented as a valuable tool to measure the quality of accounting information by analyzing the texts of explanatory notes of financial statements, and ii) a bibliometric analysis based on the journals listed in Web of Science, and Bibliometrix, an R-Tool for mapping analysis.

Results and Contributions: The in-depth literature analysis through a systematic review and a comprehensive analysis performed via bibliometrics allow this study to serve as a conceptual framework for future research. There are discussions on the potential directions of accounting research on the subject, which is strongly developed in the United States, but little explored in Brazil and other countries. Finally, the analysis suggests that the framework through which data become available to the market might restrict the application of this proxy.

Keywords: Grammatical Violations; Bibliometrix; Quality of Accounting Information; Comprehensibility; Bibliometrics.

1. Introduction

Financial statements provide a framework for companies to disclose accounting information, discussions, and explanations about the numbers reported, besides analyses that explain variations in results and balance sheets. However, the growing complexity of applying the principles of recognition, measurement, and disclosure of accounting standards and the textual language accompanying financial statements led to questions about the quality and ability of such disclosures to establish communication between a company and its investors and creditors.

Many researchers and users of financial statements argue that the increasing textual complexity of such disclosures makes it difficult for companies to bring any item to the attention of investors. Empirical evidence suggests that professional and non-professional investors cannot accurately interpret the most complex information in financial statements. Additionally, such complexity reduces efficiency when interpreting accounting information, increasing uncertainty (Gkikas, Tzafilkou, Theodoridis, Garmpis, Gkikas, 2022).

The corporate environment is affected by information asymmetry and agency problems between investors, entrepreneurs, and managers (Wang et al., 2021; Bourveau & Schoenfeld, 2017; Beyer, Cohen, Lys, Walther, 2010; Haniffa & Cooke, 2005). Information asymmetry might lead to distortions in the interpretation of financial statements. Hence, faced with incomplete or inaccurate information, external users may incur misinterpretations, undermining credibility due to the low quality of information (Scheld, Stolper, Walter, 2021; Elfeky, 2017). For Chung, Goh, Ng, and Yong (2017), the quality of financial information ensures that investors can rely upon such information. Therefore, mechanisms intended to reduce information asymmetry and improve the quality of accounting information become relevant.

While there is a substantial amount of research in the field of finance focusing on quantitative data, qualitative elements related to textual disclosure are seldom explored (Dyer, Lang, Lawrence, 2017). Among studies focusing on textual analysis in the finance field, Buehlmaier and Whited (2018) point out that measures of financial constraints using textual analysis of annual results efficiently capture the companies' volume of shares and debt. Along these lines, Hoberg and Maksimovic (2015) consider that restriction measures based on textual analysis affect companies with high cash flow levels. The textual analysis of financial statements enables one to predict a company's future cash flows (Frankel, Jennings, Lee, 2016). As highlighted by Borges and Rech (2019), evidence from Brazilian studies on textual analysis indicates that the textual quality of financial statements improves communication between companies and investors. Therefore, it is relevant to investigate mechanisms and new proxies capable of improving the quality of Brazilian financial information, benefiting society, companies, and investors.

Investors in the capital market need relevant, complete, and good-quality information. The disclosure of financial information that meets the needs of investors becomes essential for an efficient investment decision-making process. Therefore, the companies, which centralize such information, may adopt different disclosure strategies that may increase or decrease the transparency of information. Hence, conflicts of interest are expected to arise, considering that some companies may choose to disclose only part of the information or present low-quality information to investors, undermining the decision-making process. However, regardless of a company's strategy, a point seldom explored concerns the textual quality of what is disclosed. This new perspective may capture another nuance of financial statements' quality.

In this study's systematic review, Grammatical Violation (GV) is presented as a proxy capable of measuring the quality of financial information through a textual analysis of companies' standardized financial statements or specifically analyzing management reports, explanatory notes, or other structured reports containing texts released by companies. For example, a low level of GV would suggest high financial information quality, and accurately interpreting textual accounting information decreases the risk of uncertainty in decision-making, favoring investors' confidence in their investments (Gillette & Pundrich, 2020).

Research focusing on the textual analysis of financial statements contributes to understanding communication between companies and accounting information users (Borges & Rech, 2019). Li (2008) confirms this statement by showing that the linguistic characteristics of financial statements affect the companies' performance. The literature presents the effects and consequences of the textual portion of financial statements, and such evidence can help companies improve their disclosure process.

It is in this context that Grammatical Violation (GV) is inserted. It originates from linguistics, which focuses on the quality of texts, i.e., the fewer grammatical violations, the better the quality and credibility of information (Gillette & Pundrich, 2020; Campbell, Loumioti, Moerman, 2019). Gillette and Pundrich (2020) are considered precursors as they incorporated GV in financial accounting. They investigated whether GV is a timely indication of the quality and reliability of financial information, shedding light on the relationship between qualitative and quantitative aspects of corporate communication. Furthermore, this textual analysis metric differs from others, such as the Gunning fog and Flesch-Kincaid indices, widely used in the accounting literature.

GV is defined as the presence of errors related to the standard norm of a language, such as punctuation, syntax, and spelling mistakes, among others (Gillette & Pundrich, 2020; Hucks, 2015; Ghose & Ipeirotis, 2010). Few studies address the GV theme in the accounting and finance field in the Brazilian literature, configuring a gap. Hence, further research is needed on the quality of accounting information, considering the GV level of financial information as a proxy among Brazilian companies. This systematic review can serve as a conceptual basis and theoretical support for future studies. It will present directions in accounting research on this topic, which is still seldom explored in Brazil and several other countries. Most evidence comes from the United States, as researchers can access several reports and structured financial information sources, allowing automated processing (e.g., algorithm) and textual analysis tools.

Thus, this study presents a systematic literature review and a bibliometric analysis, building a theoretical framework to lay down the paths for future studies that will bring new evidence to the field of accounting sciences. Therefore, it has been established that low levels of grammatical violation suggest a greater quality of financial statements, decreasing information asymmetry and favoring better capital allocation.

In such an environment, investors need a mechanism capable of reducing information asymmetry between a company and its investors to make efficient decisions about the potential capital contribution in a company. Thus, the proxy developed in previous studies on the grammatical violation (GV) can be used to complement or add to previous studies addressing earnings quality, such as, for example, the quality of accruals, earnings management, and conservatism, among others. In addition, this GV proxy has practical effects on the market, considering that companies could process their reports periodically to reduce the risk of disclosing financial statements with grammatical errors.

This systematic literature review and bibliometric analysis on Grammatical Violation (GV) in financial statements signal to companies, to the market in general, and, mainly, to investors that the quality of accounting information, indirectly obtained through the level of GV, can be used strategically, as an indicator of reliable information and credibility. Therefore, this aspect allows investors to make more accurate decisions and constitutes another means for reducing information asymmetry. In addition, publicly traded companies can benefit from the technology used in GV algorithms to improve the quality of all sorts of accounting reports.

The following sections were organized into (2) grammatical violation and financial information quality; (3) grammatical violation in financial statements; (4) grammatical violation: systematic literature review; (5) grammatical violation: literature bibliometrics; and (6) final considerations.

2. Grammatical violations and the quality of financial information

2.1 Textual Analysis of Accounting Information

Changes in accounting standards impacted the process of preparing financial statements, considerably increasing the volume of reports companies must present to external users. Consequently, companies face challenges in disclosing information about their businesses in a clear and informative manner (Wang, Brabenec, Gao, Tang, 2021). Lehavy, Feng, and Merkley (2011) corroborate such a statement. The increase in the amount of disclosure required, coupled with the difficult task of disclosing business transactions to investors, has led to concerns about the quality of disclosure and the ability of external users to make sound decisions based on such information.

Accounting information plays two essential roles in market economies. First, it allows investors to assess the return potential of investment opportunities (*ex-ante*), and second, it enables investors to monitor the use of the capital invested (*ex-post*) (Beyer et al., 2010). These roles suggest that the quality of financial information affects the investors' decision-making. Hence, information must be well-written and comprehensible to facilitate users' understanding. In this context, the financial information in narrative texts is relevant to the stock market because investors use it in decision-making (Izco, Cabestre, Olalla, 2021). A similar analysis conducted by Silva and Fernandes (2009) indicates that Brazilian companies report fewer material facts about economic and financial performance in addition to more-difficulty-to-read accounting reports. However, comprehensibility is an essential quality of the information presented in financial statements so that users instantly and adequately assimilate it, whether from a corporate (Helms, Holscher, Nelde, Schneider, 2021) or an environmental perspective (Voigt, Machado, Meurer, 2020).

Companies operating in the capital market must disclose their information to investors in a more attractive and explanatory manner. In this sense, Gao, Lin, and Sias (2018) state that the level of reliability of financial information is a relevant aspect of interest to researchers, companies, and investors. Thus, research in the accounting and finance field focuses on verifying whether the qualitative attributes of corporate disclosure affect the decision-making of investors and information intermediaries.

Companies operating in a highly competitive global stock market should adopt strategies to maximize economic and financial performance to gain a competitive advantage over competitors (Alipour, 2019; Bonsall, Leone, Miller, Rennekamp, 2017). The international literature suggests that this competitiveness also applies to the Brazilian stock market, as companies disclose different amounts of financial information with varying quality levels. In such a context, information asymmetry decreases trust and increases the risk to which investors are exposed (Sethuraman, 2019; Li, 2017).

Next, Table 1 describes the topics that are trending in articles on textual analysis in the field of accounting in the international literature.

Table 1
Tendencies in Textual Analysis

Tendencies in Textual Analysis	Authorship
Relationship between behavioral aspects of language and employee performance.	Erickson, Holderness, and Thornock, 2022.
Influence of linguistic characteristics on information quality and content of accounting reports.	Brown, Crowley, and Elliott, 2020.
Analysis of style characteristics for disclosing accounting information and their relationship with investors' interpretation.	Grant, Hodge, and Sinha, 2018.
Relationship between textual characteristics and probability of earnings management.	Kim & Zhou, 2017.
Companies that have a higher level of earnings management in a given year and accounting reports with a lower degree of readability.	Lo, Ramos, and Rogo, 2017.
Translation of international accounting standards and practical problems of linguistic equivalence within the institutional context of IFRS.	Kettunen, 2017.
International accounting standards applied in countries with different languages can cause dialectical inconsistencies and impair the quality of accounting information.	Doupnik & Richter, 2003.

Source: Developed by the authors.

Table 1 presents the topics addressed in the articles, showing a need for studies to develop this line of research, i.e., the textual analysis of accounting information considering grammatical violations.

3. Grammatical violations in financial statements

Previous studies applying content and textual analysis in the accounting field started a new front of research interpreting texts, graphs, and words, among other elements extracted from financial statements or complementary reports, mainly because such evidence is still incipient and inconclusive (Souza & Borba, 2022). This portion of the literature processed analyses manually or with little automation until data processing advanced, and algorithms and systems allowed the automated analysis of the explanations provided in financial statements (Efretuei, Usoro, Koutra, 2021).

Recently, studies such as Gkikas et al. (2022) presented evidence that texts processed by algorithms capture, to some extent, helpful information for users, as they are associated with the price of assets traded on stock exchanges. In other words, it is possible to quantify, via analysis tools using machine learning, the quality of texts disclosed to external users and, thus, understand the impact of this factor on investors.

A text is a relevant tool for transmitting information between users; texts written with higher quality can significantly influence investors' decision-making, especially in a context with high information asymmetry (Si, Jiang, Fang, Usman, 2020). Being able to understand a text, despite its grammatical mistakes, directly depends on what is being read and the context in which information is transmitted (Williams, 1981). The transmission of information through writing texts creates an efficient communication network between the author and readers (Helms et al., 2021).

Nascimento and Henz (2021) consider that building knowledge of written language involves understanding and reflecting on textual production as an element constituted by norms. A set of skills is needed to produce a text, such as spelling accuracy, the proper use of ideas, words, phrases, grammar, and punctuation, besides assessment and revision (Brandenburg, 2015; Galbraith & Stillman, 2006). All this knowledge serves to generate texts that readers are capable of understanding entirely. However, errors may occur in this process and jeopardize the accurate interpretation of the meaning intended.

The main reasons justifying grammatical violations in Brazilian texts include factors such as the complexity of the spelling system, lack of knowledge of grammatical norms, stagnation of the writing learning process, lack of reading, and inattention. These factors lead to agreement, accent marks, and spelling errors, impairing the quality of a text (Nascimento & Henz, 2021).

Using correctly written texts that obey grammatical and spelling rules may indicate the user's personality characteristics and behavioral aspects, which become a relevant source of information for companies. Regardless of whether a lack of knowledge or negligence causes mistakes, they may have negative consequences for companies (Scott, Sinclair, Short, Bruce, 2014). In the context of accounting information quality, users may use the textual quality of financial statements as a parameter to assess the quality of financial information a company reports. For Hucks (2015), the characteristics of a text used in financial statements provide a significant indication of the quality of information. In this context, the GV level of financial information configures evidence of the quality of accounting information.

For Hucks (2015), grammatical violations indicate potential future default, longer financing terms, higher default rates, and lower yields for creditors. In such a context, GV is related to the debtor's credibility by assessing the textual quality a company presents to creditors. Therefore, companies with a high level of GV may incur increased costs due to the low quality of financial statements. Especially in the capital market context, grammatical violations are a relevant source of information for reducing information asymmetry, which becomes even more latent when the level of asymmetry is high and reliable information is limited (Shore, Tashchian, Forrester, 2021).

GV is defined as the presence of violations of a language's standard norm, such as errors in punctuation, syntax, spelling, and agreement, among others (Gillette & Pundrich, 2020). Grammatical and spelling errors affect how readers interpret texts (Kreiner, Schnakenberg, Green, Costello, Mcclin, 2002), hindering understanding and affecting the judgment of those analyzing a text (Shore et al., 2021). In this sense, one's educational level is a significant indicator of the probability of making such mistakes. Therefore, people with more years of schooling are less likely to make spelling mistakes than those with fewer years of schooling. Such a difference in educational level impacts the authors' behavioral differences (Hargittai, 2006).

For Gubala, Larson, and Melonçon (2020), the quality of redaction affects the business environment; companies tend to accept texts that comply with grammatical and spelling standards better. Assessing the quality of accounting information through the GV proxy is consistent because the GV level conveys information about the competence, awareness, and limitations of text information. This notion can be extended to the business environment as grammatical violations in financial statements reveal how reliable reports are and provide cues on the quality of financial statements.

GV levels may influence a company's sales volume. A higher level of GV is negatively associated with sales, compared to a lower level of GV, which tends to positively affect the volume of sales (Ghose & Ipeirotis, 2010). Therefore, the disclosure of financial statements must be timely and accurate to allow a more precise interpretation and, consequently, efficient decision-making for internal and external users (Silva & Fernandes, 2009).

Readers consider grammatically correct texts (without GV) more comprehensible (Liu, Wei, Gao, 2022). Hence, this evidence highlights the need to adopt correct redaction in texts, showing the importance of textual quality and its role in users' behavior (Queen & Boland, 2015). Furthermore, according to Boland and Queen (2016) and Everard and Galletta (2005), spelling mistakes give readers a negative impression of the competence of the individual who wrote the text and his/her attention to detail.

In line with the grammatical quality of a text, Gao et al. (2018) examined how writing style is associated with creditor behavior. They found that more aggressive creditors are more likely to grant financing, and at lower rates, to clients whose redaction is more comprehensible. In addition, customers with more comprehensible writing are less likely to default; hence, the quality of the textual writing facilitates obtaining financing.

Thus, verifying the level of GV in financial statements enables identifying evidence on the accuracy of the financial information in reports and serves as a new, more practical investigation method for measuring the quality of financial information (Gillette & Pundrich, 2020). Therefore, the assumption is that grammatical violations provide cues about financial statements' quality, as they correlate with the time, effort, and resources dedicated to preparing these reports.

4. Grammatical Violations: A Systematic Literature Review

A Systematic Literature Review and Bibliometric Analysis were used to identify the characteristics of Brazilian and international literature on Grammatical Violation (GV). The objective was for these tools to complement each other, broadening information on results. The methodological procedures of the two tools are described in the sections below (4.1 and 5.1, respectively).

The main difference between these tools is the step concerning variable parameterization. In the Systematic Literature Review, the researcher can specifically define the characteristics s/he wants to look for in the articles selected, whereas, in a Bibliometric Analysis, the parameters to identify the characteristics of articles are established in advance and cannot be manipulated.

4.1 Methodological Procedures

A Systematic Literature Review was performed to improve the quality of the theoretical foundation on the GV theme. It consists of reviewing the literature in a planned, structured, and controlled manner to identify the primary authors, relevant publications, and trends throughout time and highlight research gaps (Marioka, Iritani, Ometto, Carvalho, 2018).

Eligibility criteria were adopted to conduct the Systematic Literature Review to ensure a replicable and transparent procedure (Moher, Liberati, Tetslaff, Altman, 2009). Accordingly, the following criteria were established:

1. **Inclusion criteria**: (a) empirical studies addressing topics on the level of grammatical violation of financial information; (b) articles from the fields of finance, business, linguistics, and psychology; and (c) preferably, articles published in A1, A2, A3, A4, and B1 of Qualis Capes 2016.
2. **Exclusion criteria**: (a) articles in fields outside the scope of this study; (b) articles not addressing the topic of this study; and (c) studies not complying with the standard structure of peer-reviewed articles (Introduction, Theoretical Framework, Methodology, Results, and Conclusion).
3. **Identification and Selection**: The articles were identified and selected by reading titles, abstracts, and keywords.
4. **Time frame**: No time frame was established to select articles.
5. **Language**: articles written in Portuguese, English, or Spanish.
6. **Publication status**: peer-reviewed scientific articles.
7. **Search criteria**: Queries were made to electronic databases, and we also searched the studies cited by the articles selected.
8. **Search descriptors**: The words and/or expressions used to search for the articles were based on a consultation of the articles' abstracts and keywords related to this study's objective. Hence, synonyms were used to expand the scope and quality of the search for articles of interest for each subject.
9. **Databases**: The search was conducted in 4 databases: Periodical Portal of Capes, Web of Science, Scopus, and ranking of the 50 journals used by the Financial Times Research.

4.2 Literature Systematic Review

The results of the Systematic Literature Review on GV are described below, according to the database (Tables 2 to 5) and total general and total per language (Tables 6 and 7, respectively).

4.2.1 Periodical Portal of Capes

Descriptors: *violação gramatical* [grammatical violation]*; erro gramatical* [grammatical error]*; erro de ortografia* [misspelling]*; erro ortográfico* [orthographic mistake]*; erro de escrita*[writing error]*; erro escrito* [written error]*; qualidade do texto* [text quality]*; qualidade textual* [textual quality], and *mineração de texto* [text mining]. The filters "contain" and "in the title" were used.

Table 2
Papers on Grammatical Violation – Capes

Descriptor	Number of Articles
Violação gramatical	0
Qualidade textual	0
Erro gramatical	1
Erro de ortografia	1
Erro ortográfico	1
Erro de escrita	1
Erro escrito	0
Mineração de texto	1
Qualidade do texto	3
Total	8

Source: Developed by the authors.

4.2.2 Web of Science

Descriptors: grammatical violation; grammatical error; spelling errors, spelling mistakes; orthographic error; writing error; text quality; textual quality, and text mining. The filters "contain"; "no subject" were used.

Table 3
Papers on Grammatical Violation - Web of Science

Descriptor	Number of Articles
Grammatical violation	2
Grammatical error	7
Spelling errors	1
Spelling mistakes	1
Orthographic error	1
Writing error	0
Text quality	0
Text mining	0
Total	12

Source: Developed by the authors.

4.2.3 Scopus

Descriptors: grammatical violation; grammatical error; spelling errors, spelling mistakes; spelling error; writing error; text quality; textual quality, and text mining. The filters "contain" and "in the title" were used.

Table 4

Papers on Grammatical Violation – Scopus

Descriptor	Number of Articles
Grammatical violation	0
Grammatical error	5
Spelling errors	1
Spelling mistakes	0
Orthographic error	0
Writing error	2
Text quality	3
Text mining	2
Total	13

Source: Developed by the authors.

4.2.4 Financial Times Top 50 journals

Descriptors: grammatical violation; grammatical error; spelling errors, spelling mistakes; orthographic error; writing error; text quality; textual quality e text mining.

Open search – no filter parameterization.

Table 5

Articles on Grammatical Violation – Financial Times

Journal	Number of Articles
3. Accounting, Organizations and Society	5
4. Administrative Science Quarterly	1
12. Information Systems Research	2
13. Journal of Accounting and Economics	1
14. Journal of Accounting Research	3
43. Review of Accounting Studies	1
46. Review of Financial Studies	1
48. Strategic Entrepreneurship Journal	3
50. The Accounting Review	1
Total	18

Hyperlink: Link with the complete ranking

Source: Developed by the authors based on the Financial Times Top journals containing articles on the topic.

Table 5 describes the journal ranking. Of the 50 journals, only 9 contained articles addressing GV; the remaining 82% did not contain articles on GV. The complete ranking of journals is accessed through the link at the end of Table 5.

Table 6
Articles on Grammatical Violation – Financial Times

Descriptor	Number of Articles
Grammatical violation	6
Grammatical error	12
Spelling errors	0
Spelling mistakes	0
Orthographic error	0
Writing error	0
Text quality	0
Text mining	0
Total	18

Source: Developed by the authors.

4.2.5 Summary of the Systematic Literature Review

Table 7 presents the general total of articles and according to the database, and general total of articles and according to language.

Table 7
Articles on Grammatical Violation – Total

Database	Number of Articles
CAPES Periodicals	8
Web of Science	12
Scopus	13
Financial Times	18
Portuguese	12
Spanish	11
English	28
Total	51

Source: Developed by the authors.

The Systematic Literature Review revealed only three studies addressing GV in financial accounting in the international literature: Gillette and Pundrich (2020), Hucks (2015), and Ghose and Ipeirotis (2010). Furthermore, no articles addressing this topic were found in the Brazilian literature.

4.3 The Development of Textual Analysis Metrics

One of the most critical methodological limitations of research on the textual analysis of financial statements is that studies are restricted to very small samples. Such limitation usually occurs due to a difficulty in data processing, which, in most studies, is done manually. For example, Tennyson, Ingram, and Dugan (1990) examined the relationship between financial constraint and textual narrative with a sample of only 23 companies. Therefore, studies on textual analysis using automated data collection can contribute to developing textual analysis metrics as it enables the analysis of more extensive samples.

Among the articles found in the Systematic Literature Review, section 4.3.1 analyzes the three studies in the international literature addressing GV in the accounting field; no Brazilian studies were found. The following articles were analyzed: Ghose and Ipeirotis (2010), Hucks (2015), and Gillette and Pundrich (2020).

4.3.1 Textual Review and Economic Aspects – Ghose and Ipeirotis (2010)

The study focused on the impact of textual reviews on the economic results of product sales. It verified how different textual factors, such as perceived usefulness, interfere with social results. Various aspects of texts, such as subjectivity levels, readability measures, spelling measures, and grammatical errors, were explored and used to identify relevant aspects based on textual analyses.

Different levels of textual characteristics were also analyzed, such as the average utility of reviews and identity text measures. The econometric analysis showed that the degrees of subjectivity, informativeness, readability, and linguistic correction of the texts are relevant aspects that might affect a company's sales volume and also perceptions concerning the usefulness of the products sold.

The search's main results show that non-objective and subjective phrases are negatively associated with the volume of product sales compared to more objective and less subjective phrases; i.e., users deem objective and non-subjective phrases to be more informative and valuable.

Finally, this study highlights that it is possible to accurately predict the impact of textual analysis on the volume of product sales and perceived usefulness by analyzing the relative importance of three categories of textual resources: the resources of the textual sender, subjectivity text resources, and textual readability feature. This article is the first to integrate econometrics into the textual analysis of grammatical and spelling aspects to estimate the economic impact of textual aspects.

4.3.2 Textual Review and the Obtaining of Loans – Hucks (2015)

The study mentioned above verified whether the textual characteristics of loan applicants would reveal the issuer's behavioral aspects. The author analyzed whether the spelling and grammatical quality of those applying for a bank loan (the debtor) affects the decisions of companies granting loans (lenders). The authors consider that spelling errors probably do not represent a strategic choice, and this textual deficiency may impair communication between creditors and debtors.

The study revealed that spelling errors directly and negatively affect the lenders' decisions regarding capital allocation and that such errors also predict adverse outcomes for lenders, such as the incidence of high default rates and decreased rates of return. Consistent with this finding, misspellings in borrowers' loan application reports also predict the outcome of lenders' risk analysis; i.e., the more spelling errors in a text, the higher the interest rates on loans granted and the greater the difficulty in obtaining a loan.

The author also pointed out that the presence of spelling errors in loan applications is a sign of the borrowers' low level of financial capacity, and this aspect is used in the creditors' decision-making, decreasing the chances of borrowers obtaining a loan. Thus, borrowers may intentionally provide more readable texts with fewer grammatical errors to increase their chances of obtaining loans.

Thus, the study previously mentioned directly contributes to textual analysis in disseminating information in the credit market literature, establishing that the characteristics of textual language provide an essential cue about the quality of texts, the applicant's financial capacity, and the likelihood of obtaining loans.

4.3.3 Textual Review and Quality of Information – Gillette and Pundrich (2020)

The study above analyzed whether grammatical violations, including the number of typographical, syntax, punctuation, spelling, and other similar errors present in annual financial statements, provide a timely indication of the quality of financial reports. To the extent that grammatical violations capture the time, effort, and resources devoted to preparing financial records, the authors predicted that grammatical violations are correlated with the quality of accounting information.

This hypothesis was validated through the analysis of financial statements, which revealed that grammatical violations are negatively correlated with empirical metrics of the quality of a company's internal relationship with the information environment, the time companies have to file their financial statements, and how closely a company monitors the process of generating these accounting reports.

Consistent with this prediction, the authors above showed that grammatical violations positively correlate with the likelihood of reformulating accounting reports. In general, the study identified that grammatical violations might provide a timely indication of the quality of financial statements, in their qualitative and quantitative aspects, improving the disclosure process between companies and external users.

The authors claim that the users of financial statements depend on cues suggesting the quality of accounting reports to make investment, loan, or service contracting decisions. Consequently, studies focused on textual analysis can contribute to advancing this line of research on the quality of accounting information, especially considering the relevance of the quality of the texts in these reports.

The results suggest that grammatical violations can be used as an indicator of the quality of accounting reports within the context of disclosures of accounting information. Consequently, grammatical violations might provide a valuable measure of the reliability of accounting information in various forms of corporate communication, such as management forecasts and press releases.

The above article, which is the most recent in this line of research (2020), reports findings relevant to investors, researchers, regulators, and other users of financial statements and served as the basis for preparing this systematic literature review, intended to generate a theoretical framework for the construction of future empirical research on the influence of grammatical violations in Brazilian financial statements. The article also released several appendices on Python automation scripts for the data collection process.

REPeC – Revista de Educação e Pesquisa em Contabilidade, ISSN 1981-8610, Brasília, v.17, n. 2, art. 1, p. 117-139, Apr./Jun. 2023

129

5. Grammatical violations: literature's bibliometrics

5.1 Methodological Procedures

This section presents a bibliometric analysis intended to map and quantify the state of the art of the Grammatical Violation topic in the international and Brazilian literature. Section 4 presented several databases, and the current section presents a bibliometric analysis in the Web of Science database through the Bibliometrix Library, a package for the R software.

The Bibliometrix tool is open source and was developed by Aria and Cuccurullo (2017) to perform comprehensive analyzes of scientific mapping through bibliometric analyzes that are more consistent with the state of the art. Furthermore, because it is programmed by R software, Bibliometrix is flexible and can be rapidly updated and integrated with other R statistical packages.

The bibliometric analysis technique studies the characteristics of production, dissemination, and use of registered information. Hence, mechanisms using statistical and mathematical methods are intended to measure the scientific production of journals, authors, and other information from the literature (Macias-Chapula, 1998). Araújo (2006) shares this understanding, stating that bibliometrics is a resource that uses statistics and mathematics to map production rates, aspects of the literature, and dissemination of knowledge.

The following descriptors were the criteria used to select articles in the Web of Science database: *violação gramatical* (2 articles) and grammatical violation (344 articles); fields (all); knowledge area (all) and year (all).

5.2 Bibliometric Analysis of the Literature

Based on the sample of articles, the following bibliometric analysis was performed.
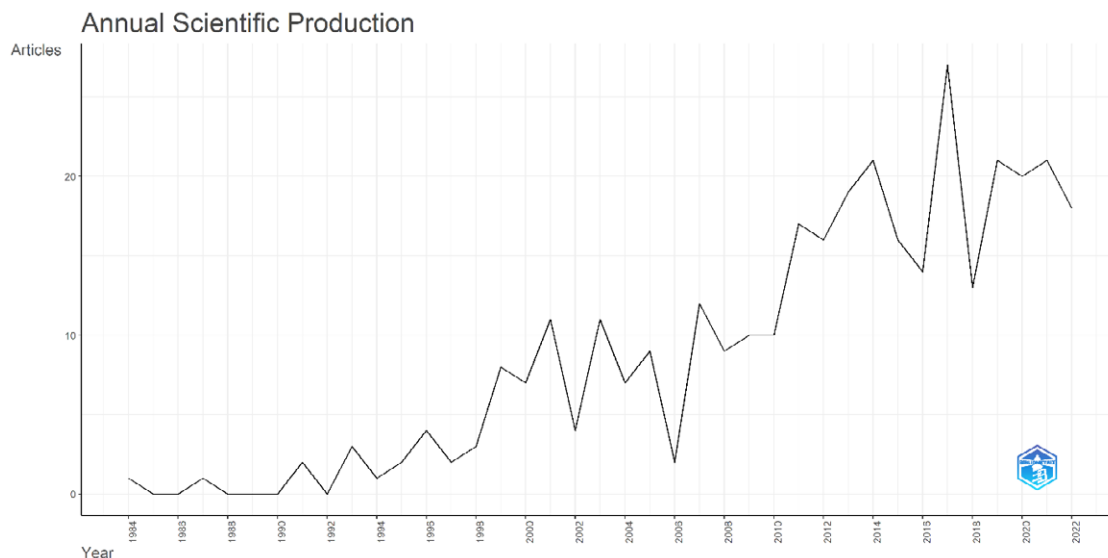

Figure 1. Total of Articles over the years

Source: *Bibliometrix*.

Figure 1 shows the total number of articles per year. The temporal range of the articles was from 1984 to 2022, with a total of 346 articles and 789 authors. The highest number of articles was published in 2017. The graph shows a trend towards increased productivity in the last six years.
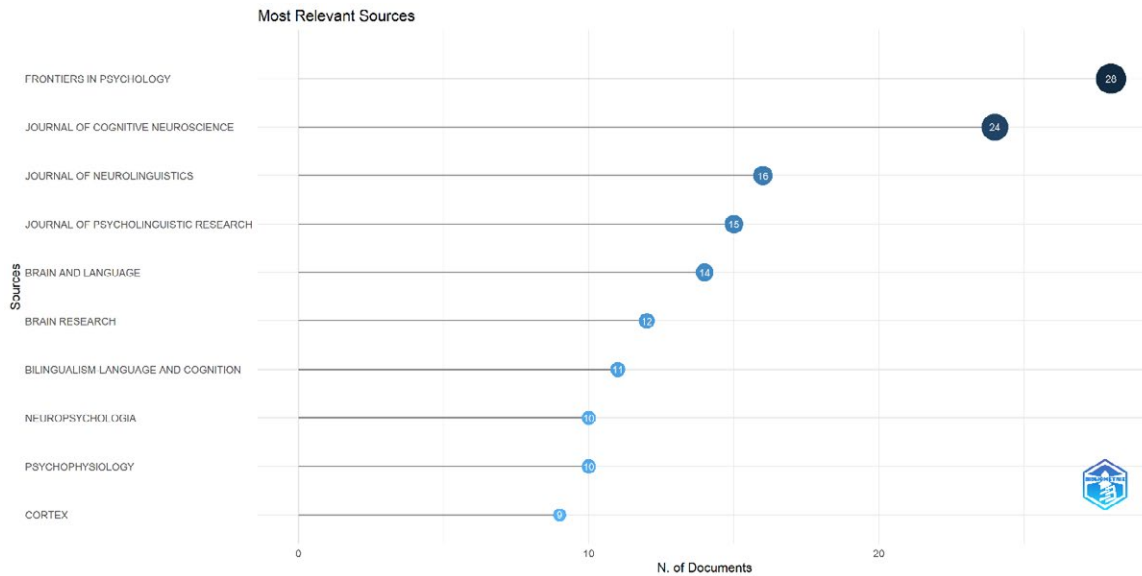
Figure 2. Total Journals

Source: Bibliometrix.

Figure 2 shows the total number of journals. Overall, 97 journals were analyzed; the three most productive journals were: Frontier In Psychology (28 articles), Journal of Cognitive Neuroscience (24 articles), and Journal of Neurolinguistics (16 articles). The other most productive journals were from the Linguistics, Psychology, and Computing fields. No results were found in financial accounting.

Figures 1 and 2 reinforce the notion that, with the increase in computers and servers' data processing capacity and with new tools for textual processing and machine learning, the field began to develop more intensely; there are many opportunities for research and application in the accounting field.
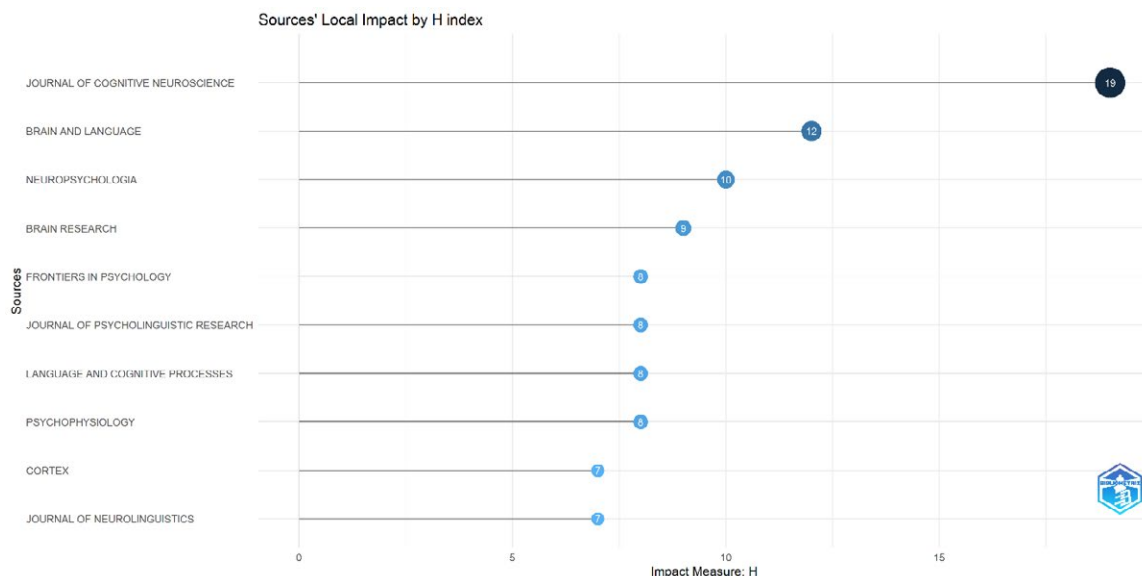
Figure 3. Journals' impact factor

Source: Bibliometrix.

Figure 3 presents the journals' impact factor (H-index); the publication with the highest impact was the Journal of Cognitive Neuroscience.
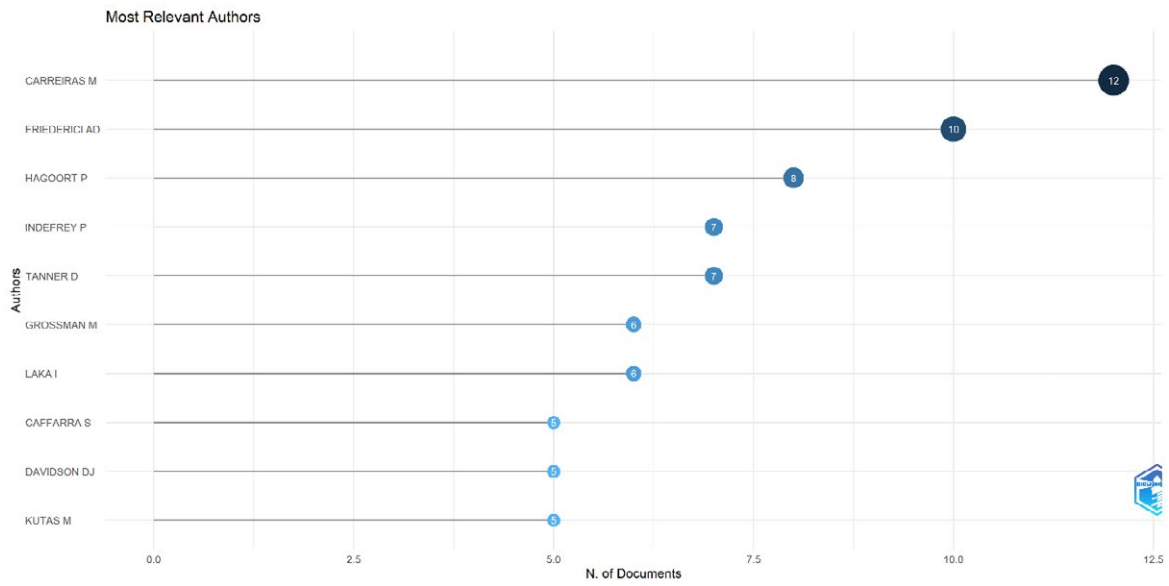
Figure 4. Most relevant authors

Source: Bibliometrix.

Figure 4 presents the most relevant authors. The 3 most productive authors were Carreiras (12 articles), Frederici (10 articles), and Hagdort (8 articles).
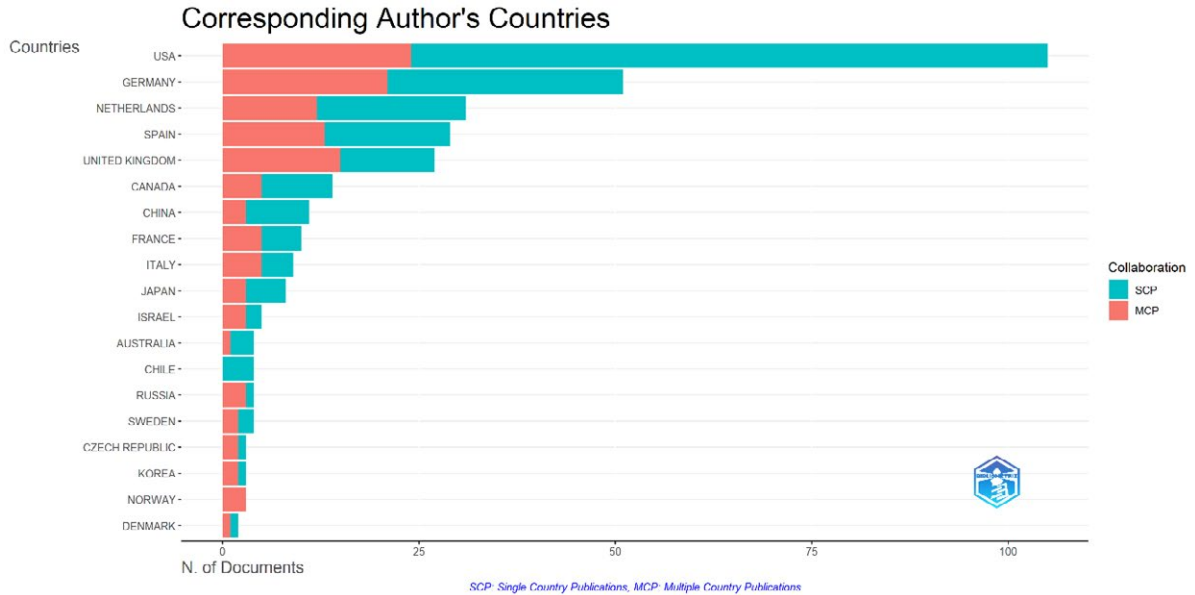

Figure 5. Chart of the most productive countries

Source: Bibliometrix.
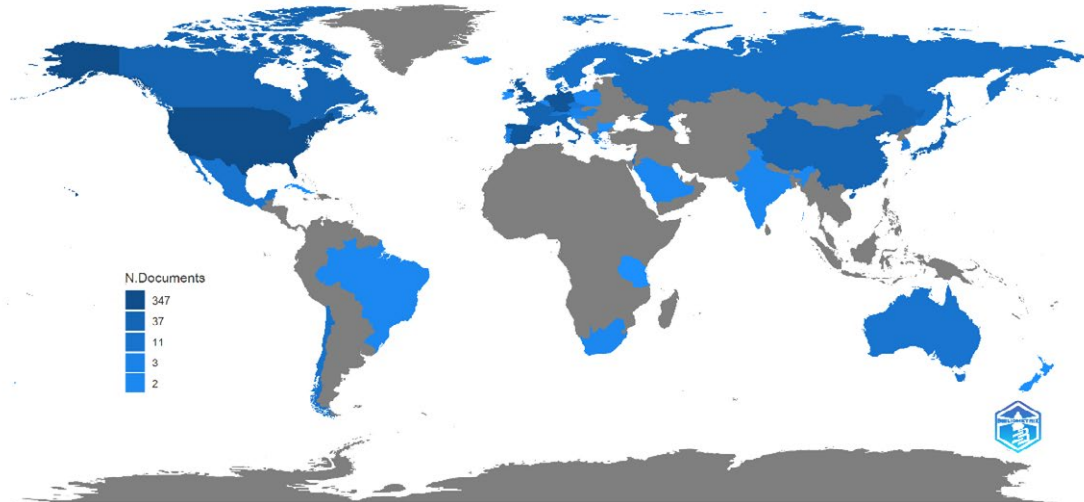
Country Scientific Production

Figure 6. Map of the most productive countries

Source: Bibliometrix.

Figures 5 and 6 present the most productive countries. The United States of America stands out, with 105 articles. Figure 6 depicts countries with high productivity (dark blue color), low productivity (light blue color), and zero productivity (gray color). As shown in the map, Brazil has low productivity, with only two articles, both in the linguistics field.
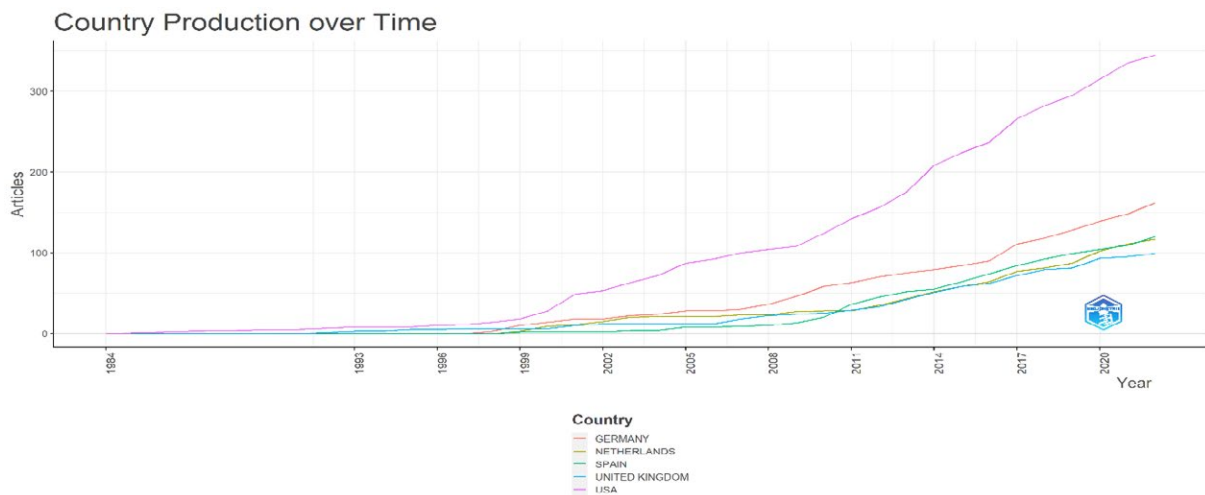

Figure 7. Production of articles per year in each country

Source: Bibliometrix.

Figure 7 presents the articles in each country per year. According to the analysis of the countries' production, from 1984 to 2022, the topic of Grammatical Violation in the international literature has been rising and growing over the last six years.
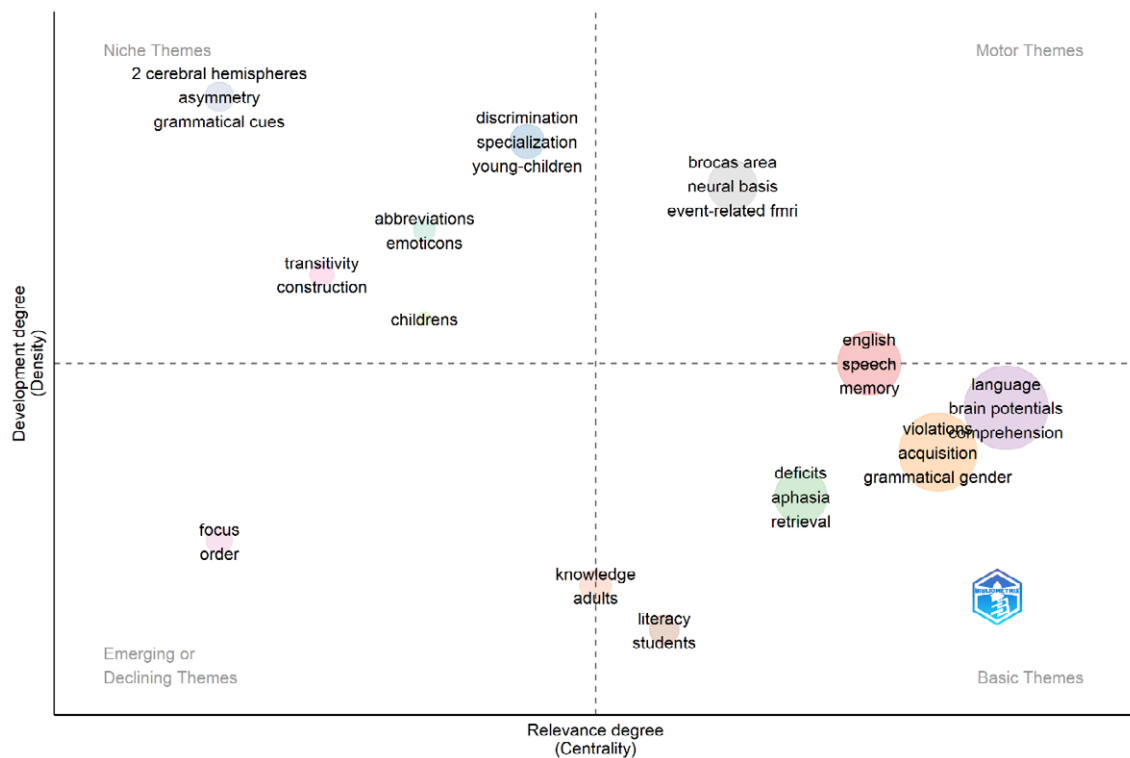
Figure 8. The topics most frequently addressed

Source: Bibliometrix.

Figure 8 presents the topics most frequently addressed in the articles: linguistic comprehension and grammatical vocabulary growth.
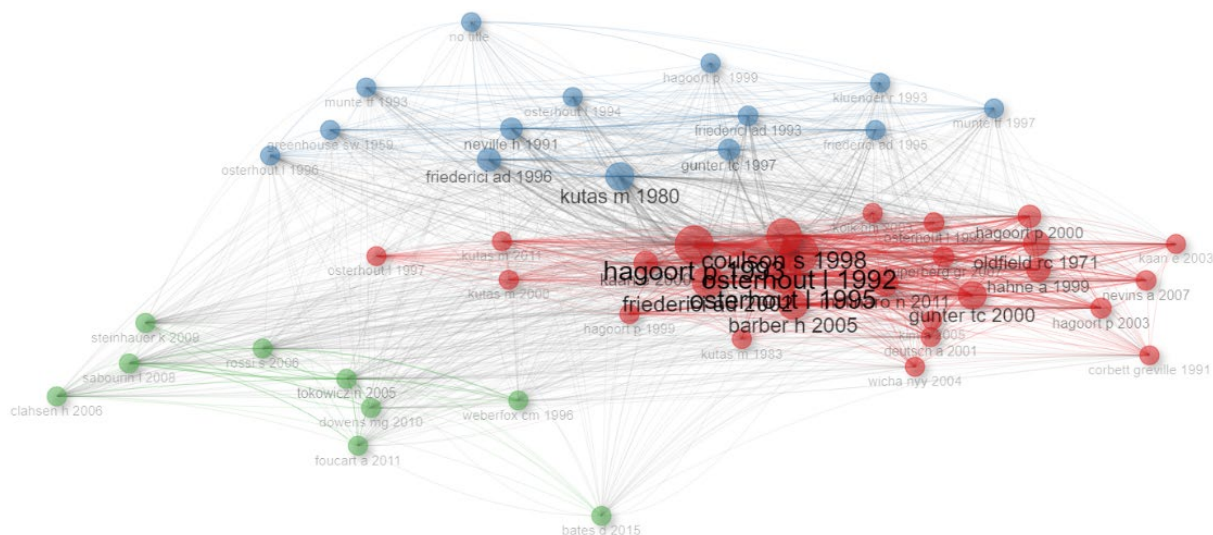

Figure 9. Co-citation Network

Source: Bibliometrix.

REPeC – Revista de Educação e Pesquisa em Contabilidade, ISSN 1981-8610, Brasília, v.17, n. 2, art. 1, p. 117-139, Apr./Jun. 2023

134

Figure 9 describes the co-citation network among the authors of articles on Grammatical Violation. The five most expressive authors in terms of co-citation were: Osterhout (1992), Hagoort (1993), Osterhout (1995), Coulson (1998), and Friederici (2002). Overall, the 50 most relevant authors in the field were highlighted and classified into three groups: strong citation network among authors (red color in the graph), average citation network among authors (blue color in the graph), and low citation network among authors (green color in the graph).

The results of the Systematic Literature Review and the Bibliometric Analysis reveal that researchers should pay attention and develop the topic of Grammatical Violation in Accounting. Future empirical evidence has the potential to fill such a gap in the literature on this topic. For example, studies can explore the quality of companies' accounting information through the analysis of Grammatical Violations, even if indirectly, to generate empirical data, which are still incipient, collaborating with the development of the topic and creating other opportunities for future research.

6. Final considerations

In this study, Grammatical Violation (GV) was presented as a proxy to indirectly measure the quality of financial information, decreasing information asymmetry and minimizing conflicts of interest between companies and investors. In this context, low levels of grammatical violation suggest a high degree of financial information quality, which might favor decreased information asymmetry.

Several studies in financial accounting focus on analyzing the role of quantitative variables, such as accounting, financial, and economic indexes, while qualitative aspects are seldom explored. Such a fact highlights a research gap in the international and Brazilian literature, though international studies are advancing. As few studies focus on the qualitative aspects of financial information, investigating the mechanisms capable of measuring the quality of financial information via the textual qualitative aspects of financial statements is a relevant contribution to financial accounting.

For this reason, this study is expected to be helpful when addressing a textual analysis variable originating from the linguistics field within financial accounting. Note that fewer grammatical violations indicate that a company has a higher level of financial information quality, increasing the credibility of information.

This study also signals companies, the market in general, and, mainly, investors that the quality of accounting information verified through the GV level of financial information can be used strategically to indicate reliable information and credibility. Analyses based on such a proxy can support investors' more accurate decisions and decrease information asymmetry. Thus, these aspects are relevant for publicly held companies, which can internally validate the quality of their financial statements produced by departments that establish relationships with investors and communicate with the market.

This Systematic Literature Review and Bibliometric Analysis can serve as a theoretical framework for creating a GV proxy for the Brazilian accounting field. In this context, economic and financial factors that affect the level of grammatical violation in financial statements disclosed by Brazilian companies can be analyzed. This study is the first to analyze this topic in Brazil; hence, future studies are suggested to address grammatical violations in financial information, providing empirical evidence of a subject still incipient in international literature, particularly in Brazilian literature.

References

Alipour, M., Ghanbari, M., Jamshidinavid, B., & Taherabadi, A. (2019). The relationship between environmental disclosure quality and earnings quality: a panel study of an emerging market. *Journal of Asia Business Studies, 13*(2), 326-347. https://doi.org/10.1108/JABS-03-2018-0084

Araújo, C. A. (2006). Bibliometria: evolução histórica e questões atuais. *Em questão*, *12*(1), 11-32. https://seer.ufrgs.br/index.php/EmQuestao/article/view/16

Aria, M., & Cuccurullo, C. (2017). bibliometrix: An R-tool for comprehensive science mapping analysis. Journal of informetrics, 11(4), 959-975. https://doi.org/10.1016/j.joi.2017.08.007

Beyer, A., Cohen, D. A., Lys, T. Z., & Walther, B. R. (2010). The financial reporting environment: Review of the recent literature. *Journal of accounting and economics*, *50*(2-3), 296-343. https://doi.org/10.1016/j.jacceco.2010.10.003

Boland, J. E., & Queen, R. (2016). If you're house is still available, send me an email: Personality influences reactions to written errors in email messages. *PloS one*, *11*(3), 1-17. https://doi.org/10.1371/journal.pone.0149885

Bonsall, S. B., Leone, A. J., Miller, B. P., & Rennekamp, K. (2017). A plain English measure of financial reporting readability. *Journal of Accounting and Economics*, *63*(2-3), 329-357. https://doi.org/10.1016/j.jacceco.2017.03.002

Borges, G., & Rech, I. J. (2019). Determinantes da legibilidade das notas explicativas de companhias brasileiras. *Revista de Gestão, Finanças e Contabilidade*, *9*(3), 31-51. https://doi.org/10.18028/rgfc.v9i3.7522

Bourveau, T., & Schoenfeld, J. (2017). Shareholder activism and voluntary disclosure. *Review of Accounting Studies*, *22*(3), 1307-1339. https://doi.org/10.1007/s11142-017-9408-0

Brandenburg, L. C. (2015). Testing the recognition and perception of errors in context. *Business and Professional Communication Quarterly*, *78*(1), 74-93. https://doi.org/10.1177/2329490614563570

Brown, N. C., Crowley, R. M., & Elliott, W. B. (2020). What are you saying? Using topic to detect financial misreporting. *Journal of Accounting Research*, *58*(1), 237-291. https://doi.org/10.1111/1475-679X.12294

Buehlmaier, M. M., & Whited, T. M. (2018). Are financial constraints priced? Evidence from textual analysis. *The Review of Financial Studies*, *31*(7), 2693-2728. https://www.jstor.org/stable/48615517

Campbell, D., Loumioti, M., & Wittenberg-Moerman, R. (2019). Making sense of soft information: Interpretation bias and loan quality. *Journal of Accounting and Economics*, *68*(2-3), 18-110. https://doi.org/10.1016/j.jacceco.2019.101240

Chung, S. G., Goh, B. W., Ng, J., & Yong, K. O. (2017). Voluntary fair value disclosures beyond SFAS 157's three-level estimates. *Review of accounting studies*, *22*(1), 430-468. https://doi.org/10.1007/s11142-016-9384-9

Doupnik, T. S., & Richter, M. (2003). Interpretation of uncertainty expressions: a cross-national study. *Accounting, Organizations and Society*, *28*(1), 15-35. https://doi.org/10.1016/S0361-3682(02)00010-7

Dyer, T., Lang, M., & Stice-Lawrence, L. (2017). The evolution of 10-K textual disclosure: Evidence from Latent Dirichlet Allocation. *Journal of Accounting and Economics*, *64*(2-3), 221-245. https://doi.org/10.1016/j.jacceco.2017.07.002

Elfeky, M. I. (2017). The extent of voluntary disclosure and its determinants in emerging markets: Evidence from Egypt. *The Journal of Finance and Data Science*, *3*(1-4), 45-59. https://doi.org/10.1016/j.jfds.2017.09.005

Erickson, D., Holderness Jr, D. K., Olsen, K. J., & Thornock, T. A. (2022). Feedback with feeling? How emotional language in feedback affects individual performance. *Accounting, Organizations and Society*, *99*, 101329. https://doi.org/10.1016/j.aos.2021.101329

Everard, A., & Galletta, D. F. (2005). How presentation flaws affect perceived site quality, trust, and intention to purchase from an online store. *Journal of Management Information systems*, *22*(3), 56-95. https://doi.org/10.2753/MIS0742-1222220303

Frankel, R., Jennings, J., & Lee, J. (2016). Using unstructured and qualitative disclosures to explain accruals. *Journal of Accounting and Economics*, *62*(2-3), 209-227. https://doi.org/10.1016/j.jacceco.2016.07.003

Galbraith, P., & Stillman, G. (2006). A framework for identifying student blockages during transitions in the modelling process. *ZDM*, *38*(2), 143-162. https://doi.org/10.1007/BF02655886

Gao, Q., Lin, M., & Sias, R. W. (2018). Words matter: The role of texts in online credit markets. *Journal of Financial and Quantitative Analysis, forthcoming*, 01-58. https://dx.doi.org/10.2139/ssrn.2446114

Ghose, A., & Ipeirotis, P. G. (2010). Estimating the helpfulness and economic impact of product reviews: Mining text and reviewer characteristics. *IEEE transactions on knowledge and data engineering*, *23*(10), 1498-1512. https://doi.org/10.1109/TKDE.2010.188

Gillette, J., & Pundrich, G. (2020). Grammatical Violations and Financial Reporting Quality. *Available at SSRN 3496434*. https://dx.doi.org/10.2139/ssrn.3496434

Gomes, A. C., de Macedo, H. C., & da Costa Marques, J. A. V. (2018). Efeitos da estrutura de capital sobre o gerenciamento de resultados contábeis no mercado brasileiro de capitais. *Sistemas & Gestão*, *13*(4), 469-481. https://doi.org/10.20985/1980-5160.2018.v13n4.1411

Gkikas, D. C., Tzafilkou, K., Theodoridis, P. K., Garmpis, A., & Gkikas, M. C. (2022). How do text characteristics impact user engagement in social media posts: Modeling content readability, length, and hashtags number in Facebook. *International Journal of Information Management Data Insights*, *2*(1), 1-9. https://doi.org/10.1016/j.jjimei.2022.100067

Grant, S. M., Hodge, F. D., & Sinha, R. K. (2018). How disclosure medium affects investor reactions to CEO bragging, modesty, and humblebragging. *Accounting, Organizations and Society*, *68*, 118-134. https://doi.org/10.1016/j.aos.2018.03.006

Gubala, C., Larson, K., & Melonçon, L. (2020). Do writing errors bother professionals? An analysis of the most bothersome errors and how the writer's ethos is affected. *Journal of Business and Technical Communication*, *34*(3), 250-286. https://doi.org/10.1177/1050651920910205

Haniffa, R. M., & Cooke, T. E. (2005). The impact of culture and governance on corporate social reporting. *Journal of Accounting and Public Policy*, *24*(5), 391-430. https://doi.org/10.1016/j.jaccpubpol.2005.06.001

Hargittai, E. (2006). Hurdles to information seeking: Spelling and typographical mistakes during users' online behavior. *Journal of the Association for Information Systems*, *7*(1), 1-16. https://doi.org/10.17705/1jais.00076

Helms, N., Hölscher, R., Nelde, M., & Schneider, J. (2021). The quality of risk reports: Integrating requirement levels of standard setters into text analysis. *Journal of Corporate Accounting & Finance*, *32*(3), 7-23. https://doi.org/10.1002/jcaf.22494

Hoberg, G., & Maksimovic, V. (2015). Redefining financial constraints: A text-based analysis. *The Review of Financial Studies*, *28*(5), 1312-1352. https://doi.org/10.1093/rfs/hhu089

Hucks, R. J. (2015). *Voluntary Involuntary Disclosure*. Unpublished Doctoral Thesis, Doctor of Philosophy (Business Administration), Universidade de Michigan. Ann Arbor, Estados Unidos da América.

Izco, M. Á., Cabestre, R. F. J., & Olalla, R. C. (2021). Readabilty in management reports: extension and good governance practices: La legibilidad en los informes de gestión: extensión y buenas prácticas de gobierno corporativo. *Revista de Contabilidad-Spanish Accounting Review*, *24*(1), 19-30. https://doi.org/10.6018/rcsar.363171

Kettunen, J. (2017). Interlingual translation of the International Financial Reporting Standards as institutional work. *Accounting, Organizations and Society*, *56*, 38-54. https://doi.org/10.1016/j.aos.2016.10.001

Kim, J., Kim, Y., & Zhou, J. (2017). Languages and earnings management. *Journal of Accounting and Economics*, *63*(2-3), 288-306. https://doi.org/10.1016/j.jacceco.2017.04.001

Kreiner, D. S., Schnakenberg, S. D., Green, A. G., Costello, M. J., & McClin, A. F. (2002). Effects of spelling errors on the perception of writers. *The Journal of general psychology*, *129*(1), 5-17. https://doi.org/10.1080/00221300209602029

Lehavy, R., Li, F., & Merkley, K. (2011). The effect of annual report readability on analyst following and the properties of their earnings forecasts. *The Accounting Review*, *86*(3), 1087-1115. https://doi.org/10.2308/accr.00000043

Li, F. (2008). Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and economics*, *45*(2-3), 221-247. https://doi.org/10.1016/j.jacceco.2008.02.003

Li, Y. (2017). Voluntary disclosure and investment in environmental technology. *Journal of Economic Behavior & Organization*, *133*, 331-341. https://doi.org/10.1016/j.jebo.2016.11.002

Liu, S., Wei, K., & Gao, B. (2022). Power of information transparency: How online reviews change the effect of agglomeration density on firm revenue. *Decision Support Systems, 153*, 1-13. https://doi.org/10.1016/j.dss.2021.113681

Lo, K., Ramos, F., & Rogo, R. (2017). Earnings management and annual report readability. *Journal of accounting and Economics*, *63*(1), 1-25. https://doi.org/10.1016/j.jacceco.2016.09.002

Macias-Chapula, C. A. (1998). O papel da informetria e da cienciometria e sua perspectiva nacional e internacional. *Ciência da informação*, *27*(2), 134-140. https://doi.org/10.18225/ci.inf.v27i2.794

Morioka, S. N., Iritani, D. R., Ometto, A. R., & Carvalho, M. M. D. (2018). Revisão sistemática da literatura sobre medição de desempenho de sustentabilidade corporativa: uma discussão sobre contribuições e lacunas. *Gestão & Produção*, *25*, 284-303. https://doi.org/10.1590/0104-530X2720-18

Moher, D., Liberati, A., Tetslaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA Statement. *Annals of Internal Medicine*, 151(4), 264-269. https://doi.org/10.1136/bmj.b2535

Nascimento, J. F., & Henz, R. R. (2021). A ortografia e os níveis de escrita: o erro em textos de sujeitos escolarizados. *Confluência: Revista do Instituto de Língua Portuguesa*, (61), 226-248. https://doi.org/10.18364/rc.2021n61.399

Queen, R., & Boland, J. E. (2015). I think your going to like me: Exploring the role of errors in email messages on assessments of potential housemates. *Linguistics Vanguard*, *1*(1), 283-293. https://doi.org/10.1515/lingvan-2015-0011

Scheld, D., Stolper, O., & Walter, A. (2021). Double Dutch finally fixed? A large-scale investigation into the readability of mandatory financial product information. *Journal of Consumer Policy*, *44*(2), 151-178. https://doi.org/10.1007/s10603-021-09486-0

Sethuraman, M. (2019). The effect of reputation shocks to rating agencies on corporate disclosures. *The Accounting Review, 94*(1), 299-326. https://doi.org/10.2308/accr-52114

Scott, G. G., Sinclair, J., Short, E., & Bruce, G. (2014). It's not what you say, it's how you say it: language use on Facebook impacts employability but not attractiveness. *Cyberpsychology, Behavior, and Social Networking*, *17*(8), 562-566. https://psycnet.apa.org/doi/10.1089/cyber.2013.0584

Si, H., Jiang, S., Fang, Y., & Usman, M. (2020). Can readability of loan description affect loan success rate and loan cost?: A textual analysis of P2P loan description. *Engineering Economics*, *31*(3), 302-313. http://dx.doi.org/10.5755/j01.ee.31.3.25760

Silva, C. A. T., & Fernandes, J. L. T. (2009). Legibilidade dos fatos relevantes no Brasil. *Revista de Administração Contemporânea-RAC Electronica*, *3*(1), 142-159. https://doi.org/10.58969/259471 25.1.1.2017.16

Shore, T., Tashchian, A., & Forrester, W. R. (2021). The influence of resume quality and ethnicity cues on employment decisions. *Journal of Business Economics and Management*, *22*(1), 61-76. https://doi.org/10.3846/jbem.2020.13670

Tennyson, B. M., Ingram, R. W., & Dugan, M. T. (1990). Assessing the information content of narrative disclosures in explaining bankruptcy. *Journal of Business Finance & Accounting*, *17*(3), 391-410. https://doi.org/10.1111/j.1468-5957.1990.tb01193.x

Voigt, F. H., Machado, J. V., & Meurer, A. M. (2020). Nível de evidenciação de custos e investimentos ambientais e legibilidade dos relatórios da administração de empresas do setor de papel e celulose. *Revista Mineira de Contabilidade*, *21*(3), 51-62. https://doi.org/10.51320/rmc.v21i3.1171

Wang, C., Brabenec, T., Gao, P., & Tang, Z. (2021). The Business Strategy, Competitive Advantage and Financial Strategy: A Perspective from Corporate Maturity Mismatched Investment. *Journal of Competitiveness*, *13*(1), 164-181. https://doi.org/10.7441/joc.2021.01.10

Williams, J. M. (1981). The phenomenology of error. *College composition and communication*, *32*(2), 152-168. https://doi.org/10.2307/356689