

Violação Gramatical e Grau de Legibilidade Das Demonstrações Contábeis No Brasil

PAULO HENRIQUE AMARAL RODY

*Doutor em Ciências Contábeis pela Universidade Federal do Espírito Santo (UFES)
Pós-doutorando em Ciências Contábeis pela FUCAPE Business School
paulohrody@gmail.com*

JOSÉ ELIAS FERES DE ALMEIDA

*Doutor em Controladoria e Contabilidade pela Universidade de São Paulo (FEA/USP)
Professor associado da Universidade Federal do Espírito Santo (UFES)
feresdealmeida@gmail.com*

Resumo

Este estudo analisa a associação entre o grau de Legibilidade e o nível de Violação Gramatical das demonstrações contábeis das companhias de capital aberto brasileiras. A Violação Gramatical é uma *proxy* capaz de medir indiretamente a qualidade das informações contábeis, dentro de um ambiente de assimetria de informações e de problemas de agência. As variáveis de Violação Gramatical e de Legibilidade foram mensuradas de forma automatizada via *Python* com informações extraídas de relatórios da administração e notas explicativas. A amostra compreende as empresas não financeiras listadas na [B]³, do período de 2010 a 2021. Nossos resultados indicam que: quanto maior o nível de Legibilidade, em média, menor o grau de Violação Gramatical das companhias; quanto maior o Tamanho da companhia, em média, menor o grau de Violação Gramatical das demonstrações contábeis; quanto maior o endividamento, em média, maior o grau de Violação Gramatical das demonstrações contábeis das companhias; e companhias auditadas por Big Four possuem menor grau de Violação Gramatical das demonstrações contábeis. Este artigo contribui para a literatura como um dos poucos estudos brasileiros da área a utilizar ferramentas de processamento e automatização de texto para interpretar padrões de linguagem dentro do ambiente contábil, além de sugerir que o menor nível de Violação Gramatical das demonstrações contábeis pode ser utilizado como indicador de informações contábeis confiáveis e de credibilidade.

Palavras-chave: Violação Gramatical; Legibilidade; Qualidade das Demonstrações Contábeis.

Realização



UFRJ
UNIVERSIDADE FEDERAL
DO RIO DE JANEIRO



**Universidade
Federal
Fluminense**



1. Introdução

O ambiente de informações corporativas é afetado pela assimetria de informação e pelos problemas de agência entre companhias e investidores. Nesse sentido diante de informações incompletas ou imprecisas, usuários externos, podem ter interpretações equivocadas, o que prejudicaria a credibilidade devido à baixa qualidade das informações (Barkemeyer, Comyns, Figge, Napolitano, 2014).

A Violação Gramatical pode ser entendida como a presença não intencional de erros gramaticais e ortográficos nos textos (Gillette & Pundrich, 2020). A Legibilidade corresponde a complexidade do texto e mensura a dificuldade de compreensão, representando a capacidade que o texto possui de permitir ao leitor interpretar o seu significado de forma adequada (Telles & Salotti, 2021).

Ainda não foram encontrados estudos na literatura analisando a relação entre esses dois atributos textuais (Legibilidade e Violação Gramatical) das demonstrações contábeis brasileiras e internacionais. De um lado, a Violação Gramatical pode trazer redução da capacidade do agente em compreender a mensagem que o usuário da informação contábil deseja transmitir, devido a presença de erros ortográficos e gramaticais no texto. Por outro lado, a Legibilidade possibilita o aumento da compreensão desses usuários mediante textos menos complexos, redigidos de forma mais objetiva e direta.

A literatura contábil recente (Hucks, 2015; Gillette & Pundrich, 2020) tem utilizado ferramentas de processamento de texto com o objetivo de mensurar uma perspectiva diferente de um atributo da utilidade (qualidade) das demonstrações contábeis para a tomada de decisão. Todavia, distorções textuais podem interferir no processo decisório de investidores, analistas, credores e dentre outros. Por isso, é relevante investigar o impacto que essas distorções textuais podem gerar nas demonstrações contábeis e também de que forma elas afetam outras variáveis contábeis.

Esta pesquisa tem como objetivo identificar a associação entre o grau de Legibilidade e o nível de Violação Gramatical das demonstrações contábeis, considerando os efeitos das Violações Gramaticais sobre a qualidade textual das informações das companhias de capital aberto listadas na Brasil Bolsa Balcão - [B3] entre 2010 e 2021. Assim, este estudo gerou evidências empíricas que podem contribuir com o preenchimento dessa lacuna na literatura mediante a criação da hipótese de Pesquisa (H1): quanto maior o nível de Legibilidade das notas explicativas e dos relatórios da administração, menor o grau de Violação Gramatical das companhias abertas.

Em geral, as evidências encontradas por este artigo sinalizam que os índices de Legibilidade estão associados de forma negativa com a Violação Gramatical, o que sugere que quanto maior o nível de Legibilidade, em média, menor o grau de Violação Gramatical das companhias. Em testes adicionais o Gerenciamento de Resultados afeta positivamente a Violação Gramatical, inferindo que maiores níveis de Gerenciamento de Resultados geram maior grau de Violação Gramatical das companhias. Essas evidências possibilitam a redução da assimetria informacional e, consequentemente, diminuição dos riscos de interpretações equivocadas dos usuários, o que proporciona aumento da confiança dos investidores, aprimorando a alocação de capital.

O estudo contribui com a literatura ao gerar informações relevantes sobre os reflexos

Realização



UFRJ
UNIVERSIDADE FEDERAL
DO RIO DE JANEIRO



**Universidade
Federal
Fluminense**



gramaticais dos relatórios contábeis sobre a qualidade da informação contábil. As evidências também podem contribuir com os investidores que podem utilizar a Violação Gramatical de forma complementar na análise da qualidade das demonstrações contábeis, bem como, reguladores no monitoramento das empresas.

2. Referencial Teórico

2.1. Violação Gramatical das Demonstrações Contábeis

Os elementos textuais podem ser considerados uma ferramenta relevante de transferência de informações entre os usuários, principalmente em um cenário com elevada assimetria informacional, pois quando são redigidos com qualidade podem gerar influência significativa na tomada de decisão dos investidores (Si, Jiang, Fang, Usman, 2020). Diante de um texto escrito com erros gramaticais, a compreensão do erro por parte do leitor está diretamente ligada ao que está sendo lido e ao contexto da informação que se transmite (Williams, 1981). A transmissão de informação por meio da escrita cria uma rede de comunicação eficiente entre o autor e o leitor (Morais, 2020; Helms, Holscher, Nelde, Schneider, 2021).

No Brasil, dentre os principais motivos para justificar a ocorrência de Violação Gramatical de um texto destacam-se a complexidade do sistema ortográfico, o desconhecimento da norma gramatical, a estagnação do processo de aprendizagem da escrita, a falta de leitura e a desatenção ao escrever. Esses fatores levam a erros de concordância, acentuação e ortografia, o que prejudica a qualidade do texto (Nascimento & Henz, 2021; Rody & Almeida, 2023). Portanto, redigir textos obedecendo as normas ortográficas e gramaticais é essencial para a qualidade compreensiva que se deseja transmitir (Everard & Galletta, 2005).

A Violação Gramatical é definida como a presença nos textos de violações das normas gramaticais, como erros de pontuação, ortografia, concordância e outros erros de escrita de um texto (Gillette & Pundrich, 2020). Erros de ortografia afetam a forma como os leitores interpretam os textos (Kreiner, Schnakenberg, Green, Costello, Mcclin, 2002). Ou seja, os erros gramaticais dificultam a compreensão do texto e afetam o julgamento de quem os analisa (Shore, Tashchian, Forrester, 2021). Nesse sentido, o nível de educação é um indicador significativo da probabilidade de cometer esses erros. Portanto, indivíduos com maior grau de escolaridade são menos propensos a cometer erros de ortografia quando comparados àqueles com menor grau de escolaridade. Essa relação impacta nas diferenças de comportamento dos emissores de um texto (Hargittai, 2006).

Para Gubala, Larson e Melonçon (2020), a qualidade da escrita tem efeito sobre o ambiente de negócios, uma vez que os textos que obedecem às normas gramaticais e ortográficas possuem maior aceitação entre as companhias. É consistente avaliar a qualidade da informação contábil por meio da *proxy* Violação Gramatical das demonstrações contábeis, pois o nível de Violação Gramatical transmite informações sobre a competência, a consciência e as limitações das informações contidas no texto. Isso é possível de ser estendido para o ambiente de negócios, a fim de verificar se as Violações Gramaticais das demonstrações contábeis revelam a confiabilidade desses relatórios, servindo como um sinal oportuno para a qualidade das demonstrações contábeis.

Textos com escrita correta, isto é, sem Violação Gramatical, são interpretados pelo leitor como mais compreensíveis (Liu, Wei, Gao, 2022). Essa evidência destaca a necessidade de se

Realização



UFRJ
UNIVERSIDADE FEDERAL
DO RIO DE JANEIRO



**Universidade
Federal
Fluminense**



Programa de Pós-Graduação em Administração - UFF

adotar uma escrita correta nos textos, dada a importância da qualidade textual e o seu papel no comportamento dos usuários (Queen & Boland, 2015). Os mesmos autores afirmam que os erros de ortografia também podem ser usados para formar impressões negativas sobre competência e atenção aos detalhes de quem emite o texto. Nessa mesma linha, para Gubala *et al.* (2020), a qualidade da escrita revela importantes aspectos sobre o autor.

2.2. Legibilidade das Demonstrações Contábeis

A Legibilidade corresponde a complexidade do texto, concentra-se no texto para medir a dificuldade de compreensão. A complexidade do texto é a capacidade que o texto possui de permitir ao leitor interpretar o seu significado de forma adequada (Telles & Salotti, 2021). As características das narrativas de Legibilidade das demonstrações contábeis podem gerar interpretações divergentes e causar distorções na interpretação das informações sobre a realidade econômica da companhia por parte dos investidores. Portanto, redigir parágrafos longos e com palavras de difícil compreensão pode desviar do real significado que se deseja transmitir nas demonstrações contábeis (Gkikas, Tzafilkou, Theodoridis, Garmpis, Gkikas, 2022).

As fórmulas de Legibilidade buscam mensurar a complexidade de um texto por meio do uso de palavras frequentes e estruturas sintáticas complexas. Nesse sentido, argumenta-se que a Legibilidade capta o grau de complexidade textual, isto é, textos com maior Legibilidade têm uma leitura menos complexa enquanto textos com menor Legibilidade possuem uma leitura mais complexa (Souza & Borba, 2022).

O processo de tomada de decisão é afetado pelas informações contábeis, que precisam ser compreensíveis para que possam ser úteis aos usuários (Telles & Salotti, 2021). Uma das funções da contabilidade é divulgar informações confiáveis, dessa forma aumentar a Legibilidade das demonstrações contábeis pode elevar a credibilidade junto aos investidores na evidenciação dessas informações (Holtz & Santos, 2020). De acordo com Monteiro, Rengel, Sousa e Borba (2021), há uma relação significativa e positiva entre o tamanho das companhias e a Legibilidade. Essa relação se justifica porque companhias, principalmente as de grande porte, possuem maior necessidade de emitir relatórios contábeis com maior grau de qualidade aos usuários externos, sobretudo os investidores.

Não obstante, o nível de Legibilidade de um texto pode ser controlado de forma intencional por uma companhia, em uma tentativa de relativizar as informações relevantes que poderiam ser apresentadas e melhor compreendidas caso fossem redigidas com maior nível de clareza. Nesse contexto, o elevado grau de Legibilidade poderia ser capaz de alterar as previsões dos usuários internos e externos ao possibilitar interpretações diferentes para o mesmo texto, dependendo da interpretação e capacidade de leitura de cada usuário (Voigt, Machado, Meurer, 2020).

3. Metodologia

3.1. Amostra Da Pesquisa

Este trabalho possui abordagem quantitativa. A técnica de análise de dados foi a Regressão Linear Múltipla com dados em painel de efeito fixo para firma e ano, com erro padrão clusterizado na firma, considerando que a variável dependente do modelo econométrico é contínua (Wooldridge, 2002). A amostra deste estudo é composta de todas as companhias não financeiras da [B]3, do período entre 2010 e 2021. A quantidade inicial de observações da

Realização

amostra de 6.528 (544 companhias x 12 anos) resultou em uma quantidade final de 965 a 1.018 observações (uma média de 85 companhias x 12 anos). Essa redução na quantidade final da amostra ocorreu devido a companhias com informações faltantes, inserção das *proxies* na base de dados e tratamento das variáveis na base de dados final.

3.2. Variáveis de Interesse Da Pesquisa

Para a mensuração das *proxies* Legibilidade (LE) e Violação Gramatical (VG), os arquivos com extensão em *.pdf* foram convertidos para o formato *.txt* por meio de *scripts* no *Python*, sendo as tabelas, sumários, cabeçalhos, imagens e figuras excluídas da análise dos arquivos para não provocar distorções na mensuração das referidas *proxies*.

3.2.1. Violação Gramatical

O banco de dados das Notas Explicativas (NE) e dos Relatórios da Administração (RA) das Demonstrações Contábeis das companhias, na primeira etapa de coleta de dados, foi criado de forma automatizada pela análise textual do *Software Python*. Assim, foram criados três códigos em *Python*, dois desses códigos baixaram as NE e as RA do sítio da [B]³ e o terceiro código acessou o *LanguageTool* para extrair as VG detectadas pelo programa para cada NE e RA de cada companhia.

Ressalta-se que o *LanguageTool* é uma biblioteca na qual as regras gramaticais foram validadas por profissionais da área de Linguística, ou seja, tais regras passaram pela avaliação e aprovação desses profissionais especialistas. Além disso, essa biblioteca também foi utilizada por outros estudos, como, por exemplo, a pesquisa de Gillette e Pundrich (2020), que foi utilizada como base para este estudo.

A mensuração de VG foi realizada conforme a Equação 1 e dividida em duas sub-Equações, uma com mensuração de VG com base nas NE e outra com base nos RA. Esta equação para mensuração de VG utilizou no denominador o número de palavras complexas, assim como foi adotado na pesquisa de Gillette e Pundrich (2020).

$$VG = \frac{\text{Número de VG praticadas por cada companhia}}{\text{Número total de palavras complexas de cada companhia}} \quad (\text{Equação 1})$$

Em que:

VGNE = Violações Gramaticais das Notas Explicativas; e

VGRA = Violações Gramaticais dos Relatórios da Administração.

Após a criação do *script* no *Python*, que mensurou o índice de Violação Gramatical, foi necessária uma etapa profunda e criteriosa de tratamento das informações internas mensuradas por esse índice, com o objetivo de garantir a consistência e qualidade do referido índice.

A apresentação qualitativa das Violações Gramaticais no *Python* possui a seguinte composição por meio da geração de um relatório em formato *.txt* para cada companhia da amostra: (1) número da Violação Gramatical (informação que foi utilizada para quantificar o total de Violações Gramaticais por companhia, por ano e o total geral da amostra); (2) descrição da regra gramatical que foi violada; (3) sugestão de correção da regra gramatical que foi violada; e (4) trechos das notas explicativas e dos relatórios da administração que foram constatadas as Violações Gramaticais.

A partir desse relatório foi realizado o tratamento dos falso-positivos, isto é, de determinadas violações gramaticais apontadas pelo *Python*, que não necessariamente configuram uma violação gramatical. Pois nem sempre são de fato um erro gramatical ou ortográfico e, muitas vezes, podem ser apenas divergências de estilo de escrita. Com isso, foram gerados dois relatórios para tratamento dos falso-positivos, um relatório com as regras que de

Realização

fato configuram violações gramaticais e ortográficas, que podem prejudicar a compreensão dos usuários e outro relatório com as regras que não configuram violações gramaticais e ortográficas.

Após a finalização desses relatórios e a conclusão do tratamento dos falso-positivos, foram utilizadas, para a mensuração do índice de Violação Gramatical, as regras gramaticais que realmente configuram violações gramaticais e que podem prejudicar a compreensão dos usuários. Esse relatório possui 66 regras gramaticais e ortográficas que foram consideradas para a mensuração do índice de Violação Gramatical das notas explicativas e dos relatórios da administração.

3.2.2. Legibilidade

A *proxy* Legibilidade (LE) foi coletada por meio das Notas Explicativas (NE) e dos Relatórios da Administração (RA) das Demonstrações Contábeis das companhias de forma automatizada mediante a criação de dois *scripts* que baixaram as NE e as RA de modo automático do sítio da [B]³.

A complexidade das demonstrações contábeis foi mensurada pela *proxy* LE, obtida por meio de seis índices utilizados pela literatura para estimar o nível de Legibilidade dos relatórios contábeis: (1) Índice de Facilidade de Leitura de Flesch (FLF), (2) Índice Gulpease (IG), (3) Nível de Graduação de Flesch-Kincaid (NG), (4) Índice de Nebulosidade de Gunning (ING), (5) Índice de Leiturabilidade Automatizado (ILA) e (6) Índice de Coleman-Liau (ICL). Assim, este estudo proporciona uma ampla análise das *proxies* de Legibilidade aplicadas na área contábil. Todos os índices de Legibilidade foram mensurados por meio de *script* no *Python*.

As equações dos índices são adaptadas do inglês para a língua portuguesa, todas consultadas no sistema de Análise de Legibilidade Textual (ALT). O sistema ALT disponibiliza de forma aberta a mensuração dos índices de Legibilidade.

Conforme consta no sistema ALT esses 6 índices de Legibilidade usam duas escalas: 0 a 100 (Teste de facilidade de leitura de Flesch e Índice Gulpease); e 0 a 20 (Nível de instrução de Flesch-Kincaid, Índice de Nebulosidade de Gunning, Índice de Leiturabilidade Automatizado, Índice de Coleman-Liau, entre outros). Os índices da escala de 0-100 possuem uma interpretação por ordem numérica crescente, em que zero indica um texto de difícil compreensão e cem sugere um texto de fácil compreensão. Os índices da escala 0-20 têm uma interpretação por ordem numérica decrescente, em que zero indica um texto de fácil compreensão e 20 sugere um texto de difícil compreensão.

3.2.2.1. Teste de Facilidade de Leitura de Flesch (Flesch Reading Ease)

O Teste de Facilidade de Leitura de Flesch (FLF) é utilizado para determinar a dificuldade de compreensão do texto mediante uma fórmula que mede a facilidade de leitura textual em uma escala que varia de 0 a 100, na qual zero indica um texto de difícil compreensão e cem sugere um texto de fácil compreensão (Flesch, 1948). A fórmula do índice FLF é especificada a seguir, conforme a Equação 2.

$$FLF = 206,835 - (84,6 \times CP) - (1,015 \times CF) \text{ (Equação 2)}$$

Em que:

FLF = é a medida de facilidade de leitura: medido pelo índice FLF;

CP = comprimento da palavra: medido pelo número de sílabas dividido pelo número de palavras; e

CF = comprimento da frase: medido pelo número de palavras dividido pelo número de

Realização

frases.

3.2.2.2. Índice Gulpease

O Índice de Gulpease (IG) é utilizado para determinar o nível de compreensão do texto, a partir de uma fórmula que gera um resultado numa escala que varia de 0 a 100, na qual zero indica um texto de difícil compreensão e cem sugere um texto de fácil compreensão (Gulpease, 1988). A fórmula do IG é especificada a seguir, conforme a Equação 3.

$$IG = \frac{89 + 300 \times (\text{Quantidade de sentenças}) - 10 \times (\text{Quantidade de letras})}{\text{Quantidade de palavras}}$$

(Equação 3)

3.2.2.3. Nível de Graduação de Flesch-Kincaid (Flesch-Kincaid Grade Level)

O Nível de Graduação de Flesch-Kincaid (NG) é utilizado para determinar o nível de compreensão do texto mediante uma fórmula que gera um resultado numa escala que varia de 0 a 20, na qual zero indica um texto de fácil compreensão e 20 sugere um texto de difícil compreensão (Kincaid *et al.*, 1975). A fórmula do NG é especificada a seguir, conforme a Equação 4.

$$NG = 0,36 \times \frac{(\text{Quantidade de palavras})}{\text{Quantidade de sentenças}} + 10,4 \times \frac{(\text{Quantidade de sílabas})}{\text{Quantidade de palavras}} - 18$$

(Equação 4)

3.2.2.4. Índice de Nebulosidade de Gunning (Gunning Fog Index)

O Índice de Nebulosidade de Gunning (ING) é utilizado para determinar o nível de compreensão do texto a partir de uma fórmula que gera um resultado numa escala que varia de 0 a 20, na qual zero indica um texto de fácil compreensão e 20 sugere um texto de difícil compreensão (Gunning, 1952). A fórmula do ING é especificada a seguir, conforme a Equação 5.

$$ING = 0,49 \times \frac{(\text{Quantidade de palavras})}{\text{Quantidade de sentenças}} + 19 \times \frac{(\text{Quantidade de palavras complexas})}{\text{Quantidade de palavras}}$$

(Equação 5)

3.2.2.5. Índice de Leiturabilidade Automatizado (Automated Readability Index - ARI)

O Índice de Leiturabilidade Automatizado (ILA) é utilizado para determinar o nível de compreensão do texto a partir de uma fórmula que gera um resultado numa escala que varia de 0 a 20, na qual zero indica um texto de fácil compreensão e 20 sugere um texto de difícil compreensão (Senter & Smith, 1967). A fórmula do ILA é especificada a seguir, conforme a Equação 6.

$$ILA = 4,6 \times \frac{(\text{Quantidade de letras})}{\text{Quantidade de palavras}} + 0,44 \times \frac{(\text{Quantidade de palavras})}{\text{Quantidade de sentenças}} - 20$$

(Equação 6)

3.2.2.6. Índice de Coleman-Liau (Coleman-Liau Index)

O Índice de Coleman-Liau (ICL) é utilizado para determinar o nível de compreensão do texto a partir de uma fórmula que gera um resultado numa escala que varia de 0 a 20, na qual zero indica um texto de fácil compreensão e 20 sugere um texto de difícil compreensão (Coleman, 1975). A fórmula do ICL é especificada a seguir, conforme a Equação 7.

$$ICL = 5,4 \times \frac{(\text{Quantidade de letras})}{\text{Quantidade de palavras}} - 21 \times \frac{(\text{Quantidade de sentenças})}{\text{Quantidade de palavras}} - 14$$

(Equação 7)

3.3. Modelo Econométrico da Pesquisa

Realização



UFRJ
UNIVERSIDADE FEDERAL
DO RIO DE JANEIRO



**Universidade
Federal
Fluminense**



Para verificar o efeito da relação entre o nível de Violação Gramatical e o grau de Legibilidade das informações contábeis das companhias de capital aberto no Brasil, estimou-se o modelo econométrico de Regressão Linear Múltipla, conforme descrito na Equação 8.

$$VG_{it} = \beta_0 + \beta_1 LE_{it} + \beta_2 GR_{it} + \beta_3 GR \times LE_{it} + \beta_4 BIG_{it} + \beta_5 ROA_{it} + \beta_6 TAM_{it} + \beta_7 ENDIV_{it} + \beta_8 MTB_{it} + \beta_9 RET_{it} + \beta_{10} CF_{it} + \beta_{11} TAU_{it} + \beta_{12} + \varepsilon_{it}$$

(Equação 8)

Esse modelo econométrico foi estimado em outros 12 submodelos econométricos, considerando as seguintes alterações: (1) os 6 distintos índices de Legibilidade (LE), em que VG e LE foram mensurados com base nas Notas Explicativas; e (2) os 6 distintos índices de Legibilidade (LE), em que VG e LE foram mensurados com base nos Relatórios da Administração. Todas as variáveis da pesquisa estão descritas na Tabela 1 abaixo. Para tratar os *outliers* da amostra, todas as variáveis foram *Winsorizadas* ao nível de 1%.

Tabela 1

Variáveis do Modelo Econométrico

Sigla	Descrição	Mensuração	Fonte	Sinal Esperado	Literatura
Variável Dependente					
VG	Violação Gramatical	Número de VG praticadas por companhia dividido pelo total de palavras complexas de cada companhia.	CVM		Gillette e Pundrich (2020).
Variável Independente de Interesse					
LE	Legibilidade	Escala de 0 a 100, classificando o texto, por nível de facilidade de leitura, do nível de legibilidade mais difícil para o mais fácil.	CVM	(-)	Lehavy <i>et al.</i> (2011).
Variáveis Independentes de Controle					
GR	Gerenciamento de Resultados	<i>Accruals</i> discricionários estimados pelo Modelo de Jones Modificado, sendo o resíduo da regressão em valor absoluto.	Comdinheiro	(+)	Dechow <i>et al.</i> (1995).
ROA	Retorno sobre o Ativo	Razão entre o lucro operacional sobre o ativo médio de cada companhia.	Comdinheiro	(-)	Elfeky (2017).
TAM	Tamanho da Empresa	Logaritmo natural do ativo total de cada companhia.	Comdinheiro	(-)	Karajeh, Ibrahim e Lode (2017).
ENDIV	Endividamento Total	Divisão de empréstimos e financiamentos de curto e longo prazo pelo ativo total de cada companhia.	Comdinheiro	(-)	Egbunike e Tarilaye (2017).
MTB	<i>Market-to-Book</i>	Razão entre o valor de mercado sobre o patrimônio líquido de cada companhia.	Comdinheiro	(-)	Almeida, Lopes e Corrar (2013).
TAU	Troca de Auditor	Variável <i>dummy</i> , sendo 1 para companhias que trocaram de auditor e 0 para as demais.	Comdinheiro	(+)	Borba <i>et al.</i> (2019).
CF	Conselho Fiscal	Variável <i>dummy</i> , sendo 1 para companhias que possuem	Comdinheiro	(-)	Borba <i>et al.</i> (2019).

Realização

		conselho fiscal e 0 para as demais.			
RET	Retorno das Ações	Preço da ação em t - preço da ação em t-1, dividido pelo preço da ação em t-1.	Comdinheiro	(-)	Segura, Cabral e Tardio (2021)
TCA	Tamanho do Conselho Administrativo	Logaritmo natural da quantidade de membros do conselho administrativo de cada companhia.	Comdinheiro	(-)	Borba <i>et al.</i> (2019).
BIG	Big Four	Variável <i>dummy</i> , sendo 1 para companhias auditadas por Big Four e 0 para as demais.	Comdinheiro	(-)	Borba <i>et al.</i> (2019).
Variável de Interação					
GR*LE	Interação entre Gerenciamento de Resultados e Legibilidade			(-)	Lehavy <i>et al.</i> (2011).
Segregação das Variáveis VG e LE pelos Índices de Legibilidade					
VGNE	Violação Gramatical - Notas Explicativas				
INGNE	Legibilidade Gunning Fog - Notas Explicativas				
ILANE	Legibilidade leitura automatizado - Notas Explicativas				
NGNE	Legibilidade Flesch-Kincaid - Notas Explicativas				
IGNE	Legibilidade Gulpease - Notas Explicativas				
FLFNE	Legibilidade facilidade de leitura de Flesch - Notas Explicativas				
ICLNE	Legibilidade Coleman-Liau - Notas Explicativas				
VGRA	Violação Gramatical – Relatórios da Administração				
INGRA	Legibilidade Gunning Fog - Relatórios da Administração				
ILARA	Legibilidade leitura automatizado - Relatórios da Administração				
NGRA	Legibilidade Flesch-Kincaid - Relatórios da Administração				
IGRA	Legibilidade Gulpease - Relatórios da Administração				
FLFRA	Legibilidade facilidade de leitura de Flesch - Relatórios da Administração				
ICLRA	Legibilidade Coleman-Liau - Relatórios da Administração				

Fonte: Elaborado pelos Autores.

Sobre o modelo econométrico, por possuir variável dependente quantitativa que tem valor inicial igual a zero, como análise de robustez, as regressões foram analisadas por dois estimadores (efeito fixo e de censura), ambos utilizados para modelos de regressão linear para dados em painel. A principal diferença é que o efeito de censura é uma alternativa para variáveis dependentes quantitativas cujo valor inicial é zero e, assim, é realizado a censura da variável dependente em função de seu valor inferior.

Também como análise de robustez complementar, nós utilizamos o modelo de Kothari, Leone e Wasley (2005) em complemento ao modelo de Dechow *et al.* (1995). A diferença entre eles reside no fato de que o modelo de Kothari *et al.* (2005) inclui o Retorno sobre os Ativos (ROA) na estimação do gerenciamento de resultados para controlar pela performance das firmas.

4. Análise dos Resultados

4.1. Resultados da Estatística Descritiva

Realização

A Tabela 2 apresenta a estatística descritiva sobre o resumo do comportamento das variáveis presentes no modelo econométrico, contendo o número de observações, a média, o desvio padrão, assim como os valores mínimo e máximo de cada variável.

Tabela 2

Estatística Descritiva das Variáveis dos Modelos

Variável	Obs.	Média	Desv. Pad.	Min	Max	p25	p50	p75
VGNE	965	6,362	3,182	0,000	26,985	4,441	6,010	7,816
INGNE	965	14,63	3,801	9,939	28,030	12,081	13,396	15,741
ILANE	965	11,422	3,654	6,564	23,878	9,029	10,095	12,702
NGNE	965	12,358	2,767	8,561	22,364	10,624	11,298	13,176
IGNE	965	59,793	11,022	36,385	85,523	51,888	59,919	66,920
FLFNE	965	36,746	9,441	0,000	49,302	33,940	39,862	42,814
ICLNE	965	13,529	1,367	9,432	17,773	12,751	13,311	14,224
VGRA	1018	6,897	3,654	0,000	31,746	4,515	6,558	8,895
INGRA	1018	19,377	4,087	10,126	31,815	16,811	19,859	21,712
ILARA	1018	14,942	3,772	5,729	26,286	12,652	15,312	17,131
NGRA	1018	14,988	2,852	8,981	23,917	12,991	15,173	16,559
IGRA	1018	50,786	6,621	40,601	76,133	46,950	48,742	52,400
FLFRA	1018	33,961	9,119	2,363	53,577	29,394	34,264	39,649
ICLRA	1018	13,045	1,159	8,795	16,209	12,514	13,180	13,750
GR	1018	0,056	0,068	0,000	0,555	0,010	0,036	0,077
BIG FOUR	1018	0,373	0,484	0,000	1,000	0,000	0,000	1,000
RET	1018	0,279	1,669	-0,838	13,103	-0,296	-0,031	0,300
ROA	1018	0,059	0,102	-0,518	0,388	0,027	0,066	0,111
TAM	1018	22,076	1,564	17,474	25,910	20,950	22,035	23,142
ENDIV	1018	0,320	0,211	0,000	1,382	0,169	0,306	0,429
MTB	1018	1,908	2,244	-5,249	13,323	0,661	1,331	2,607
CF	1018	0,822	0,383	0,000	1,000	1,000	1,000	1,000
TAU	1018	0,516	0,500	0,000	1,000	0,000	1,000	1,000
TCA	1018	1,914	0,330	1,099	2,773	1,609	1,946	2,197

Fonte: Elaborado pelos Autores.

Pela análise da Tabela 2, a *proxy* de Violação Gramatical das Notas Explicativas (VGNE), numa escala de 0% a 100%, variou de 0% a 26,98%. Resultado que reflete, descritivamente, um baixo nível de VGNE das Companhias de capital aberto no Brasil. Para a *proxy* de Violação Gramatical dos Relatórios da Administração (VGRA), numa escala de 0% a 100%, variou de 0% a 31,74%. Resultado que também reflete, descritivamente, um baixo nível de VGRA das Companhias de capital aberto no Brasil. Comparando esses resultados descritivos de VGNE e VGRA, observa-se que o nível de Violação Gramatical das Companhias, com base nos Relatórios da Administração, foi superior ao nível de Violação Gramatical considerando as Notas Explicativas.

4.2. Resultados da Correlação Entre As Variáveis

A matriz de correlação de Spearman e Pearson das variáveis presentes no modelo econométrico, em que constam a existência de correlação positiva ou negativa entre as variáveis. Os coeficientes de correlação de Spearman são apresentados abaixo da diagonal, enquanto os coeficientes de correlação de Pearson são apresentados acima da diagonal.

Os resultados da matriz de correlação de Spearman e Pearson demonstraram correlação entre as variáveis de interesse e de controle com as variáveis dependentes e não apresentaram problemas de multicolinearidade. Com o objetivo de obter indícios iniciais da relação entre Legibilidade e Violação Gramatical, nossos resultados indicaram correlação significativa e

Realização

negativa dos seguintes índices de Legibilidade: INGNE, ILANE, ILANE, NGNE, ICLNE e ICLRA. E também correlação significativa e positiva dos seguintes índices de Legibilidade: IGNE e FLFNE.

Os índices de Legibilidade com correlação significativa e negativa com Violação Gramatical sugerem indícios de associação entre essas variáveis, o que corrobora com a pesquisa de Gillette e Pundrich (2020). As demais variáveis de interesse e de controle também apresentaram correlação significativa. A tabela de matriz de correlação por economia de espaço no artigo, foi suprimida e pode ser divulgada mediante solicitação.

4.3. Resultados Legibilidade e Violação Gramatical

Conforme descrito na subseção da Metodologia, esses 6 índices de Legibilidade usam duas escalas: 0 a 100 (Teste de facilidade de leitura de Flesch e Índice Gulpease); e 0 a 20 (Nível de instrução de Flesch-Kincaid, Índice de Nebulosidade de Gunning, Índice de Leiturabilidade Automatizado, Índice de Coleman-Liau, entre outros). Os índices da escala de 0-100 possuem uma interpretação por ordem numérica crescente, em que zero indica um texto de difícil compreensão e cem sugere um texto de fácil compreensão. Os índices da escala 0-20 têm uma interpretação por ordem numérica decrescente, em que zero indica um texto de fácil compreensão e 20 sugere um texto de difícil compreensão.

A Tabela 3 descreve o resultado da regressão linear múltipla com dados em painel de efeito fixo de uma série temporal de 12 anos (2010-2021) das empresas não financeiras listadas na [B]³.

Tabela 3

Regressões Legibilidade e Violação Gramatical

Painel A: Notas Explicativas						
Var. Dep.: VG	INGNE	FLFNE	ICLNE	ILANE	NGNE	IGNE
LE	-0,084 (0,081)	0,096*** (0,03)	-0,782** (0,39)	-0,123 (0,089)	-0,173 (0,115)	0,072 (0,044)
GR	4,014** (1,678)	1,343 (1,392)	-9,561 (14,573)	3,789** (1,614)	4,010** (1,680)	-3,472 (7,849)
GR*LE	-0,123*** (0,055)	0,04* (0,022)	0,805 (1,066)	-0,133** (0,052)	-0,14** (0,063)	0,086 (0,128)
TAM	-0,558** (0,253)	-0,543** (0,260)	-0,595** (0,271)	-0,559** (0,254)	-0,562** (0,254)	-0,714*** (0,265)
BIG FOUR	-0,367* (0,187)	-0,39** (0,178)	-0,416** (0,183)	-0,361* (0,185)	-0,360* (0,185)	-0,376** (0,184)
RET	0,013 (0,033)	0,009 (0,034)	0,02 (0,037)	0,012 (0,033)	0,014 (0,034)	0,006 (0,037)
TAU	-0,556** (0,242)	-0,549** (0,229)	-0,473** (0,2)	-0,558** (0,235)	-0,559** (0,240)	-0,517** (0,222)
CF	0,43 (0,444)	0,316 (0,412)	0,327 (0,374)	0,408 (0,434)	0,409 (0,433)	0,281 (0,41)
ROA	1,286 (0,991)	1,238 (0,960)	0,901 (0,912)	1,299 (0,980)	1,289 (0,977)	1,049 (0,992)
ENDIV	-0,359 (0,829)	-0,563 (0,789)	-0,456 (0,838)	-0,390 (0,831)	-0,390 (0,834)	-0,737 (0,812)
MTB	-0,062 (0,046)	-0,08 (0,049)	-0,068 (0,048)	-0,067 (0,046)	-0,068 (0,046)	-0,068 (0,047)
TCA	-0,234 (0,594)	-0,208 (0,570)	-0,243 (0,578)	-0,205 (0,593)	-0,204 (0,590)	-0,234 (0,608)
CONSTANTE	20,112***	15,163**	30,515***	20,296***	21,079***	18,191***

Realização

	(5,398)	(6,260)	(7,981)	(5,493)	(5,452)	(6,425)
OBSERVAÇÕES	965	965	965	965	965	965
R-QUADRADO	0,05	0,121	0,116	0,06	0,062	0,092
PROB > F	0,009	0,000	0,036	0,004	0,003	0,000
ANO EF	Sim	Sim	Sim	Sim	Sim	Sim
FIRMA EF	Sim	Sim	Sim	Sim	Sim	Sim
Painel B: Relatórios da Administração						
Var. Dep.: VG	INGRA	FLFRA	ICLRA	ILARA	NGRA	IGRA
LE	0,092* (0,052)	0,028 (0,024)	-0,221 (0,236)	0,081 (0,055)	0,07 (0,069)	-0,037 (0,037)
GR	12,866* (7,498)	-7,448** (3,086)	-33,141* (18,87)	9,483 (5,848)	10,981 (7,978)	-10,783 (12,682)
GR*LE	-0,736* (0,384)	0,169** (0,083)	2,402* (1,43)	-0,726* (0,387)	-0,83 (0,53)	0,179 (0,249)
TAM	-0,739** (0,372)	-0,786** (0,367)	-0,776** (0,382)	-0,748** (0,373)	-0,757** (0,371)	-0,786** (0,375)
BIG FOUR	-0,493* (0,266)	-0,439 (0,27)	-0,414 (0,276)	-0,487* (0,265)	-0,474* (0,27)	-0,454* (0,261)
RET	-0,074 (0,049)	-0,08 (0,051)	-0,067 (0,047)	-0,075 (0,049)	-0,074 (0,049)	-0,076 (0,047)
TAU	0,126 (0,272)	0,17 (0,267)	0,092 (0,266)	0,138 (0,27)	0,152 (0,271)	0,139 (0,271)
CF	-0,565 (1,001)	-0,584 (0,985)	-0,551 (0,981)	-0,573 (1,005)	-0,587 (1,005)	-0,616 (1,006)
ROA	1,141 (1,616)	0,961 (1,69)	1,087 (1,644)	1,093 (1,634)	1,02 (1,642)	1,098 (1,649)
ENDIV	3,102** (1,262)	3,104** (1,293)	3,214** (1,291)	3,109** (1,267)	3,123** (1,275)	3,058** (1,271)
MTB	0,029 (0,083)	0,024 (0,082)	0,029 (0,08)	0,031 (0,083)	0,029 (0,083)	0,035 (0,081)
TCA	0,741 (0,868)	0,844 (0,89)	0,791 (0,882)	0,734 (0,873)	0,766 (0,878)	0,704 (0,889)
CONSTANTE	19,685** (7,74)	21,277*** (7,616)	25,047*** (7,988)	20,446*** (7,684)	20,733*** (7,695)	24,331*** (7,904)
OBSERVAÇÕES	1018	1018	1018	1018	1018	1018
R-QUADRADO	0,053	0,054	0,054	0,051	0,049	0,045
PROB > F	0,000	0,000	0,000	0,000	0,000	0,000
ANO EF	Sim	Sim	Sim	Sim	Sim	Sim
FIRMA EF	Sim	Sim	Sim	Sim	Sim	Sim

Erros-padrão nos parênteses

*** $p < .01$, ** $p < .05$, * $p < .1$

Fonte: Elaborado pelos Autores.

A Tabela 3 – Painel A, indica que a variável LE para os índices (FLFNE e ICLNE) tem associação forte com a variável dependente VG (coeficientes 0,096 e 0,782 e p-valores 0,01 e 0,05, respectivamente), sendo FLFNE positivo e ICLNE negativo. Logo, considerando FLFNE, para cada 1% incremental de LE, a VG aumentará em 0,096 o seu grau de Violação Gramatical. Resultado que, para o índice FLFNE, rejeita a hipótese de pesquisa (H1). Esse resultado sugere que, para o FLFNE, índice que elevado valor infere-se alta Legibilidade, quanto maior o nível de legibilidade, em média, maior o grau de violação gramatical das notas explicativas das companhias.

A Tabela 3 – Painel B, indica que a variável LE para o índice (INGRA) tem associação

Realização

fraca e positiva com a variável dependente VG (coeficiente 0,092 e p-valor 0,10). Logo, considerando INGRA, para cada 1% incremental de LE, a VG aumentará em 0,092 o seu grau de Violação Gramatical. Resultado que, para o índice INGRA, rejeita a hipótese de pesquisa (H1). Esse resultado sugere que, para o INGRA, índice que elevado valor infere-se baixa Legibilidade, quanto maior o nível de legibilidade, em média, maior o grau de violação gramatical dos relatórios da administração das companhias.

O resultado que rejeitou a hipótese de pesquisa (H1) para o índice FLFNE contraria as evidências da pesquisa de Gillette e Pundrich (2020), que verificou que as Violações Gramaticais são um sinal da qualidade das demonstrações contábeis porque estão correlacionados com o tempo, esforço e recursos dedicados na preparação desses relatórios. Logo, o nível de Violação Gramatical das notas explicativas, embora possa gerar evidências sobre a qualidade das informações contábeis presentes nesses relatórios, essa relação pode não se confirmar, dependendo das características das informações textuais a que se reportam.

Pela análise da Tabela 3 – Painel A, o resultado de ICLNE confirmou a hipótese de pesquisa (H1), então, para cada 1% incremental de LE, a VG diminuirá em 0,78 o seu grau de Violação Gramatical. O resultado de LE do índice ICLNE (coeficiente 0,782 e p-valor 0,05), índice que elevado valor infere-se baixa Legibilidade, sugere que quanto maior o nível de legibilidade, em média, menor o grau de violação gramatical das notas explicativas das companhias. O que significa a não rejeição da hipótese de pesquisa (H1) para o índice ICLNE.

O resultado do índice ICLNE, que não rejeitou a hipótese de pesquisa (H1), corrobora com as evidências da pesquisa de Gillette e Pundrich (2020), que verificou que as Violações Gramaticais são um sinal da qualidade das demonstrações contábeis porque estão correlacionados com o tempo, esforço e recursos dedicados na preparação desses relatórios. Portanto, o nível de Violação Gramatical das notas explicativas permite gerar evidências sobre a qualidade das informações contábeis presentes nesses relatórios, servindo como um novo método de investigação mais prático para a mensuração da qualidade dessas informações.

Conforme a Tabela 3 – Painel A, a variável GR tem associação positiva com a variável dependente VG para os índices (INGNE, coeficiente 4,01 e p-valor 0,05), (ILANE, coeficiente 3,78 e p-valor 0,05) e (NGNE, coeficiente 4,01 e p-valor 0,05). Logo, quando a empresa aumenta 1% seu grau de gerenciamento de resultados, considerando o índice INGNE, a VG aumentará em 4,01, em média, o seu grau de Violação Gramatical. Esse resultado indica que quanto maior o gerenciamento de resultados, em média, maior o grau de violação gramatical das notas explicativas das companhias.

Conforme a Tabela 3 – Painel B, a variável GR para os índices (INGRA, FLFRA e ICLRA) tem associação com a variável dependente VG (coeficientes 12,86, 7,44 e 33,14 e p-valores 0,10, 0,05 e 0,010, respectivamente), sendo INGRA positivo e FLFRA e ICLRA negativos. Logo, para INGRA, quando a empresa aumenta 1% seu grau de gerenciamento de resultados, a VG aumentará em 12,86, o seu grau de Violação Gramatical. Esse resultado, para INGRA, índice que elevado valor infere-se baixa Legibilidade, indica que quanto maior o gerenciamento de resultados, em média, maior o grau de violação gramatical dos relatórios da administração das companhias. Já para FLFRA, índice que elevado valor infere-se alta Legibilidade, e ICLRA, índice que elevado valor infere-se baixa Legibilidade, o resultado indica que quanto maior o gerenciamento de resultados, em média, menor o grau de violação gramatical dos relatórios da administração das companhias.

Realização



UFRJ
UNIVERSIDADE FEDERAL
DO RIO DE JANEIRO



**Universidade
Federal
Fluminense**



Programa de Pós-Graduação em Administração - UFF

Com base na Tabela 3 – Painel A, a variável TAM tem associação negativa com a variável dependente VG para os índices (INGNE, coeficiente 0,55 e p-valor 0,05), (FLFNE, coeficiente 0,54 e p-valor 0,05), (ICLNE, coeficiente 0,59 e p-valor 0,05), (ILANE, coeficiente 0,55 e p-valor 0,05), (NGNE, coeficiente 0,56 e p-valor 0,05) e (IGNE, coeficiente 0,71 e p-valor 0,01). Logo, considerando ICLNE, quando a empresa aumenta 1 Real em seu ativo total, a VG diminuirá em 0,59, em média, o seu grau de Violação Gramatical. Esse resultado indica que quanto maior o tamanho, em média, menor o grau de violação gramatical das notas explicativas das companhias.

Com base na Tabela 3 – Painel B, a variável TAM tem associação negativa com a variável dependente VG para os índices (INGRA, coeficiente 0,73 e p-valor 0,05), (FLFRA, coeficiente 0,78 e p-valor 0,05), (ICLRA, coeficiente 0,77 e p-valor 0,05), (ILARA, coeficiente 0,74 e p-valor 0,05), (NGRA, coeficiente 0,75 e p-valor 0,05) e (IGRA, coeficiente 0,78 e p-valor 0,05). Logo, considerando FLFRA, quando a empresa aumenta 1 Real em seu ativo total, a VG diminuirá em 0,78, em média, o seu grau de Violação Gramatical. Esse resultado indica que quanto maior o tamanho, em média, menor o grau de violação gramatical dos relatórios da administração das companhias.

Considerando a Tabela 3 – Painel A, a variável BIG FOUR tem associação negativa com a variável dependente VG para os índices (INGNE, coeficiente 0,36 e p-valor 0,10), (FLFNE, coeficiente 0,39 e p-valor 0,05), (ICLNE, coeficiente 0,41 e p-valor 0,05), (ILANE, coeficiente 0,36 e p-valor 0,10), (NGNE, coeficiente 0,36 e p-valor 0,10) e (IGNE, coeficiente 0,37 e p-valor 0,05). Esse resultado indica que companhias auditadas por BIG FOUR, em média, possuem menor grau de violação gramatical das notas explicativas.

Considerando a Tabela 3 – Painel B, a variável BIG FOUR tem associação negativa com a variável dependente VG para os índices (INGRA, coeficiente 0,49 e p-valor 0,10), (ILARA, coeficiente 0,48 e p-valor 0,10), (NGRA, coeficiente 0,47 e p-valor 0,10) e (IGRA, coeficiente 0,45 e p-valor 0,10). Esse resultado indica que companhias auditadas por BIG FOUR, em média, possuem menor grau de violação gramatical dos relatórios da administração.

Pela análise da Tabela 3 – Painel B, a variável ENDIV tem associação positiva com a variável dependente VG para os índices (INGNE, coeficiente 3,10 e p-valor 0,05), (FLFNE, coeficiente 3,10 e p-valor 0,05), (ICLNE, coeficiente 3,21 e p-valor 0,05), (ILANE, coeficiente 3,10 e p-valor 0,05), (NGNE, coeficiente 3,12 e p-valor 0,05) e (IGNE, coeficiente 3,05 e p-valor 0,05). Logo, considerando INGNE, quando a empresa aumenta 1 Real em seu montante de dívida, a VG aumentará em 3,10, em média, o seu grau de Violação Gramatical. Esse resultado indica que quanto maior o endividamento, em média, maior o grau de violação gramatical dos relatórios da administração das companhias.

De acordo com a Tabela 3 – Painel A, a variável TAU tem associação negativa com a variável dependente VG para os índices (INGNE, coeficiente 0,55 e p-valor 0,05), (FLFNE, coeficiente 0,54 e p-valor 0,05), (ICLNE, coeficiente 0,47 e p-valor 0,05), (ILANE, coeficiente 0,55 e p-valor 0,05), (NGNE, coeficiente 0,55 e p-valor 0,05) e (IGNE, coeficiente 0,51 e p-valor 0,05). Esse resultado indica que companhias que realizaram rodízio de auditor possuem menor grau de Violação Gramatical.

5. Conclusão

Realização

Este estudo analisou a associação entre o grau de legibilidade e o nível de violação gramatical das demonstrações contábeis, considerando os efeitos das violações gramaticais sobre a qualidade textual das informações das companhias de capital aberto brasileiras. Os resultados sugerem que dentro do ambiente brasileiro, a *proxy* Violação Gramatical foi apresentada como uma *proxy* capaz de medir, de forma indireta, a qualidade das informações contábeis, reduzindo a assimetria de informações e minimizando conflitos de interesses entre companhias e investidores. Resultado similar encontrado pela pesquisa de Gillette e Pundrich (2020), a qual identificou que as Violações Gramaticais são um sinal da qualidade das demonstrações contábeis para as companhias norte-americanas.

A relação entre Legibilidade e Violação Gramatical, quando analisadas as notas explicativas, identificou que, quanto maior o nível de legibilidade, menor o grau de Violação Gramatical das companhias. Esse resultado foi sustentado pelos seguintes índices de legibilidade: Índice de Nebulosidade de Gunning e Índice de Coleman-Liau.

Já a associação entre Legibilidade e Violação Gramatical dos relatórios da administração, verificou que quanto maior o nível de legibilidade, menor o grau de Violação Gramatical dos relatórios da administração das companhias. Esse resultado foi sustentado pelo seguinte índice de legibilidade: Índice de Gulpease.

Sobre esse resultado, considerando a associação negativa entre Legibilidade e Violação Gramatical, as companhias podem reforçar sua estrutura de revisão de suas demonstrações contábeis para reduzirem erros que aumentam as Violações Gramaticais ou reduzem a Legibilidade.

Ainda sobre a relação entre Legibilidade e Violação Gramatical, considerando o maior número de índices de Legibilidade não rejeitados por parte das Notas Explicativas, esse resultado sugere que a relação entre Legibilidade e Violação Gramatical é mais expressiva com as Notas Explicativas do que com os Relatórios da Administração, possivelmente, devido a maior variabilidade em função do maior volume de páginas das Notas Explicativas em comparação com os Relatórios da Administração.

As evidências deste artigo geram informações relevantes na literatura contábil sobre os reflexos gramaticais dos relatórios contábeis sobre a qualidade da informação contábil e, também, sobre o desempenho econômico e financeiro das companhias brasileiras. Com destaque para a relevância da qualidade da escrita para uma comunicação eficiente entre companhia e os usuários das informações contábeis.

Os resultados deste artigo contribuem com a literatura e com o mercado, ao destacar que elevado nível de Violação Gramatical e baixo nível de Legibilidade podem diminuir a capacidade de compreensão da informação contábil. Para os auditores, é possível avaliar se alto grau de Violação Gramatical e baixa Legibilidade prejudicam seus trabalhos de auditoria e de revisão das demonstrações contábeis. Já para os investidores, textos redigidos com maior qualidade permitem diminuir os riscos de interpretações equivocadas das demonstrações contábeis, o que pode proporcionar aumento da confiança no processo de alocação de capital.

Este estudo não está isento de limitações. Assume-se que devido a estrutura padronizada de divulgação que as notas explicativas e os relatórios da administração possuem, foi encontrado pouca variabilidade entre os anos nas *proxies* Violação Gramatical e Legibilidade. Esse fato pode ter diminuído a capacidade dessas *proxies* capturarem completamente o que se esperava delas, obtendo menos evidências. Essa baixa variabilidade das *proxies* pode ter

reduzido o poder explicativo dos modelos econométricos e gerado resultados menos significativos entre as variáveis.

Por isso, sugere-se que pesquisas futuras abordem o tema Violação Gramatical das demonstrações contábeis, utilizando relatórios contábeis com formas de evidenciação mais variadas. Além disso, nós recomendamos que realizem a coleta de dados em outros mercados, assim como este artigo fez com o mercado financeiro brasileiro, para que o tema possa evoluir e se consolidar, gerando contribuições significativas para a literatura de qualidade da informação contábil.

Referências

- Almeida, J. E. F. D., Lopes, A. B., & Corrar, L. J. (2013). Gerenciamento de resultados para sustentar a expectativa do mercado de capitais: impactos no índice market-to-book. *ASAA-Advances in Scientific and Applied Accounting*, 4(1), 44-62. <https://asaa.anpcont.org.br/index.php/asaa/article/view/37>
- Análise de Legibilidade Textual (ALT): um software para análise de legibilidade de textos em Língua Portuguesa. Disponível em <https://doi.org/10.48550/arXiv.2203.12135> (2021).
- Barkemeyer, R., Comyns, B., Figge, F., & Napolitano, G. (2014). CEO statements in sustainability reports: Substantive information or background noise? *Accounting Forum*, 38(4), 241-257. <https://doi.org/10.1016/j.accfor.2014.07.002>
- 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>
- Borba, A. E., Dos Santos, D. A., Dagostini, L., & Schvirck, E. (2019). A Influência das características do conselho de administração, comitê de auditoria e conselho fiscal no audit delay. *Revista Gestão Organizacional*, 12(3), 158-175. <https://doi.org/10.22277/rgo.v12i3.5203>
- Coleman, M., & Liao, T. L. (1975). A computer readability formula designed for machine scoring. *Journal of Applied Psychology*, 60(2), 283-284. <https://doi.org/10.1037/h0076540>
- Consoni, S., & Lima, G. A. S. F. (2017). Voluntary disclosure and earnings management: evidence from the brazilian capital market. *Revista Contabilidade & Finanças*, 28(74), 249-263. <https://doi.org/10.1590/1808-057x201703360>
- Dechow, P. M., Sloan, R. G., & Sweeney, A. P. (1995). Detecting earnings management. *The Accounting Review*, 70(2), 193-225. <https://www.jstor.org/stable/248303>
- 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>
- Egbunike, A. P., & Tarilaye, N. (2017). Firm's specific attributes and voluntary environmental disclosure in Nigéria: evidence from listed manufacturing companies. *Academy of Accounting and Financial Studies Journal*, 21(3), 1-9.
- 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>
- Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology*, 32(3), 221-233. <https://doi.org/10.1037/h0057532>

- Gillette, J., & Pundrich, G. (2020). Grammatical Violations and Financial Reporting Quality. Available at SSRN 3496434. <https://ssrn.com/abstract=3496434>
- 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>
- 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>
- Gunning, R. (1952) *The Technique of Clear Writing*. McGraw-Hill, New York.
- 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>
- Holtz, L., & dos Santos, O. M. (2020). Legibilidade das notas explicativas das empresas brasileiras de capital aberto. *Enfoque: Reflexão Contábil*, 39(1), 57-73. <https://doi.org/10.4025/enfoque.v39i1.45275>
- Hucks, R. J. (2015). *Voluntary Involuntary Disclosure*. Unpublished Doctoral Thesis, Doctor of Philosophy (Business Administration), Universidade de Michigan. Ann Arbor, Estados Unidos da América.
- Índice de Gulpease. Disponível em: https://farfalla-project.org/readability_static/ (1988).
- Karajeh, A. I., Ibrahim, M. Y. B., & Lode, N. A. B. (2017). Impact of shareholder structure on voluntary disclosure in malaysian companies. *Global Business and Management Research*, 9(1), 142-153. <https://doi.org/10.1080/01559982.2019.1605874>
- Kincaid, J. P., Fishburne Jr, R. P., Rogers, R. L., & Chissom, B. S. (1975). Derivation of new readability formulas (automated readability index, fog count and flesch reading ease formula) for navy enlisted personnel. *Naval Technical Training Command Millington TN Research Branch*. <https://stars.library.ucf.edu/istlibrary/56>
- Kothari, S. P., Leone, A. J., & Wasley, C. E. (2005). Performance matched discretionary accrual measures. *Journal of accounting and economics*, 39(1), 163-197. <https://doi.org/10.1016/j.jacceco.2004.11.002>
- 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>
- 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>

- Monteiro, J. J., Rengel, R., De Sousa, A. M., & Borba, J. A. (2021). Não basta ler, é preciso compreender: um enfoque na legibilidade do resultado e do EBITDA. *Revista Universo Contábil*, 16(3), 31-49. <http://dx.doi.org/10.4270/ruc2020314>
- Morais, A. I. (2020). Are changes in international accounting standards making them more complex? *Accounting Forum*, 44(1), 35-63. <https://doi.org/10.1080/01559982.2019.1573781>
- 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>
- Rody, P. H. A., & Almeida, J. E. F. (2023). Análise Sistemática e Bibliométrica da Literatura sobre Violação Gramatical nas Demonstrações Contábeis. *Revista De Educação E Pesquisa Em Contabilidade (REPeC)*, 17(2), 118-139. <https://doi.org/10.17524/repec.v17i2.3289>
- 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>
- Segura, L. C., Cabral, C. G., & Tardio, D. G. C. (2021). A relação entre a divulgação do resultado contábil e o retorno das ações das companhias abertas brasileiras em decorrência do covid-19. *Revista Eletrônica do Departamento de Ciências Contábeis & Departamento de Atuária e Métodos Quantitativos*, 8(1), 144-161. <https://doi.org/10.23925/2446-9513.2021v8i1p144-161>
- Senter, R. J., & Smith, E. A. (1967). Automated readability index. *Cincinnati Univ OH*.
- 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>
- Souza, J. A. S. D., & Borba, J. A. (2022). Readability como medida de complexidade textual: determinantes e evidências em empresas brasileiras. *Revista Contabilidade & Finanças*, 33, 112-129. <https://doi.org/10.1590/1808-057x202114180>
- Telles, S. V., & Salotti, B. M. (2021). Intelligibility vs readability: understandability measures of financial information. *Revista Universo Contábil*, 16(2), 110-126. <http://dx.doi.org/10.4270/ruc.2020209>
- 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>
- Williams, J. M. (1981). The phenomenology of error. *College composition and communication*, 32(2), 152-168. <https://doi.org/10.2307/356689>
- Wooldridge, J. M. (2002). *Econometric analysis of cross section and panel data*. Cambridge: MIT Press.