

Harmonia entre as Informações Contidas no Texto e nos Números dos Relatórios Anuais de Empresas Brasileiras

Resumo

O presente trabalho teve o objetivo de analisar a harmonia entre as informações transmitidas nas seções narrativas dos relatórios anuais e o desempenho financeiro das empresas. A amostra foi composta por 120 empresas dentre as listadas na BM&FBovespa no ano de 2009, sendo as 60 com maior variação positiva e as 60 com maior variação negativa de resultado contábil líquido. Foram selecionadas palavras-chave referentes a três temas centrais: rentabilidade, crescimento e gestão e por meio de análise de conteúdo foi avaliado o sentido que estas palavras-chave apareciam nos relatórios (positivo quantitativo, positivo qualitativo, negativo quantitativo e negativo qualitativo). Obtidas as frequências, foram realizadas duas regressões logísticas para comparação do texto com os números, sendo uma pela frequência bruta de palavras e outra pela ponderação de termos. Os resultados indicam que, nas informações ligadas ao tema rentabilidade, o texto é harmônico com os números. Nas informações referentes ao crescimento, a harmonia é parcial. E quando se trata de informações ligadas à gestão, existe conflito entre as seções narrativas e o desempenho das empresas. Finalmente, constatou-se que quanto mais subjetivas são as informações, maior o conflito.

Palavras-chave: Relatórios Anuais; Dissonância cognitiva; Análise de Conteúdo.

Marcelo Sanches Pagliarussi

Doutor em Ciência e Engenharia de Materiais pela Universidade Federal de São Carlos e Professor Doutor da Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto da Universidade de São Paulo. **Contato:** Av. Bandeirantes 3900, Monte Alegre, Ribeirão Preto, SP, CEP.: 14040-905.
E-mail: marcelosp@usp.br

Thiago Neiva Guimarães

Mestre em Ciências Contábeis pela Fundação Instituto Capixaba de Pesquisas em Contabilidade, Economia e Finanças e Professor Adjunto das Faculdades Unidas do Norte de Minas. **Contato:** Rua Lírio Brant, 787, Melo, Montes Claros, MG, CEP.: 39401-063.
E-mail: thiaguimaraes@gmail.com

Tiago Alves Ferreira

Mestrando em Ciências Contábeis pela Fundação Instituto Capixaba de Pesquisas em Contabilidade, Economia e Finanças. **Contato:** Av. Fernando Ferrari, 1358, Boa Vista, Vitória, ES, CEP.: 29075-505.
E-mail: tiagorev@gmail.com

Os autores agradecem à Fapes – Fundação de Apoio à Ciência e Tecnologia do Espírito Santo, pelo apoio financeiro para realização deste trabalho.

1. Introdução

Tradicionalmente, a avaliação do desempenho de uma organização se baseia na análise de informações numéricas, como os números contábeis. Entretanto, as informações textuais também podem ser úteis para avaliar o potencial de desempenho futuro de uma organização (Clatworthy & Jones, 2003). Tais informações contextualizam os números contábeis por meio de uma narrativa que descreve as atividades e ações desenvolvidas pelas empresas em seus mercados-alvo, seus planos futuros, projetos, linhas de produto, esforços de redução de custos, venda de ativos, etc.

Recentemente, um corpo de pesquisa substancial foi desenvolvido com o foco na análise das informações textuais divulgadas pelas organizações e sua utilidade para tomada de decisão em um contexto de mercado de capitais. Os estudos envolvem fenômenos diversos, como a legibilidade dos relatórios anuais corporativos em função do desempenho (Li, 2008); o uso de táticas de gerenciamento da reputação corporativa em relatórios anuais (Vargas, Almeida & Maria Júnior, 2014; Ribeiro, Pagliarussi, Silva & Silva Júnior, 2011; Tessarolo, Pagliarussi & da Luz, 2010; da Luz, Pagliarussi, Teixeira & Baptista, 2009); o tom dos textos financeiros (Loughran & McDonald, 2011); e a evidenciação de informações de caráter estratégico nos relatórios anuais (Pagliarussi & Liberato, 2011).

O foco do presente estudo é a harmonia entre as informações textuais e as numéricas apresentadas nos relatórios de empresas listadas no Brasil. A pertinência do questionamento a respeito da harmonia entre os dois conjuntos de informação deriva da maior discricionariedade permitida aos gestores das organizações na elaboração da narrativa (Balata & Breton, 2005). Ao desenvolver estudo semelhante com empresas canadenses, os autores observaram um moderado nível de divergência entre o texto e os números, porém analisaram apenas as cartas do presidente de 30 empresas.

Em outro estudo sobre o tema, Smith (1998) demonstrou, por meio de um experimento, que a existência de mensagens conflitantes impactou as decisões tomadas pelos sujeitos da pesquisa. Beynon, Clatworthy e Jones (2004), buscando verificar a correspondência do texto em relação aos números, desenvolveram um método de *data mining* para tentar classificar, por meio de uma análise textual de cartas do presidente, empresas rentáveis e não rentáveis do Reino Unido e tiveram sucesso na aplicação do método, com alta acurácia preditiva e um elevado nível de classificações corretas.

Visando contribuir para o entendimento do relacionamento entre as informações textuais e as informações numéricas apresentadas nos relatórios de empresas listadas, o presente trabalho teve o objetivo de analisar se há harmonia nas mensagens transmitidas nas seções narrativas dos relatórios anuais, em relação aos números contábeis referentes ao desempenho da empresa. A sustentação teórica do presente trabalho se baseia nas proposições de Milgrom (1981), relativas ao jogo de persuasão, no qual um interessado em fornecer informações tenta influenciar a decisão de outros; e de Grossman e Hart (1980), que afirmam que a parte que possui informação relevante nada tem a ganhar, ocultando ou mentindo a respeito da informação, desde que seja possível ao receptor verificar a se as informações são verdadeiras e completas.

O presente estudo contribui para a literatura que trata da coerência entre informações textuais e informações numéricas por meio da análise de diferentes agrupamentos temáticos em relação aos estudos anteriores de Vargas, Almeida e Maria Júnior (2014), Ribeiro, Pagliarussi, Silva e Silva Júnior (2011), Tessarolo, Pagliarussi e da Luz (2010), e da Luz, Pagliarussi, Teixeira e Baptista (2009). Além de analisarmos os temas Rentabilidade, Crescimento e Gestão, toda a seção narrativa contida nos relatórios anuais foi objeto de análise, e não apenas a carta do presidente, como o estudo de Balata e Breton (2005).

Para atingir os objetivos propostos, foram coletados no *site* da BM&FBovespa os relatórios anuais de todas as empresas listadas no ano de 2009 e selecionados os 120 com maior variação de resultado contábil líquido, sendo os 60 com maior variação positiva e os 60 com maior variação negativa. A técnica de análise de conteúdo foi aplicada aos 120 relatórios, por meio da análise da frequência de palavras-chave. Tais palavras-chave foram pré-selecionadas em correspondência aos temas de Rentabilidade, Crescimento e Gestão e sua frequência foi categorizada, levando-se em consideração a natureza, quantitativa ou qualitativa, e a polaridade, positiva ou negativa. A partir das frequências, foram obtidas as variáveis para a aplicação da técnica de regressão logística.

Foram realizadas duas regressões logísticas, sendo uma pela frequência bruta da ocorrência de cada palavra-chave nos textos e outra pela ponderação de termos desenvolvida por Manning e Schütze (1999). Os resultados mostraram que, para as informações referentes à rentabilidade, o texto é consistente com os números. Já para as informações referentes a crescimento e gestão existe um conflito entre as duas seções.

2. Referencial Teórico e Desenvolvimento das Hipóteses

2.1 Dispositivos Legais Brasileiros Sobre Relatórios Anuais da Administração

No Brasil, segundo o Art. 133 da Lei n.º 6.404(1976), todas as empresas classificadas como sociedades anônimas devem publicar, anualmente, entre outros, o relatório da administração sobre os negócios sociais e os principais fatos administrativos do exercício findo, em conjunto com as demonstrações financeiras. Nas seções narrativas dos relatórios, as empresas podem explicar com maior riqueza de detalhes os fatos ocorridos durante o exercício e projeções futuras.

O parecer de orientação da Comissão de Valores Mobiliários (CVM) n.º 15, de 28 de dezembro de 1987 (CVM, 1987), define também os tópicos e o conteúdo mínimo que deveria estar contido nas seções narrativas, como a descrição dos negócios, produtos e serviços; comentários sobre a conjuntura econômica geral; recursos humanos, números de empregados, investimento em treinamento, fundos de seguridade e planos sociais; investimentos realizados; pesquisa e desenvolvimento; novos produtos e serviços; proteção ao meio ambiente; reformulações administrativas; investimentos em controladas e coligadas; direitos dos acionistas e dados de mercado; perspectivas e planos para o exercício em curso e exercícios futuros.

Ainda, segundo esse parecer, p.2: “a divulgação de informações úteis, fidedignas e detalhadas, que possibilitem o conhecimento da companhia e de seus objetivos e políticas, é um direito essencial do acionista”. Outro trecho diz que o texto dos relatórios deve:

[...]ser redigido com simplicidade de linguagem para ser acessível ao maior número possível de leitores, devendo ser evitados adjetivos e frases tais como “excelente resultado”, “ótimo desempenho”, “baixo endividamento”, “excelentes perspectivas”, a menos que corroborado por dados comparativos ou fatos (CVM, 1987, p.2).

Além disso, o Art. 25 da Instrução CVM n.º 308 (CVM, 1999), diz que o auditor independente deve verificar “se as informações e análises contábeis e financeiras apresentadas no relatório da administração da entidade estão em consonância com as demonstrações contábeis auditadas”.

Percebe-se a importância dos relatórios como fonte de informação para todos os *stakeholders* da empresa. Por exemplo, a maior parte das informações citadas pelos analistas é retirada nas seções narrativas (Rogers e Grant, 1997).

2.2 Estudos Anteriores

Tratando-se de seções narrativas, as pesquisas na área podem ser categorizadas em dois tipos: legibilidade e análise de conteúdo (Beynon, Clatworthy & Jones, 2004). Os estudos em legibilidade procuram avaliar a dificuldade no entendimento dos textos, enquanto que na análise de conteúdo existe uma variedade de métodos, entre os quais está a interpretação do conteúdo do texto por quem elabora o trabalho, com contagem de palavras ou frases de temas predeterminados com o auxílio de *software* (Beynon, Clatworthy & Jones, 2004).

Entre os estudos envolvendo análise de conteúdo, foco do presente artigo, Balata e Breton (2005) analisaram as cartas do presidente de 30 empresas canadenses entre 1993 e 1998. Eles questionaram a harmonia entre as informações contidas nas cartas de presidente e os números contábeis, verificando a

presença de informações contraditórias que tendiam a produzir um efeito dissonante. Os autores criaram um índice do nível de otimismo contido nas cartas e nas demonstrações financeiras, analisando três temas centrais: Rentabilidade, Crescimento e Gestão. Para cada tema, foram determinadas palavras-chave para avaliar o otimismo na parte textual e indicadores financeiros associados aos temas para avaliar o otimismo contido nas demonstrações financeiras. Como resultados, eles encontraram um moderado nível de divergência entre seções narrativas e os números contábeis.

Outro estudo que verificou o conflito entre as informações contidas nas seções narrativas e nas demonstrações financeiras foi realizado por Smith (1998). O autor submeteu 60 estudantes graduados em Contabilidade, concluindo o MBA a um experimento. Nele, os alunos tinham que tomar decisões a partir de informações financeiras e, posteriormente, também de narrativas de empresas fracassadas e empresas de sucesso. Como resultados, o autor concluiu haver evidências de conflito significativo entre as mensagens transmitidas pelas informações contábeis e pelas narrativas. Para ele, a narrativa pode induzir a erro na tomada de decisão.

Clatworthy e Jones (2003) examinaram 100 cartas de presidentes de firmas listadas no Reino Unido, 50 com variação de desempenho positiva e 50 com variação negativa, para saber se firmas com aumento ou declínio de desempenho reportam boas e más notícias de maneiras diferentes. De maneira geral, os resultados sugerem que as firmas preferem enfatizar os aspectos positivos de seu desempenho, tomando para si o crédito pelas boas notícias enquanto culpam o ambiente externo pelas más notícias.

Tessarolo, Pagliarussi e da Luz (2010) utilizaram a análise de conteúdo para investigar se existiria um padrão de atribuição causal nas seções narrativas dos relatórios anuais e de que forma o contexto econômico e o desempenho da empresa influenciaria tal padrão. Os resultados encontrados pelos autores sugerem que, no Brasil, os relatórios anuais são marcados por níveis significativos de atribuições em causa própria, ou seja, tentam criar uma reputação corporativa positiva, mesmo obtendo um desempenho negativo num contexto externo favorável. Os autores afirmam ainda que ficou evidente a intenção de os gestores divulgarem as seções narrativas do relatório anual de forma oportunista. Em estudo com temática semelhante, da Luz *et al.* (2009) encontraram que as empresas listadas na BM&FBovespa utilizam de táticas de gerenciamento de impressões em seus relatórios anuais, independentemente da adesão a níveis diferenciados de governança corporativa. Já Ribeiro *et al.* (2011) observaram, em estudo de caso sobre uma empresa de papel e celulose, que as informações textuais apresentadas procuravam priorizar e enfatizar a identidade institucional da empresa, ao intensificar, circunstanciar e reforçar suas atitudes, em vez de informar de maneira neutra suas realizações ou planos.

Finalmente, Vargas, Almeida e Maria Júnior (2014) observaram a existência de gerenciamento de impressão nos relatórios de administração de empresas listadas na BM&FBovespa. Os autores constataram que há diferença na quantidade de linhas e palavras presentes nas seções que discutem lucro e prejuízo. Também observaram que, ao tratar de resultados negativos (prejuízo), reduz-se a frequência de termos que expressam esse resultado e aumenta o destaque de termos como Ebitda e margem bruta. Também foram observados indícios de tentativa de desvio da atenção dos usuários dos resultados para outros indicadores.

2.3 Desenvolvimento das Hipóteses

Os resultados das pesquisas utilizando seções narrativas ainda são contraditórios, não havendo consenso a respeito da consistência dos textos com o desempenho ou se, de fato, há uma apresentação enviesada das informações visando ao gerenciamento da reputação corporativa.

Na visão de Smith (1998), predomina o gerenciamento de reputação nos documentos divulgados pelas empresas. Com base nas teorias da Agência e Sinalização, o autor afirma que as empresas buscam construir uma narrativa de modo a ofuscar a percepção da audiência quando o desempenho financeiro é ruim. Abrahamson e Amir (1996), por sua vez, encontraram evidências de coerência entre os dois conjuntos de informação, números e narrativa.

Conforme teoriza Spence (1973), em um ambiente de mercado existem características que são fixas e outras modificáveis. O autor argumenta que, em um contexto de informação assimétrica, essas características modificáveis são transmitidas por meio de sinais, capazes de alterar crenças e as impressões de indivíduos que as recebem. Ou seja, se existe assimetria de informação, indivíduos e/ou empresas podem fornecer aos interessados a informação necessária a fim de eliminar e/ou minimizar a assimetria. Mas as informações devem ser críveis, sendo que o interesse de quem as fornece é potencial fonte de viés.

No que Milgrom (1981) chama de jogo de persuasão, quem fornece informações a um indivíduo na tentativa de influenciar suas decisões deveria revelar todas as informações que possui sobre o “produto”, pois ele presume que as informações podem ser verificadas e o “comprador” pode detectar se há informações que foram ocultadas. Porém, se existirem informações que não podem ser verificadas, o “vendedor” poderá passá-las da maneira que lhe for mais conveniente. Corroboram esta ideia Grossman e Hart (1980) ao dizerem que não é vantagem para o “vendedor” reter informações na tentativa de enganar. Os autores sugerem que, assumindo que fraudar ou mentir é ilegal, e o custo de transmissão das informações é irrelevante, então o “vendedor” vai divulgar, voluntariamente, todas as informações que ele tem acerca do “produto”.

No contexto do presente estudo, assume-se que se as informações forem passíveis de verificação, elas serão transmitidas de forma transparente. Caso contrário, poderá haver tentativas de gerenciamento de reputação por meio da escolha dos conjuntos de informação a serem transmitidos a respeito da situação da empresa.

Portanto, para atender ao objetivo da pesquisa, que foi verificar se há harmonia nas mensagens transmitidas por meio do texto, representado pelas seções narrativas dos relatórios anuais, comparativamente ao desempenho da empresa, representado pela variação no seu lucro líquido, assumimos as seguintes hipóteses:

H1a: *Quanto melhor o desempenho da empresa, maior a frequência de informações quantitativas positivas a respeito dos temas Rentabilidade, Crescimento e Gestão;*

H1b: *Quanto pior o desempenho da empresa, maior a frequência de informações quantitativas negativas a respeito dos temas Rentabilidade, Crescimento e Gestão;*

Espera-se que o desempenho da empresa tenha uma relação diretamente proporcional com a frequência de informações quantitativas no seu relatório. Assim, entende-se que, quanto às informações quantitativas, o texto dos relatórios será transparente e coerente com os números.

H2a: *Independente do desempenho da empresa, espera-se que não haja diferença na frequência de informações qualitativas negativas e positivas sobre os temas Rentabilidade, Crescimento e Gestão.*

Pelo fato de as informações qualitativas serem mais subjetivas e não estarem diretamente associadas aos números, espera-se que haja falta de harmonia entre as informações qualitativas e quantitativas, no caso das empresas com variação negativa no lucro, pois espera-se que estas, mesmo não tendo um bom desempenho, apresentem informações qualitativas positivas e negativas com frequência semelhante às empresas com variação positiva no lucro.

3. Metodologia

3.1 Dados e Amostra

A construção da amostra utilizada contou inicialmente com todas as empresas listadas na BM&FBovespa e que tiveram seus relatórios anuais publicados referentes ao ano de 2009. O ano de 2009 foi selecionado por duas razões: primeiramente, o desempenho das empresas não estava ainda seriamente impactado pela crise mundial iniciada ao final de 2008, o que provavelmente afetaria a narrativa construída para justificar o desempenho das organizações. Também, a partir de 2010, em decorrência da adoção das normas internacionais de contabilidade no Brasil, as empresas listadas na BM&FBovespa passaram a divulgar suas informações narrativas por meio do Formulário de Referência, com estrutura bastante diversa das informações anuais até então exigidas. Para efeitos de comparabilidade com os estudos anteriores, o ano de 2009 se mostrou o mais apropriado para a condução do estudo.

A coleta dos relatórios foi feita diretamente no *site* da BM&FBovespa e contou com uma amostra inicial de 389 relatórios. Conforme Clatworthy e Jones (2001), relatórios anuais que contenham um total menor que 300 palavras de conteúdo são pouco representativos, e, portanto, devem ser excluídos da amostra considerada. Baseando-se nesta premissa, foram excluídos 25 relatórios que apresentaram a condição, restando então 364 relatórios.

Na etapa seguinte, foram extraídos da base de dados Economatica os dados referentes à variação do resultado contábil líquido (lucro/prejuízo) das empresas, para determinar aquelas com maior variação positiva e negativa. Os relatórios que não possuíam as informações necessárias para o cálculo da variação do resultado contábil líquido (resultado líquido para os anos de 2008 e 2009), que totalizaram cinco, foram excluídos, sobrando finalmente 359 relatórios passíveis de análise.

Foram então selecionados 120 relatórios, sendo os 60 com maior variação positiva do resultado contábil líquido e os 60 com maior variação negativa do resultado contábil líquido. Foram escolhidos estes dois grupos extremos para captar as empresas que estavam passando naquele momento por mudanças mais radicais no seu desempenho financeiro (Clatworthy & Jones, 2003, 2006).

3.2 Análise de Conteúdo

A técnica de análise de conteúdo foi utilizada para extrair informações das seções narrativas dos relatórios anuais. A análise de conteúdo tem por objetivo utilizar o conteúdo de mensagens para evidenciar indicadores que infiram sobre outra realidade que não a da mensagem apenas (Bardin, 2004).

Para avaliar os tipos de mensagem transmitidas nas seções narrativas desses relatórios, foram determinados grupos de palavras-chave que remetem aos temas rentabilidade, crescimento e gestão das empresas, conforme Balata e Breton (2005). As palavras-chave utilizadas para rentabilidade foram: custo, despesa, faturamento, ganho, lucro, margem, perda, preço, prejuízo, receita, rentabilidade, resultado, retorno e valor. Para o crescimento, temos: adição, alienação, aquisição, crescimento, demanda, desenvolvimento, fusão, incorporação, investimento, mercado, oferta, vendas e volume. No caso de gestão, as palavras-chave são: competitividade, controle, diferenciação, estratégia, foco, gestão, iniciativa, inovação, liderança, produtividade, reestruturação e reorganização.

Para localizar as palavras-chave nos relatórios, foi utilizado o software ATLAS.ti para a análise de conteúdo, que, além de localizar no texto os trechos onde as palavras-chave aparecem, gera um relatório com a contagem das palavras, tanto por total de palavras-chave em cada relatório, quanto por total de cada uma das palavra-chave somadas em todos os relatórios. Neste processo, foram inseridos no *software* apenas os radicais das palavras, para que pudessem ser capturadas palavras derivadas relacionadas às palavras-chave, como, por exemplo, “rent”, que pode ser relativo à rentabilidade, rentável, rentáveis.

Feito isso, foi gerado um relatório com a frequência das palavras-chave, que totalizaram 17.348 aparições de alguma das palavras mencionadas. Outro relatório foi gerado a partir do *software* com os trechos dos relatórios anuais em que aparecia cada uma das palavras-chave, dando origem a um arquivo do editor de textos com 1.405 páginas.

Os autores avaliaram manualmente este arquivo para determinar em que contexto cada uma das palavras aparecia no decorrer do texto (positivo quantitativo, positivo qualitativo, negativo quantitativo ou negativo qualitativo) para cada um dos temas centrais (rentabilidade, crescimento e gestão). Palavras que não estavam inseridas em nenhum dos contextos mencionados, ou seja, aquelas que apareciam em títulos e subtítulos, as que tinham um sentido neutro no texto ou ainda aquelas com radical semelhante, mas que não correspondiam à palavra-chave (por exemplo, “rent”: **corrente**), foram descartadas, restando um total de 7.173 palavras válidas.

Para ilustrar como ocorreu esta avaliação, a seguir são apresentados alguns trechos onde aparecem as palavras-chave referentes aos temas propostos em cada um dos sentidos avaliados. No que se refere ao tema de Rentabilidade, tem-se como exemplo de Rentabilidade positiva quantitativa: “No ano de 2009, o faturamento da companhia foi de R\$ 299.195, 131% a mais do que o apresentado em 2008”. Neste trecho, a palavra-chave “faturamento” está relacionada com o trecho em destaque “131% a mais”, ou seja, uma notícia positiva e de forma quantitativa. Um exemplo de Rentabilidade positiva qualitativa: “As despesas financeiras líquidas obtiveram uma diminuição devido à melhora de capital de giro [...]”. Neste trecho, a palavra despesa está associada a “diminuição”, portanto aparece em um contexto positivo e de forma qualitativa.

Para exemplificar Rentabilidade negativa quantitativa, temos: “O lucro bruto de R\$ 85,1 milhões em 2008 caiu 9,5% para R\$ 77,1 milhões em 2009”. Neste trecho, a palavra-chave “lucro” aparece associada à palavra “caiu”, ou seja, em um contexto negativo e seguido por valores numéricos, que dão um sentido quantitativo ao trecho. Rentabilidade negativa qualitativa: “[...] e os altos custos de matérias-primas, como enxofre, amônia e derivados de petróleo, geraram uma considerável redução nas margens de rentabilidade da empresa em 2009”. No trecho citado, o termo Rentabilidade aparece associado à palavra “redução”, ou seja, um contexto negativo para empresa, em um sentido qualitativo.

Com relação ao tema Crescimento, temos um exemplo com contexto positivo quantitativo: “[...] e vendas que somaram 2,9 milhões de toneladas, volume recorde, que representa aumento de 16,7% também em comparação ao exercício anterior”. Neste trecho, as palavras-chave “vendas” e “volume” aparecem associadas à palavra “aumento”, ou seja, um sentido positivo e de forma quantitativa. Para Crescimento positivo qualitativo, um exemplo é o trecho: “[...] nossa presença em todas as regiões do país e exposição a diferentes segmentos de renda permitem continuarmos crescendo de uma maneira sólida e significativa nos próximos anos”. Nesta citação, o termo “crescendo” aparece em um sentido positivo qualitativo.

Já no contexto de Crescimento negativo quantitativo, temos: “No ano de 2009, as vendas no mercado interno totalizaram 3,2 milhões de toneladas, representando uma retração de 22% em relação ao mesmo período de 2008, fruto da desaceleração na demanda registrada, principalmente, no primeiro semestre do ano”. Neste trecho, o termo “vendas” vem associado à palavra “retração”, acompanhada de valores numéricos, ou seja, no sentido negativo quantitativo. No tema Crescimento, com sentido negativo qualitativo: “Os investimentos em 2009 foram substancialmente reduzidos [...]”. Neste trecho, a palavra-chave “investimentos” está associada à palavra “reduzidos”, o que demonstra uma notícia negativa por parte da empresa e de forma qualitativa.

Para o tema Gestão, um exemplo no contexto positivo quantitativo: “Mantivemos projetos Kaizen em diversos setores da produção, tornando os processos adequados e reduzindo tempos nas operações, com resultados de melhoria de produtividade entre 50% e 200%”. Neste trecho, o termo “produtividade” aparece associado à palavra “melhoria” e aos percentuais “50% e 200%”, dando um sentido positivo e quantitativo à frase. No tema Gestão, com sentido positivo qualitativo aparece o seguinte trecho: “[...] criação de valor para todo o público com o qual se relaciona, além de consolidar sua liderança nos mercados em que atua”.

Neste trecho, a palavra-chave “liderança” aparece associada a “consolidar”, ou seja, um contexto positivo para a empresa e de forma qualitativa.

Um exemplo de Gestão no contexto negativo quantitativo: “[...] queda de produtividade, principalmente da soja, de 10,5%, no ano-safra 2008/09 em comparação com o ano-safra anterior”. Na frase citada, o termo produtividade aparece ligado à palavra “queda” e aos valores de “10,5%”. Sendo assim, um contexto negativo quantitativo. Já para o tema Gestão, no contexto negativo qualitativo, temos: “[...] a conseqüente queda da atividade produtiva obrigaram a Usiminas a fazer ajustes no seu quadro de funcionários e colaboradores”. Neste caso, a palavra-chave “produtiva” está associada à palavra “queda”, o que dá um sentido negativo e qualitativo ao trecho.

3.3 Teste das Hipóteses Elaboradas

Tabuladas as frequências e os sentidos em que aparecem todas as palavras-chave, tanto por tema quanto por relatório, foram realizadas regressões logísticas, técnica apropriada quando a variável dependente é binária (não métrica) e as variáveis independentes são métricas (Hair, Babin, Anderson & Tatham, 2007). Assim, os dois grupos de interesse como uma variável binária são codificados com valores 0 e 1. Neste trabalho, consideramos empresas com variação negativa no lucro, como 0, e empresas com variação positiva no lucro, como 1.

Para realização das regressões, temos, de um lado, a variável binária referente à variação do lucro (0 para as empresas com variação negativa no lucro e 1 para as empresas com variação positiva no lucro) e, do outro lado, os coeficientes relativos às características das palavras-chave (positiva quantitativa e qualitativa, negativa quantitativa ou qualitativa) extraídas das seções narrativas. O modelo que representa a regressão vem a seguir:

$$y = \beta_0 + \beta_1 \text{Rentpq} + \beta_2 \text{Rentpql} + \beta_3 \text{Rentnegqt} + \beta_4 \text{Rentnegql} + \beta_5 \text{Crescpqt} + \beta_6 \text{Crescpql} + \beta_7 \text{Crescnegqt} + \beta_8 \text{Crescnegql} + \beta_9 \text{Gestpq} + \beta_{10} \text{Gestpql} + \beta_{11} \text{Gestnegqt} + \beta_{12} \text{Gestnegql} + \varepsilon$$

Onde:

- y:** 0 = empresas com variação negativa no lucro e; 1 = empresas com variação positiva no lucro;
- Rentpq:** Rentabilidade positiva quantitativa = soma das frequências de todas as palavras-chave associadas ao tema Rentabilidade que aparecem no texto do relatório com sentido positivo e quantitativo;
- Rentpnq:** Rentabilidade positiva qualitativa = soma das frequências de todas as palavras-chave associadas ao tema Rentabilidade que aparecem no texto do relatório com sentido positivo e não quantitativo;
- Rentnegq:** Rentabilidade negativa quantitativa = soma das frequências de todas as palavras-chave associadas ao tema Rentabilidade que aparecem no texto do relatório com sentido negativo e quantitativo;
- Rentnegnq:** Rentabilidade negativa qualitativa = soma das frequências de todas as palavras-chave associadas ao tema Rentabilidade que aparecem no texto do relatório com sentido negativo e não quantitativo;
- Crescpq:** Crescimento positivo quantitativo = soma das frequências de todas as palavras-chave associadas ao tema Crescimento que aparecem no texto do relatório com sentido positivo e quantitativo;
- Crescpnq:** Crescimento positivo qualitativo = soma das frequências de todas as palavras-chave associadas ao tema Crescimento que aparecem no texto do relatório com sentido positivo e não quantitativo;
- Crescnegq:** Crescimento negativo quantitativo = soma das frequências de todas as palavras-chave associadas ao tema Crescimento que aparecem no texto do relatório com sentido negativo e quantitativo;

- Crescnegq:** Crescimento negativo qualitativo = soma das frequências de todas as palavras-chave associadas ao tema Crescimento que aparecem no texto do relatório com sentido negativo e não quantitativo;
- Gestpnq:** Gestão positiva quantitativa = soma das frequências de todas as palavras-chave associadas ao tema Gestão que aparecem no texto do relatório com sentido positivo e quantitativo;
- Gestpnq:** Gestão positiva qualitativa = soma das frequências de todas as palavras-chave associadas ao tema Gestão que aparecem no texto do relatório com sentido positivo e qualitativo;
- Gestnegq:** Gestão negativa quantitativa = soma das frequências de todas as palavras-chave associadas ao tema Gestão que aparecem no texto do relatório com sentido negativo e quantitativo;
- Gestnegq:** Gestão negativa qualitativa = soma das frequências de todas as palavras-chave associadas ao tema Gestão que aparecem no texto do relatório com sentido negativo e não quantitativo.

Foram realizadas duas regressões logísticas. A primeira pela frequência bruta das palavras-chave, com os coeficientes sendo obtidos pela soma do número absoluto da ocorrência de cada palavra-chave nos relatórios.

A segunda regressão teve os seus coeficientes calculados a partir da ponderação de termos. A ponderação reflete a importância de um termo em um documento, baseando-se no número de ocorrências desse termo (Manning, Raghavan & Schütze, 2008; Loughran & McDonald, 2011). Para isso, atribui-se a cada termo do documento (relatório) um peso de acordo com o seu número de aparições no texto. A forma mais simples seria atribuir como peso o próprio número de ocorrências do termo t no documento d , pois uma maior frequência sugere maior importância da palavra no texto (Manning, Raghavan & Schütze, 2008).

Porém, Manning e Schütze (1999) afirmam que essa importância precisa ser atenuada. Um exemplo dado pelos autores é de que um termo que aparece três vezes não significa que ele seja três vezes mais importante que outro que apareça apenas uma vez. Outro detalhe importante e de utilização comum para definir a importância de um termo é saber em quantos relatórios ele aparece, em relação ao número total de relatórios. Definidos esses critérios, temos a seguir a fórmula de ponderação de termos desenvolvida por Manning e Schütze (1999).

$$W(i, j) = \begin{cases} (1 + \ln(tf_{i,j})) \ln \frac{N}{df_i}, & \text{se } tf_{i,j} \geq 1 \\ 0 & , \text{ se } tf_{i,j} = 0 \end{cases}$$

Onde:

$W_{(i,j)}$: peso do termo i , no relatório j ;

$tf_{i,j}$: frequência da palavra-chave i , no relatório j ;

N : número total de relatórios;

df_i : número de relatórios que contém a palavra-chave (i).

4. Análise dos Dados

A seguir, apresenta-se a análise dos dados da pesquisa, tendo as estatísticas descritivas e as regressões logísticas com as respectivas classificações estatísticas.

A Tabela 1 mostra as estatísticas descritivas referentes à soma da frequência bruta das palavras-chave.

Tabela 1

Estatística descritiva por frequência bruta

Variáveis	Total			Variação negativa no lucro			Variação positiva no lucro		
	Obs.	Méd.	D.Padrão	Obs.	Méd.	D.Padrão	Obs.	Méd.	D.Padrão
Rentpqt	120	5,167	5,402	60	3,917	5,099	60	6,417	5,447
Rentpql	120	9,292	9,467	60	7,717	8,567	60	10,867	10,115
Rentnegqt	120	3,392	4,235	60	4,367	4,723	60	2,417	3,456
Rentnegql	120	2,742	3,781	60	3,517	4,048	60	1,967	3,349
Crescpqt	120	6,675	8,385	60	5,733	8,273	60	7,617	8,459
Crescpql	120	18,092	19,527	60	15,567	19,727	60	20,617	19,155
Crescnegqt	120	1,375	2,521	60	1,483	2,734	60	1,267	2,306
Crescnegql	120	2,083	3,034	60	2,45	2,861	60	1,717	3,179
Gestpqt	120	0,025	0,157	60	0,033	0,181	60	0,167	0,129
Gestpql	120	10,775	15,749	60	9,350	15,491	60	12,200	16,005
Gestnegqt	120	0,075	0,735	60	0,150	1,039	60	0	0
Gestnegql	120	0,833	0,306	60	0,117	0,372	60	0,050	0,220

Legenda: Rentpqt (Rentabilidade positiva quantitativa); Rentpql (Rentabilidade positiva qualitativa); Rentnegqt (Rentabilidade negativa quantitativa); Rentnegql (Rentabilidade negativa qualitativa); Crescpqt (Crescimento positivo quantitativo); Crescpql (Crescimento positivo qualitativo); Crescnegqt (Crescimento negativo quantitativo); Crescnegql (Crescimento negativo qualitativo); Gestpqt (Gestão positiva quantitativa); Gestpql (Gestão positiva qualitativa); Gestnegqt (Gestão negativa quantitativa); Gestnegql (Gestão negativa qualitativa).

Fonte: Elaborada pelos autores

Analisando a Tabela 1, observa-se que a frequência de informações não quantitativas parece ser maior que a quantidade de informações quantitativas nos relatórios. Por exemplo, no tema Crescimento, temos média 18,10 para informações não quantitativas e média de 6,68 para informações quantitativas. Em relação às informações positivas ou negativas, as positivas parecem ser mais frequentes. Por exemplo, no tema Rentabilidade as informações positivas têm média de 9,29 e as informações negativas sobre o mesmo tema têm média de 2,74.

Percebeu-se ainda que os relatórios, de um modo geral, fornecem mais informações positivas acerca das empresas. No presente estudo, de todas as menções de alguma palavra-chave, 83,7% foram de forma positiva. Se considerarmos apenas as empresas com variação negativa no lucro, ainda, assim o percentual de palavras-chave com sentido positivo fica em 77,8%.

Após serem calculadas as estatísticas descritivas, a Tabela 2 mostra os resultados da regressão logística realizada pela frequência bruta das palavras-chave nos relatórios.

Tabela 2

Regressão logística por frequência bruta

Variável	Coef.	Erro Padrão	Z	P>z
Rentpqt	0,1872	0,0777	2,41	0,016**
Rentpql	0,1642	0,0695	2,36	0,018**
Rentnegqt	-0,2697	0,1102	-2,45	0,014**
Rentnegql	-0,2721	0,1216	-2,24	0,025**
Crescpqt	-0,0279	0,0507	-0,55	0,581
Crescpql	0,0315	0,0341	0,93	0,355
Crescnegqt	0,1414	0,1547	0,91	0,361
Crescnegql	0,0309	0,1439	0,21	0,830
Gestpqt	-1,9688	1,7561	-1,12	0,262
Gestpql	-0,0824	0,0404	-2,04	0,041**
Gestnegql	-1,9356	1,1630	-1,66	0,096*
Nº Obs = 120	LR chi2(11) = 42,74	P>chi2 = 0,0000	Pseudo R2 = 0,2569	

Legenda: Rentpqt (Rentabilidade positiva quantitativa); Rentpql (Rentabilidade positiva qualitativa); Rentnegqt (Rentabilidade negativa quantitativa); Rentnegql (Rentabilidade negativa qualitativa); Crescpqt (Crescimento positivo quantitativo); Crescpql (Crescimento positivo qualitativo); Crescnegqt (Crescimento negativo quantitativo); Crescnegql (Crescimento negativo qualitativo); Gestpqt (Gestão positiva quantitativa); Gestpql (Gestão positiva qualitativa); Gestnegql (Gestão negativa qualitativa).

Níveis de Significância: *** = 1%; ** = 5%; * = 10%

Fonte: Elaborada pelos autores

Na realização da regressão, o *software* descartou a variável Gestnegqt (Gestão negativa quantitativa) por falta de variabilidade dos dados e por isso a regressão foi feita novamente sem esta variável. Os resultados mostrados na Tabela 2, já sem esta variável descartada, apontam que o modelo foi estatisticamente significativo, pois $P > \chi^2 = 0,0000$. Adicionalmente, foram significantes ao nível de 5% as variáveis Rentpqt, Rentpql, Rentnegqt, Rentnegql e Gestpql. Também foi significativa a variável Gestnegql, porém ao nível de 10%.

Analisando os resultados, a hipótese 1a, que previa que maiores frequências de informações quantitativas positivas a respeito dos temas Rentabilidade, Crescimento e Gestão, aumentariam a probabilidade de uma empresa pertencer ao grupo de empresas com variação positiva no lucro. Somente a variável Rentpqt teve o seu coeficiente estatisticamente significativo, com o sinal conforme esperado, constituindo indícios de confirmação da hipótese. Já as variáveis Crescpqt e Gestpqt não tiveram seus coeficientes estatisticamente significantes, rejeitando a hipótese.

A hipótese 1b – quanto maior a frequência de palavras-chave quantitativas negativas a respeito dos temas Rentabilidade, Crescimento e Gestão, maior será a probabilidade de as observações pertencerem ao grupo 0 (empresas com variação negativa no lucro) – apresentou indícios de confirmação da hipótese para a variável Rentnegqt, que teve o seu coeficiente estatisticamente significativo e com o sinal esperado. As demais variáveis (Crescnegqt e Gestnegqt) não tiveram os coeficientes de suas variáveis estatisticamente significantes e por isso apresentam indícios de que as hipóteses foram rejeitadas.

Na hipótese 2a, esperava-se que as frequências fossem semelhantes. Algumas variáveis tiveram seus coeficientes sugerindo indícios de confirmação da hipótese e outras com indicativos de que a hipótese foi rejeitada. Com esta hipótese, a analogia é inversa, ou seja, as variáveis que tiveram seus coeficientes significantes (Rentpql, Gestpql, Rentnegql e Gestnegql) geram indícios de que a hipótese foi rejeitada por indicarem que as frequências não são semelhantes, conforme sugere a hipótese.

Para a variável Rentpql, maiores frequências de palavras-chave implicam em maiores probabilidades de uma observação pertencer ao grupo de empresas com variação positiva no lucro, ou seja, empresas com bom desempenho apresentam mais informações nos seus textos de forma positiva qualitativa acerca do tema Rentabilidade, o que demonstra coerência entre o texto e os números.

Para as demais variáveis (Gestpql, Rentnegql e Gestnegql), maiores frequências aumentam as chances de uma determinada observação fazer parte das empresas não com variação positiva no lucro. Nas variáveis Rentnegql e Gestnegql não há conflito de informações, pois empresas em piores situações financeiras tendem a ter mais informações negativas. Já com a variável Gestpql, existe conflito nas mensagens transmitidas, sendo que maiores frequências de informações sobre Gestão, com sentido positivo e de forma qualitativa, aumentam as chances de serem empresas com variação negativa no lucro.

Já para as variáveis Crescpql e Crescnegql, os coeficientes confirmaram a hipótese, pois não foram estatisticamente significantes e sugerem frequências semelhantes para empresas com desempenho positivo ou negativo. Ou seja, não é possível por estes resultados identificar pela frequência de palavras-chave à qual grupo de empresas pertence uma determinada observação. Neste caso há conflito, pois se esperava que empresas com variação positiva no lucro apresentassem mais informações positivas e as empresas com desempenho ruim prestassem mais informações com sentido negativo, devido a sua situação financeira.

Adicionalmente à regressão logística na qual as variáveis foram construídas por meio da frequência bruta das palavras-chave, foi realizada outra regressão, com a ponderação dos termos, atribuindo diferentes pesos, de acordo com o número de suas ocorrências dos mesmos, demonstrada na Tabela 3.

Tabela 3

Regressão logística por ponderação dos termos

Variável	Nº Obs = 120	LR chi2(11) = 48,18	P>chi2 = 0,0000	Pseudo R2 = 0,2896
	Coef.	Erro Padrão	z	P>z
Rentpqt	0,0926	0,0788	1,18	0,240
Rentpql	0,2768	0,0974	2,84	0,004***
Rentnegqt	-0,3654	0,1127	-3,24	0,001***
Rentnegql	-0,2316	0,0880	-2,63	0,008***
Crescpqt	-0,0036	0,0759	-0,05	0,962
Crescpql	0,2045	0,0904	2,26	0,024**
Crescnegqt	0,2928	0,1369	2,14	0,032**
Crescnegql	-0,0663	0,1038	-0,64	0,523
Gestpqt	0,2306	0,4287	0,54	0,591
Gestpql	-0,1557	0,0707	-2,20	0,028**
Gestnegql	-0,3655	0,3061	-1,19	0,232

Legenda: Rentpqt (Rentabilidade positiva quantitativa); Rentpql (Rentabilidade positiva qualitativa); Rentnegqt (Rentabilidade negativa quantitativa); Rentnegql (Rentabilidade negativa qualitativa); Crescpqt (Crescimento positivo quantitativo); Crescpql (Crescimento positivo qualitativo); Crescnegqt (Crescimento negativo quantitativo); Crescnegql (Crescimento negativo qualitativo); Gestpqt (Gestão positiva quantitativa); Gestpql (Gestão positiva qualitativa); Gestnegql (Gestão negativa qualitativa).

Níveis de Significância: *** = 1%; ** = 5%; * = 10%

Fonte: Elaborada pelos autores

Assim como na regressão por frequência bruta, a variável Gestnegqt (Gestão negativa quantitativa) foi desconsiderada pelo *software* na regressão por ponderação de termos por falta de variabilidade dos dados, e a regressão foi feita novamente sem esta variável. Os resultados obtidos apontam que o modelo foi estatisticamente significativo, pois $P > \chi^2 = 0,0000$. Adicionalmente, foram significantes ao nível de 1% as variáveis Rentpql, Rentnegqt e Rentnegql. Também foram significantes, porém ao nível de 5%, as variáveis Crescpql, Crescnegqt e Gestpql.

Analisando os resultados, a hipótese 1a, que diz que quanto maior a frequência de palavras-chave positivas quantitativas relacionadas à Rentabilidade, Crescimento e Gestão maiores as chances de a observação pertencer ao grupo de empresas com variação positiva no lucro, nenhuma das variáveis teve o coeficiente estatisticamente significativo, apresentando indícios de que a hipótese foi rejeitada.

A hipótese 1b – quanto maior a frequência de palavras-chave relativas aos temas Rentabilidade, Crescimento e Gestão, de forma negativa quantitativa, maior é a probabilidade de uma observação pertencer ao grupo de empresas com variação negativa no lucro – teve o coeficiente estatisticamente significativo para as variáveis *Rentnegqt* e *Crescnegqt*. Para a variável *Rentnegqt*, o sinal do coeficiente foi conforme o esperado, confirmando a hipótese. Já para a variável *Crescnegqt*, o sinal do coeficiente foi contrário ao esperado, o que implica indícios de que a hipótese foi rejeitada. Este resultado sugere que quanto maior a frequência de palavras-chave relativas ao tema Crescimento, de forma negativa quantitativa, maior é a probabilidade de uma observação pertencer ao grupo de empresas com desempenho ruim.

A variável *Gestnegqt* não teve o coeficiente estatisticamente significativo e por isso há indícios de que a hipótese foi rejeitada para esta variável. Assim, não se pode afirmar que maiores frequências de palavras-chave neste sentido aumentarão as chances de as observações fazerem parte do grupo de empresas com variação negativa no lucro.

A hipótese 2a, que sugeria que as médias de frequências são semelhantes para informações qualitativas teve os coeficientes das variáveis *Rentpql*, *Crespql*, *Gestpql* e *Rentnegql* estatisticamente significantes, constituindo indícios de que a hipótese foi rejeitada. Para as variáveis *Rentpql* e *Crespql*, quanto maiores as frequências de palavras-chave, maiores serão as probabilidades de uma observação pertencer ao grupo de empresas com variação positiva no lucro, ou seja, empresas com bom desempenho fornecem mais informações textuais de forma positiva qualitativa sobre os temas Rentabilidade e Crescimento, o que demonstra coerência entre seções narrativas e desempenho financeiro.

Com a variável *Gestpnq*, assim como na regressão anterior, existe conflito nas mensagens transmitidas, pois quanto maior a frequência de informações sobre Gestão, com sentido positivo e de forma não quantitativa, maior será a probabilidade de a observação fazer parte das empresas com variação negativa no lucro. Analisando a variável *Rentnegnq*, maiores frequências indicam maior probabilidade de uma determinada observação fazer parte das empresas com variação negativa no lucro, ou seja, empresas com pior desempenho fornecem mais informações acerca do tema Rentabilidade, com sentido negativo, de forma qualitativa. Isso demonstra coerência entre informações textuais e financeiras.

Já nas variáveis *Crescnegq* e *Gestnegq*, os coeficientes apresentam indícios de confirmação da hipótese, indicando médias semelhantes entre empresas com bom desempenho e com desempenho ruim. Assim, não é possível por estes resultados identificar pela frequência de palavras-chave a qual grupo de empresas pertence uma determinada observação. Neste caso existe conflito entre as informações, pois se esperava que as empresas com variação negativa no lucro, devido ao seu desempenho inferior, apresentassem mais informações com sentido negativo, sobre os temas Crescimento e Gestão.

Comparando-se os resultados das duas regressões realizadas, verificou-se que, em se tratando da classificação estatística, os dois modelos são bem parecidos, não apresentando grandes diferenças. O modelo por frequência bruta teve 79,17% de classificações corretas, enquanto que o modelo por ponderação de termos teve 80 % de classificações corretas.

O Pseudo R2 também aumentou com a ponderação dos termos, passando de 0,2569 na regressão por frequência para 0,2896, demonstrando um maior poder explicativo deste modelo.

Quanto às variáveis estatisticamente significantes o modelo por frequência bruta teve cinco variáveis significantes ao nível de 5%, e uma variável significativa ao nível de 10%, em um total de seis. Já o modelo por ponderação de termos teve seis variáveis significantes, sendo três ao nível de 1% e três ao nível de 5%.

Porém, se forem observadas somente as variáveis que se esperava serem significantes de acordo com as hipóteses, o modelo por frequência bruta teve duas ao nível de 5%, e o modelo por ponderação de termos teve uma variável significativa ao nível de 1% e outra ao nível de 5%, sendo esta com sinal do coeficiente contrário ao esperado.

Sendo assim, em termos gerais, podemos dizer, que apesar de modelo por ponderação de termos ter obtido melhores resultados na classificação estatística e no pseudo R2, no que tange às variáveis estatisticamente significantes esperadas, os dois modelos foram semelhantes. Isso mostrou que a ponderação

dos termos sugerida por Manning e Schütze (1999), que atenua a importância de um determinado termo que aparece com maior frequência no relatório, não trouxe melhoras significativas se comparado à simples contagem absoluta de palavras-chave.

A Tabela 4 traz um resumo das hipóteses, mostrando quais variáveis foram estatisticamente significantes e quais confirmaram ou rejeitaram as hipóteses em cada uma das regressões realizadas.

Tabela 4

Resumo dos resultados

Hipóteses	Variáveis	Freq. Bruta	Resultado	Freq. Ponderada	Resultado
1a	Rentpqt	**	confirmada	-	rejeitada
	Crescpqt	-	rejeitada	-	rejeitada
	Gestpqt	-	rejeitada	-	rejeitada
1b	Rentnegqt	**	confirmada	***	confirmada
	Cresnegqt	-	rejeitada	(**)	rejeitada
	Gestnegqt	-	rejeitada	-	rejeitada
2a	Rentpql	**	rejeitada	***	rejeitada
	Crescpql	-	confirmada	**	rejeitada
	Gestpql	**	rejeitada	**	rejeitada
	Rentnegql	**	rejeitada	***	rejeitada
	Crescnegql	-	confirmada	-	confirmada
	Gestnegql	*	rejeitada	-	confirmada

Legenda: Rentpqt (Rentabilidade positiva quantitativa); Rentpql (Rentabilidade positiva qualitativa); Rentnegqt (Rentabilidade negativa quantitativa); Rentnegql (Rentabilidade negativa qualitativa); Crescpqt (Crescimento positivo quantitativo); Crescpql (Crescimento positivo qualitativo); Crescnegqt (Crescimento negativo quantitativo); Crescnegql (Crescimento negativo qualitativo); Gestpqt (Gestão positiva quantitativa); Gestpql (Gestão positiva qualitativa); Gestnegqt (Gestão negativa quantitativa); Gestnegql (Gestão negativa qualitativa).

Níveis de Significância: *** = 1%; ** = 5%; * = 10%

Nota: (**) variável com o sinal do coeficiente contrário ao esperado.

Fonte: Elaborada pelos autores

Analisando os resultados quanto às hipóteses propostas, temos que: A hipótese 1^a, que previa que quanto melhor o desempenho da empresa, maior a frequência de informações quantitativas positivas a respeito dos temas Rentabilidade, Crescimento e Gestão, apresentou indicativos de confirmação da hipótese somente da variável Rentpqt e na regressão por frequência bruta. Nas demais variáveis e na regressão por ponderação de termos as variáveis não tiveram os seus coeficientes estatisticamente significantes e, por isso, a hipótese foi rejeitada.

Esses resultados sugerem que a mensagem transmitida foi consistente apenas no que tange ao tema Rentabilidade e, para informações referentes a este tema, não houve tentativa de gerenciamento de impressão pelas empresas. Este resultado pode ser justificado pelo fato de que quando se trata de informações que nos remetem à Rentabilidade, as palavras-chave são, em sua maioria, diretamente associadas aos números, ficando difícil passar uma falsa imagem da situação da empresa, como fica destacado no trecho a seguir: “Em 2009, houve um aumento de 78% do lucro bruto em relação a 2008, atingindo a marca de R\$ 70.388 em 2009, contra os R\$ 39.592 apresentados em 2008”. (trecho retirado do relatório de uma empresa com variação positiva no lucro).

Para os temas Crescimento e Gestão, os resultados sugerem frequências semelhantes de informação para empresas, independente do desempenho, indicando que as empresas com variação negativa no lucro transmitem informações positivas acerca destes temas tanto quanto àquelas com variação positiva no lucro.

A hipótese 1b, que previa que quanto pior o desempenho da empresa maior a frequência de informações quantitativas negativas a respeito dos temas Rentabilidade, Crescimento e Gestão, somente a variável *Rentnegqt* teve o coeficiente estatisticamente significativo nas duas regressões, o que implica dizer que houve indicativos de que a hipótese foi confirmada, sugerindo haver coerência entre o texto e o desempenho quando se trata desta variável.

A variável *Crescnegqt* teve o seu coeficiente estatisticamente significativo na regressão por ponderação, porém com o sinal inverso ao esperado, ou seja, empresas com maior frequência deste tipo de informações têm maiores probabilidades de pertencerem ao grupo de empresas com variação positiva no lucro. Isso sugere que as empresas com variação negativa no lucro evitam mencionar este tema pra não dar ênfase a sua situação financeira.

Já a variável *Gestnegqt* apresentou indicativos de que a hipótese foi rejeitada, já que esta variável foi descartada das regressões por falta de variabilidade nos dados neste sentido. Este resultado indica conflito nas mensagens transmitidas, pois sugere indícios de que as frequências deste tipo de informação é semelhante nas empresas, independente do desempenho, diferente do que se esperava na hipótese proposta.

Com estes resultados percebe-se que, assim como na hipótese 1a, somente as informações relativas ao tema Rentabilidade apresentam coerência entre seções narrativas e números.

A hipótese 2a, que previa que independente do desempenho da empresa não haveria diferença na frequência de informações qualitativas negativas e positivas sobre os temas Rentabilidade, Crescimento e Gestão, apresentou indícios de rejeição da hipótese para algumas variáveis e confirmação para outras. A analogia nesta hipótese é inversa, sendo a hipótese rejeitada quando os coeficientes das variáveis são estatisticamente significantes por indicar que as frequências não são semelhantes.

Na variável *Rentpql*, o coeficiente foi estatisticamente significativo nas duas regressões e sempre com o sinal positivo, indicando maior frequência para as empresas com variação positiva no lucro, o que indica coerência entre o texto e os números e indícios de que a hipótese foi rejeitada. A variável *Crescpql* só teve o seu coeficiente estatisticamente significativo na regressão por ponderação, com o sinal positivo, assim como a variável anterior.

A variável *Gestpql* teve o seu coeficiente estatisticamente significativo nas duas regressões. Porém, chama a atenção o fato de o coeficiente da variável ser negativo nas duas regressões. Isso indica que as frequências médias de palavras-chave sobre o tema gestão de forma positiva qualitativa foi sempre maior em empresas com variação negativa no lucro, indicando incoerência entre seções narrativas e desempenho financeiro e indícios de que a hipótese foi rejeitada. A variável *Rentnegql* também teve o seu coeficiente estatisticamente significativo nas duas regressões e ambas com o sinal do coeficiente negativo, indicando maior frequência para as empresas com variação negativa no lucro, constituindo indícios de que a hipótese 2a foi rejeitada. Com este resultado, o texto é coerente ao desempenho financeiro das empresas.

A variável *Crescnegql* não teve o seu coeficiente estatisticamente significativo em nenhuma regressão, confirmando a hipótese de semelhança entre as narrativas das empresas, independente do desempenho, e assim demonstrando conflito entre as informações fornecidas, já que se esperava que empresas com variação negativa no lucro tivessem maior frequência de informações com sentido negativo. A variável *Gestnegql* teve o seu coeficiente estatisticamente significativo na regressão por frequência bruta, com o sinal negativo, sugerindo maior frequência de informações nas empresas com variação negativa no lucro e indícios de que a hipótese foi rejeitada, indicando coerência entre as demonstrações.

Analisando esta hipótese, constatamos que, mesmo quando se trata de informações qualitativas, as informações relacionadas principalmente à rentabilidade e parte das referentes a crescimento são coerentes ao desempenho da empresa, ao passo que informações relacionadas à gestão, temos que quanto mais informações fornecidas, principalmente no sentido positivo, maior as chances de ser uma empresa com variação negativa no lucro.

Este resultado pode ser justificado pelo fato de, diferentemente das palavras-chave relativas à Rentabilidade e a algumas palavras relacionadas ao Crescimento, as palavras que nos remetem à Gestão não estão diretamente ligadas aos “números”, fazendo com que seja possível transmitir uma mensagem sobre

a situação da empresa que não seja realmente a verdadeira, como pode ser observado em destaque nos trechos a seguir: “A oferta de serviços em âmbito nacional confere à companhia a escala necessária para competir em igualdade de condições com os importantes competidores internacionais presentes no Brasil” (trecho retirado do relatório de uma empresa com variação positiva no lucro).

Em um trecho de outra empresa, temos: “As adequações realizadas e todas as ações tomadas em 2009 serviram para tornar a empresa mais produtiva e principalmente mais competitiva, o que nos qualifica para buscar novos clientes, mercados com maior valor agregado e desenvolvimento de produtos de alta complexidade” (trecho retirado do relatório de uma empresa com variação negativa no lucro).

Nestes trechos citados, percebe-se que independente da real situação da empresa, quando se trata do tema gestão, o texto pode ser igualmente positivo e otimista para as empresas, não importando qual o desempenho, como afirmaram Hildebrandt e Snyder (1981) sugerindo que, independente do desempenho financeiro da empresa, a linguagem no relatório anual será predominantemente positiva.

Constatou-se também que algumas empresas ignoram a recomendação feita pela CVM de evitar adjetivos e frases tais como “excelente resultado”, “ótimo desempenho”, “excelentes perspectivas”, quando estes não são corroborados por dados comparativos ou fatos. Alguns exemplos do descumprimento desta recomendação podem ser vistos nos trechos: “[...] que define as condições do serviço de energia nos sistemas isolados e traz excelentes perspectivas para a recuperação da saúde financeira das empresas” (trecho retirado do relatório de uma empresa com variação negativa no lucro).

5. Conclusão

O presente estudo teve como objetivo analisar a harmonia nas mensagens divulgadas nas seções narrativas dos relatórios anuais de empresas listadas na BM&FBovespa, comparativamente ao desempenho financeiro apresentado por estas empresas. Para alcançar o objetivo, foram selecionadas 120 empresas listadas na BM&FBovespa no ano de 2009, sendo as 60 com maior variação positiva e as 60 com maior variação negativa no resultado contábil líquido. Aplicou-se a análise de conteúdo nos relatórios destas empresas, analisando a frequência, natureza e polaridade de um conjunto de palavras-chave predeterminadas referentes aos temas de Rentabilidade, Crescimento e Gestão.

A análise de regressão logística foi aplicada para avaliar a associação entre variáveis obtidas a partir da análise de conteúdo com a medida de desempenho selecionada. De modo geral, os resultados indicam que, quando se trata de informações relativas à Rentabilidade, há harmonia entre texto e números. Já para as informações ligadas ao Crescimento, a harmonia entre texto e números diminui. E no caso das informações sobre o tema Gestão, a desarmonia fica mais evidente. Conclui-se, portanto, que não há uma total consistência entre texto e números, ou seja, existe um conflito nas informações, porém de forma parcial.

As palavras-chave ligadas à gestão, em sua maioria, além de serem mais subjetivas, dão uma ideia de perspectiva futura, ou seja, mesmo que o desempenho de uma determinada empresa tenha sido ruim, o seu relatório tenta trazer o foco da discussão para as perspectivas futuras. Percebeu-se que, quanto mais subjetiva a natureza da informação, maior a desarmonia com os números da empresa. Além disso, os relatórios de um modo geral transmitem predominantemente informações positivas acerca das empresas, minimizando alguma situação negativa com argumentos atenuantes.

Os resultados permitem concluir que há um grau moderado de desarmonia entre as informações textuais e numéricas divulgadas pelas empresas da amostra, e que a desarmonia observada em alguns está associada à tentativa de gerenciar a reputação corporativa no que diz respeito ao desempenho e, principalmente, no que tange à gestão da empresa, com tentativas de atenuar o foco no desempenho anterior negativo.

O presente estudo contribuiu para o entendimento dos aspectos que compõem a qualidade das informações financeiras e não financeiras apresentadas nas seções narrativas dos relatórios anuais das empresas brasileiras listadas. Os resultados obtidos reforçam a percepção alcançada por estudos anteriores de que a discricionariedade na apresentação das informações textuais nas seções narrativas pode ser ex-

cessiva. Ou seja, existe um elevado potencial de viés na apresentação de informações qualitativas sobre o desempenho e a gestão da empresa, apresentação esta que pode ser feita de modo oportunista. Consequentemente, a definição de mecanismos formais para a evidenciação e auditoria das informações textuais apresentadas nos relatórios anuais deve ser considerada quanto aos seus custos e benefícios.

Em um contexto de crescente importância do mercado de capitais brasileiro, espera-se que o presente artigo contribua para subsidiar as atividades de entidades reguladoras e analistas de mercado. O mercado de capitais brasileiro, caracterizado pela fraca proteção aos direitos de propriedade dos acionistas minoritários, elevada concentração acionária e mercado ineficiente do ponto de vista informacional estimula as necessidades de entendimento do papel das informações textuais apresentadas pelas empresas listadas neste contexto.

Entre as limitações presente no artigo, o ano de 2009 foi o ano seguinte à crise, e os resultados de algumas empresas podem ainda estar impactados por esta situação, assim como o fato de as empresas não terem sido separadas por setores. Estudos futuros podem analisar períodos de tempo mais longo, com um número maior de empresas, ou ainda utilizar outra medida para avaliar o desempenho das empresas, como o ROE (Retorno sobre o patrimônio líquido) ou o ROA (Retorno sobre os ativos), por exemplo, para verificar se os resultados se mantêm como os encontrados neste trabalho. Também há oportunidades para pesquisas futuras no mesmo tema, utilizando relatórios divulgados após a adoção completa do IFRS.

6. Referências

- Abrahamson, E., & Amir, E. (1996). The information content of the president's letter to shareholders. *Journal of Business Finance e Accounting*, 23(8), pp.1157-1182, doi: <http://dx.doi.org/10.1111/j.1468-5957.1996.tb01163.x>
- Balata, P., & Breton, G. (2005). Narratives vs numbers in the annual report: are they giving the same message to the investors?. *Review of Accounting and Finance*, 4(2), pp.5-14, doi: <http://dx.doi.org/10.1108/eb043421>
- Bardin, L. (2004). *Análise de conteúdo*. Lisboa: Edições 70.
- Beynon, M. J., Clatworthy, M. A., & Jones, M. J. (2004). The prediction of profitability using accounting narratives: a variable-precision rough set approach. *Intelligent Systems in Accounting, Finance and Management*, 12(4), pp. 227-242, doi: <http://dx.doi.org/10.1002/isaf.256>
- Breton, G., & Taffler, R. J. (2001). Accounting information and analyst stock recommendation decisions: a content analysis approach. *Accounting and business research*, 31(2), pp. 91-101, doi: <http://dx.doi.org/10.1080/00014788.2001.9729604>
- Clatworthy, M., & Jones, M. J. (2001). The effect of thematic structure on the variability of annual report readability. *Accounting, Auditing e Accountability Journal*, 14(3), pp. 311-326, doi: <http://dx.doi.org/10.1108/09513570110399890>
- _____. (2003). Financial reporting of good news and bad news: evidence from accounting narratives. *Accounting and business research*, 33(3), pp. 171-185, doi: <http://dx.doi.org/10.1080/00014788.2003.9729645>
- _____. (2006). Differential patterns of textual characteristics and company performance in the chairman's statement. *Accounting, Auditing e Accountability Journal*, 19(4), pp. 493-511, doi: <http://dx.doi.org/10.1108/09513570610679100>
- Cunha, R. K. C. D. (2009). Silva, César Augusto Tibúrcio. Análise da Facilidade de Leitura das Demonstrações Contábeis das Empresas Brasileiras: Uma Investigação do Gerenciamento de Impressões Nas Narrativas Contábeis. *Anais eletrônicos Congresso USP Controladoria e Contabilidade*, São Paulo, SP, Brasil, 9.
- CVM Comissão de Valores Mobiliários. (1987). *Parecer de Orientação CVM nº 15*, de 28/12/1987. Recuperado em 25 abril, 2013 de <http://www.cvm.gov.br/>

- CVM Comissão de Valores Mobiliários. (1999). *Nota explicativa à Instrução CVM nº. 308, de 14 de maio de 1999*. Recuperado em 8 julho, 2013 de <http://www.cvm.gov.br/>.
- Da Luz, A. T. M., Pagliarussi, M. S., Teixeira, A. M. C., & Baptista, É. C. (2009). Uma Investigação sobre a Relação entre Atribuições em Causa Própria e Governança Corporativa. *BBR-Brazilian Business Review*, 6(2), pp. 191-208, doi: <http://dx.doi.org/10.15728/bbr.2009.6.2.5>
- Grossman, S. J., & Hart, O. D. (1980). Disclosure laws and takeover bids. *The Journal of Finance*, 35(2), pp. 323-334, doi: <http://dx.doi.org/10.2307/2327390>
- Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., e Tatham, R. L. (2007). *Análise multivariada de dados*. Bookman.
- Hildebrandt, H. W., & Snyder, R. D. (1981). The Pollyanna hypothesis in business writing: initial results, suggestions for research. *Journal of Business Communication*, 18(1), pp. 5-15, doi: <http://dx.doi.org/10.1177/002194368101800102>
- Hyland, K. (1998). Exploring corporate rhetoric: metadiscourse in the CEO's letter. *Journal of Business Communication*, 35(2), pp. 224-244, doi: <http://dx.doi.org/10.1177/002194369803500203>
- Lei n.º 6.404, de 15 de dezembro de 1976. *Lei das Sociedades por Ações*. Lei 6.404, de 15/12/1976. Recuperado em 23 agosto, 2010 de http://www.planalto.gov.br/ccivil_03/Leis/L6404consol.htm.
- Li, F. (2008). Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and economics*, 45(2), pp. 221-247, doi: <http://dx.doi.org/10.1016/j.jacceco.2008.02.003>
- Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *The Journal of Finance*, 66(1), pp. 35-65, doi: <http://dx.doi.org/10.1111/j.1540-6261.2010.01625.x>
- Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to information retrieval* (Vol. 1, p. 6). Cambridge: Cambridge university press.
- Manning, C. D., & Schütze H. (1999). *Foundations of statistical natural language processing* (Ed.). MIT press.
- Milgrom, P. R. (1981). Good news and bad news: Representation theorems and applications. *The Bell Journal of Economics*, 12(2), pp. 380-391 doi: <http://dx.doi.org/10.2307/3003562>
- Pagliarussi, M.S, & Liberato, G. B. (2011) Disclosure de estratégia em relatórios anuais: uma análise de dimensões culturais, de Sistema legal e de governança corporativa em empresas de quatro países. *RAM. Revista de Administração Mackenzie (Online)*, 12(4), pp. 155-181.
- Ribeiro, J. F., Pagliarussi, M. S., Silva, A. R. L., & Silva Junior, A. (2011). Escolhas linguísticas, custos políticos e gerenciamento da imagem corporativa: o caso da Aracruz Celulose S. A.. *Cadernos EBAPE. BR*, 9(4), pp. 1136-1157.
- Rogers, R. K., & Grant, J. (1997). Content analysis of information cited in reports of sell-side financial analysts. *Journal of Financial Statement Analysis*, 3, pp. 17-31, doi: <http://dx.doi.org/10.2469/dig.v28.n2.254>
- Smith, M. (1998). Conflicting messages in annual reports, *Accountability and Performance*, 4(2), pp. 43-60,
- Smith, M., & Taffler, R. J. (2000). The chairman's statement-a content analysis of discretionary narrative disclosures. *Accounting, Auditing e Accountability Journal*, 13(5), pp. 624-647, <http://dx.doi.org/10.1108/09513570010353738>
- Spence, M. (1973). Job market signaling. *The quarterly journal of Economics*, 87(3), pp. 355-374, doi: <http://dx.doi.org/10.2307/1882010>

- Tessarolo, I. F., Pagliarussi, M. S. & da Luz, A. T. M. (2010). The justification of organizational performance in annual report narratives. *BAR-Brazilian Administration Review*, 7(2), pp. 198-212, doi: <http://dx.doi.org/10.1590/S1807-76922010000200006>
- Vargas, L. H. F., Almeida, J. E. F., & Júnior, E. M. (2014). Lucro e prejuízo sob a perspectiva da administração: como os resultados são apresentados nos relatórios de administração? *Revista de Educação e Pesquisa em Contabilidade (REPeC)*, 8(4), pp. 351-368.