

## Estatísticas descritivas e técnicas de visualização na análise de textos

Alfredo José B. Luiz<sup>1†</sup>, Carolina B. S. Luiz<sup>2</sup>

<sup>1</sup>Engenheiro Agrônomo, Dr., pesquisador na Embrapa Meio Ambiente.

<sup>2</sup>Historiadora, mestranda em História Social na FFLCH/USP. E-mail: [cbsluiz@gmail.com](mailto:cbsluiz@gmail.com).

**Resumo:** O uso de métodos quantitativos para a análise de textos possui inúmeras aplicações em diversas áreas, como na identificação de autores, na crítica e no estudo da história da literatura, no mapeamento de vocabulário e ainda na análise sociolinguística. A identificação e classificação automática de textos também é beneficiada pelas técnicas da linguística quantitativa. Por outro lado, sabe-se da dificuldade no ensino da estatística e dos métodos quantitativos, especialmente para alunos das áreas biológicas e humanas, não familiarizados às ciências exatas. Assim, o presente trabalho demonstra o uso da estatística descritiva e da visualização de dados aplicadas na análise de textos, poéticos e científicos, com o objetivo de oferecer ferramentas para profissionais que se deparem com a necessidade desse tipo de análise, mas também para disponibilizar exemplos de aplicação que possam facilitar o ensino dos conceitos estatísticos para alunos mais afeitos às ciências humanas e biológicas. O uso de estatísticas simples (como médias, frações e proporções) e de gráficos, se mostrou mais útil na distinção de poemas e resumos científicos do que apenas as suas características quantitativas absolutas.

**Palavras-chave:** Linguística quantitativa; métodos quantitativos; visualização de dados.

**Abstract:** The use of quantitative methods for the analysis of texts has innumerable applications in several areas, such as in the identification of authors, in the criticism and in the study of the history of literature, in the mapping of vocabulary and in the sociolinguistic analysis. The automatic identification and classification of texts is also benefited by quantitative linguistic techniques. On the other hand, it is known the difficulty in teaching statistics and quantitative methods, especially for students from the biological and human areas, unfamiliar with the exact sciences. Thus, the present work demonstrates the use of descriptive statistics and data visualization applied in text analysis, poetic and scientific, with the objective of offering tools for professionals who are faced with the need of this type of analysis, but also to provide examples that can facilitate the teaching of statistical concepts for students who are more interested in the human and biological sciences. The use of simple statistics (such as averages, fractions and proportions) and graphs has proved to be more useful in distinguishing poems and scientific summaries than just their absolute quantitative characteristics.

**Keywords:** Quantitative linguistics; quantitative methods; data visualization.

### Introdução

Inúmeros fatores influenciam na forma como utilizamos a linguagem, em especial quando escrevemos. Existe todo um campo de estudo da sociolinguística que busca determinar e quantificar como se dão as diferenças no uso da palavra na composição de textos escritos. Embora trabalhando com entrevistas gravadas e não com textos escritos, Oushiro (2012) associou o gênero, a faixa etária e o nível de escolaridade dos autores com alguns aspectos da forma de uso da linguagem. Ao final, a autora concluiu que a sistematização dos procedimentos das análises quantitativas é central para o avanço do campo da Sociolinguística Variacionista.

---

† Autor correspondente: [alfredo.luiz@embrapa.br](mailto:alfredo.luiz@embrapa.br).

Oliveira (2009), também usa o exemplo de fala, e não de escrita, e levanta hipótese sobre a influência do gênero do autor sobre características linguísticas. Além disso, esse autor apresenta uma interessante análise sobre um estudo do modelo estatístico utilizado nos estudos variacionistas e apresenta os aspectos teóricos estatísticos mais relevantes que contribuem para a análise da variação linguística. Segundo ele, a análise quantitativa está na base dos estudos sociolinguísticos variacionistas e a identificação da variabilidade presente nos textos, como um fenômeno dependente de determinadas variáveis associadas aos autores, se torna possível a partir do desenvolvimento de técnicas eficazes de análise de tal variabilidade.

Das diferenças 'sociais' entre autores, podemos considerar a área de atuação e a época em que escreveram os textos. Por exemplo, textos produzidos com objetivos artísticos, como poemas, devem ser quantitativamente separáveis de textos produzidos com objetivos técnicos, como resumos de trabalhos científicos. Isso se deve tanto ao tema dos escritos como a formação dos seus autores. No presente trabalho, essas diferenças (poemas versus resumos) são escolhidas como objeto de estudo do uso de técnicas quantitativas. A formação dos autores (humanas e biológicas, respectivamente, em geral) é abordada como possível obstáculo ao aprendizado da estatística, que pode ser superado pelo uso de exemplos mais afeitos às áreas de cada grupo de discentes.

Sabe-se, por exemplo, que os indivíduos com maior tendência para as ciências humanas e biológicas apresentam, em geral, resistência ou receio em relação à disciplina de Estatística. Essa opinião, além de ser consenso praticamente geral, também é comprovada por estudos da área educacional. No trabalho de Lima et al. (2017), por exemplo, afirma-se que o ensino de Estatística em cursos das Ciências Biológicas é um grande desafio para os docentes de ensino superior, e que ainda se experimenta uma forte resistência dos discentes desses cursos com relação à disciplina. Esses autores atribuem parte desse problema ao fato de que os docentes defendem (corretamente na nossa opinião) que o ensino de Estatística deve estar vinculado a um universo amplo de exercícios operacionais, entretanto, escolhem casos desvinculados de aplicações práticas ligadas aos temas de interesse dos alunos. Como, do ponto de vista desses autores, a interpretação de dados e a tomada de decisão são práticas extremamente relevantes, pois possibilitam que o discente desenvolva a capacidade de compreender melhor os fenômenos que ocorrem a sua volta, eles advogam a necessidade de ampliar a contextualização e focar na capacidade de interpretação de resultados pelos discentes.

Em outro trabalho em que foram avaliados 360 alunos, predominantemente de cursos das Ciências Sociais Aplicadas, Ciências Biológicas e Ciências Humanas (91,4%), ficou demonstrado que o reconhecimento do estudante quanto à utilidade do conteúdo é um dos fatores que influenciam a atitude em relação à Estatística (Turik et al., 2012). Esses autores ressaltaram a importância de promover uma busca continuada por formas de desenvolver o ensino de Estatística e recomendam que os docentes e as instituições de ensino explorem exemplos de aplicação úteis segundo a visão dos alunos, o que favoreceria uma visão integrada e adaptável do ensino da disciplina de acordo com o contexto de cada curso.

Sabe-se que o inverso também é verdadeiro, ou seja, alunos com maior tendência para as ciências exatas apresentam, em geral, menor domínio das disciplinas das áreas de humanas ou biológicas (Luiz, 2004), mas esse aspecto não será tema do presente trabalho.

A ideia de usar exemplos da literatura e da biologia para demonstrar a aplicabilidade de métodos estatísticos, como quase todas, não é nova. Em um artigo brilhante, Efron e Thisted (1976) reuniram três áreas: estatística; literatura e ecologia, e testaram alguns métodos usados na estimativa de espécies não observadas, um problema muito comum nos estudos ecológicos, para descobrir,

com base nos textos publicados, quantas palavras Shakespeare conhecia. Esse, inclusive, é um uso possível da análise de textos, o mapeamento de vocabulário. Entretanto, no presente estudo não ajustaremos os modelos testados por esses autores, por serem de um nível de complexidade de aplicação muito superior ao que se pretende apresentar.

Nesse sentido, o objetivo aqui é apresentar exemplos de uso de alguns métodos quantitativos simples na análise de textos. Essa linha de pesquisa possui inúmeras aplicações em diversas áreas, como na identificação de autores, tanto para efeito de crítica e história da literatura, quanto para efeitos mais pragmáticos como análise de testamentos, bilhetes de suicidas, pedidos de resgate, cartas anônimas com ameaças, etc. Também a identificação de plágio em obras artísticas ou científicas pode se beneficiar de técnicas de linguística quantitativa. Assim, o presente trabalho demonstra a aplicação da estatística descritiva e da visualização de dados na análise de textos, poéticos e científicos, com o objetivo de oferecer ferramentas para profissionais que se deparem com a necessidade desse tipo de análise, mas sobretudo para disponibilizar exemplos de aplicação que possam facilitar o ensino dos conceitos estatísticos para alunos mais afeitos às ciências humanas ou biológicas. Afinal, num artigo especialmente voltado para a forma de melhor transmitir os conceitos estatísticos aos estudantes, Stell et al. (2019) afirmam que “formalmente e informalmente, ensinar o pensamento estatístico é, talvez, a nossa maior, melhor e mais fácil oportunidade de causar um impacto positivo” e concluem com uma chamada para expandir o domínio das estatísticas para outros cursos e plataformas de comunicação.

### **Material e Métodos**

Segundo Tufte (2001), a maior qualidade de um gráfico estatístico consiste na comunicação de ideias complexas com clareza, precisão e eficiência. Nessa publicação, diversas formas de apresentação visual de dados numéricos são descritas, com suas vantagens e desvantagens e indicação de aplicações a inúmeros exemplos de diferentes tipos de dados. No presente trabalho, foram elaborados alguns gráficos nos quais foram aplicados os principais conceitos preconizados por Tufte (2001).

Como material de trabalho, foram escolhidos cinco pequenos textos, sendo três resumos de trabalhos científicos e dois poemas.

Para a comparação de textos científicos, foram escolhidos três resumos de artigos publicados em revistas ou eventos da área de pesquisa agropecuária, nos quais o autor do presente trabalho figura como coautor (Quirino et al., 1999; Schultz et al., 2016; e Watanabe et al., 2018). Dois trabalhos são semelhantes em relação a época de publicação (2016 e 2018), mas muito diferentes com relação ao tema (sensoriamento remoto e sócio economia), e o outro é mais distante dos demais, tanto na época da publicação (1999), quanto no tema da pesquisa (impacto agroambiental). Foram analisados apenas os resumos dos artigos pois o intuito é apresentar exemplos e não aprofundar um estudo comparativo entre os trabalhos. Além disso, o tamanho dos resumos, em número de palavras, é mais semelhante ao dos poemas. Comparações entre e dentro os textos poéticos e os científicos foram feitas.

Foram também comparados os poemas de mesmo título, Amor, escritos por dois dos maiores poetas brasileiros: Álvares de Azevedo (\*1831-†1852) e Hilda Hilst (\*1930-†2004). Os poemas foram escolhidos por terem o mesmo tema, serem curtos e praticamente do mesmo tamanho (em número de palavras), além da importância dos seus autores. Por outro lado, um foi escrito no século XIX e outro no XX, e por isso oferecem a possibilidade de se avaliar o efeito da época em que foram escritos, se esse é suficiente para tornar esses textos tão ou mais diferentes entre si do que em relação a textos científicos.

*Sigmae*, Alfenas, v.8, n.2, p. 67-75, 2019.

64ª Reunião da Região Brasileira da Sociedade Internacional de Biometria (RBRAS).

18º Simpósio de Estatística Aplicada à Experimentação Agrônômica (SEAGRO).

Numa pesquisa em que o objetivo principal fosse realmente a comparação de textos, haveria uma ou mais hipóteses a serem testadas, como por exemplo se os dois textos parecem ter sido escritos por uma mesma pessoa, se a área de conhecimento da pesquisa afeta a estrutura e a forma dos textos produzidos, se determinado texto foi escrito por um ser humano ou por uma máquina, se há diferença entre o vocabulário usado por diferentes autores ou em diferentes áreas, etc. No presente caso, não há uma hipótese real, pois o principal objetivo é demonstrar a aplicação da estatística e dos métodos quantitativos, mas podemos adotar, para efeitos do estudo, uma hipótese de trabalho. A suposição, então, é que, sem considerar o entendimento do conteúdo e com base apenas em estatísticas e gráficos simples produzidos pela quantificação obtida a partir dos diferentes textos, os poemas, apesar da distância no tempo em que foram escritos e das diferenças de gênero e de estilo dos autores, serão muito mais parecidos entre si do que com os resumos científicos; e que os textos científicos, apesar da distância no tempo e no tema, também serão mais parecidos entre si que com os poemas.

É importante ressaltar que, no caso dos presentes exemplos, de uma pequena quantidade de textos, cada um com poucas palavras e sem repetição (o que se daria se fossem escolhidos vários poemas de um mesmo autor ou vários resumos de uma mesma revista ou tema, por exemplo), seria mais fácil fazer a análise e classificação de forma individual e subjetiva, ou seja, uma pessoa leria os 5 textos e emitiria suas impressões de forma qualitativa. A análise quantitativa, entretanto, pode ser aplicada a um volume muito maior de dados, de forma mais automática, pois já existem programas específicos para análise quantitativa de textos (Alexa e Zuell, 2000; Franzosi et al., 2013). Por exemplo, o citado trabalho de Efron e Thisted (1976) se baseou numa pesquisa que reuniu dados de todas as obras de Shakespeare, que totalizaram 884.647 palavras, enquanto os textos aqui abordados possuem entre 107 e 213 palavras apenas.

No livro de Popescu et al. (2015) são descritas em profundidade inúmeras formas de abordagem e técnicas para a análise comparativa de poemas, no presente estudo utilizaremos os conceitos de riqueza do vocabulário, tamanho das palavras e classe de palavras. A riqueza de vocabulário pode ser avaliada pelo número de palavras diferentes utilizadas em cada texto analisado e pela quantidade de vezes que cada uma foi empregada. O tamanho das palavras é dado pelo número de letras que a palavra tem e é avaliado também pela frequência com que ocorrem os diversos comprimentos. A classe das palavras se refere a uma análise mais detalhada, para a qual é necessário um conhecimento da língua em que o texto foi escrito, de forma a permitir a classificação de cada um em verbo, substantivo, adjetivo, etc.; após a classificação, a frequência de ocorrência das classes é calculada para cada texto.

Além das medidas de cada palavra em separado (comprimento, classe, número de vezes que foi usada), Bordag e Heyer (2007) tratam da utilidade do estudo da estrutura dos textos no intuito de compará-los. Entre as diversas abordagens, por vezes complexas, apresentadas pelos autores citados, neste trabalho analisaremos, como exemplo de um componente da estrutura dos textos, o comprimento das sentenças, no caso dos resumos, e dos versos, no caso dos poemas. Esse comprimento também será avaliado pela frequência de ocorrência.

Como os textos, embora semelhantes em tamanho, são diferentes no número total de palavras, as frequências sempre serão trabalhadas na forma percentual (relativa) para permitir melhor comparação.

## Resultados

Na Tabela 1 são apresentados valores absolutos de características de cada um dos textos analisados. Para efeito de simplificação, os poemas serão chamados de PA (Azevedo, 1996) e PH (Hilst, 2017) e os resumos de RQ (Quirino et al., 1999); RS (Schutz et al., 2016) e RW (Watanabe et al. 2018). Nas análises, símbolos, como por exemplo ( . , ; : / [ ] ), não foram considerados nas contagens. Entretanto, palavras estrangeiras, nomes próprios, números, abreviaturas e siglas alfanuméricas foram contados.

Tabela 1. Valores absolutos de características dos textos analisados: poema de Alvares de Azevedo (PA); poema de Hilda Hilst (PH); resumo de Quirino et al. (RQ); resumo de Schultz et al. (RS); e resumo de Watanabe et al. (RW).

Características dos textos	PA	PH	RQ	RS	RW
Número de palavras contando repetições	107	108	164	200	213
Número de palavras sem contar repetições	67	66	108	136	126
Números de versos ou frases	17	24	5	7	8
Número de verbos	17	24	13	21	16
Número de substantivos	27	30	53	52	66
Número de pronomes	16	15	5	2	9
Número de preposições	15	20	33	34	42
Número de conjunções	9	7	13	14	13
Número de artigos	4	4	10	11	20
Número de advérbios	11	3	3	8	6
Número de adjetivos	8	5	27	12	33
Números de siglas e abreviaturas	0	0	1	27	0
Número de nomes próprios	0	0	1	3	6
Número de palavras em língua estrangeira	0	0	0	6	0
Número de numerais	0	0	1	3	2
Quantidade de números	0	0	4	7	0

Fazer uma comparação entre os textos com base nos dados da Tabela 1, os quais não sofreram nenhuma transformação ou adequação, é dificultado e até afetado pelo fato dos textos terem tamanhos (número total de palavras utilizadas) diferentes e, como resultado, poderíamos obter uma análise viesada. Por esse motivo, foram feitas transformações e cálculos simples no sentido de facilitar a comparação, conforme pode ser visto na Tabela 2.

Agora fica mais fácil perceber algumas características dos textos que podem ser úteis para diferenciá-los ou não. Por exemplo, a fração de palavras diferentes em relação ao total de palavras utilizadas variou pouco entre os textos e pelos valores apresentados não permite distinguir poemas de resumos científicos. Já a média de palavras por verso ou frase se mostra muito útil na separação entre poemas e resumos, visto que os dois poemas apresentaram valores baixos e mais parecidos entre si (6,29 e 4,50) que os resumos que apresentaram valores bem mais altos e também mais semelhantes entre si (32,80; 28,57; e 26,63) que com os dos poemas.

Mas a simples transformação dos dados originais (em valores absolutos) para porcentagens, por exemplo, como no caso das classes de palavras, não facilitou a compreensão da natureza das semelhanças ou diferenças.

Tabela 2. Valores calculados a partir dos textos analisados: poema de Alvares de Azevedo (PA); poema de Hilda Hilst (PH); resumo de Quirino et al. (RQ); resumo de Schultz et al. (RS); e resumo de Watanabe et al. (RW).

Estatísticas	PA	PH	RQ	RS	RW
Fração entre número de palavras não repetidas e total	0,626	0,611	0,650	0,668	0,592
Média de palavras por verso ou frase	6,29	4,50	32,80	28,57	26,63
Número de verbos por verso ou frase	1	1	2,6	3	2
Percentual de palavras da classe verbos	15,9	22,2	8,1	10,9	7,5
Percentual de palavras da classe substantivos	25,2	27,8	33,1	26,9	31,0
Percentual de palavras da classe pronomes	15,0	13,9	3,1	1,0	4,2
Percentual de palavras da classe preposições	14,0	18,5	20,6	17,6	19,7
Percentual de palavras da classe conjunções	8,4	6,5	8,1	7,3	6,1
Percentual de palavras da classe de artigos	3,7	3,7	6,3	5,7	9,4
Percentual de palavras da classe advérbios	10,3	2,8	1,9	4,1	2,8
Percentual de palavras da classe adjetivos	7,5	4,6	16,9	6,2	15,5
Percentual de palavras da classe siglas e abreviaturas	-	-	0,6	14,0	-
Percentual de palavras da classe nomes próprios	-	-	0,6	1,6	2,8
Percentual de palavras da classe língua estrangeira	-	-	-	3,1	-
Percentual de palavras da classe numerais	-	-	0,6	1,6	0,9
Percentual de elementos na classe números	-	-	0,6	1,6	-

Nesse caso, a matriz de resultados originada, com 12 classes e cinco textos, nos forneceu uma tal variação que fica difícil visualizar algum padrão a partir dos dados em tabela. Steel et al. (2019) ao apresentarem as, segundo eles, melhores práticas estatísticas ao alcance dos estudantes de graduação e pós-graduação, recomendam que, antes de mais nada, e sempre, se faça um gráfico com os dados. Assim, a Figura 1 foi elaborada com o intuito de verificar se a distribuição percentual das palavras em classes forma um padrão visual que permita a distinção ou o agrupamento dos textos analisados.

A Figura 1 é apenas uma das representações gráficas possíveis dos dados que dispomos. Mas ela exemplifica bem a ideia de que, como diz Tufte, (2001), muitas vezes uma boa representação visual dos dados é mais eficiente para a compreensão dos mesmos que tabelas com grande quantidade de números e até mesmo que de alguns testes estatísticos. Tanto os dados individuais como os das médias, quando apresentados na forma de gráficos de teia ou radar, permitem diferenciar os textos pelos padrões visuais dos mesmos. Observando e analisando com mais atenção, é possível verificar variáveis que pouco influem nessa diferenciação (substantivos, por exemplo) e outras com grande capacidade de separação (siglas, números, nomes próprios e palavras de língua estrangeira, por exemplo, que poderiam inclusive formar um grupo).

Se os textos tivessem sido analisados apenas a partir dos dados aqui apresentados, numéricos e gráficos deles derivados, sem que se soubesse de antemão que provinham de poemas e resumos científicos, certamente seria possível separá-los em dois grupos, que coincidiriam com os tipos de textos utilizados.

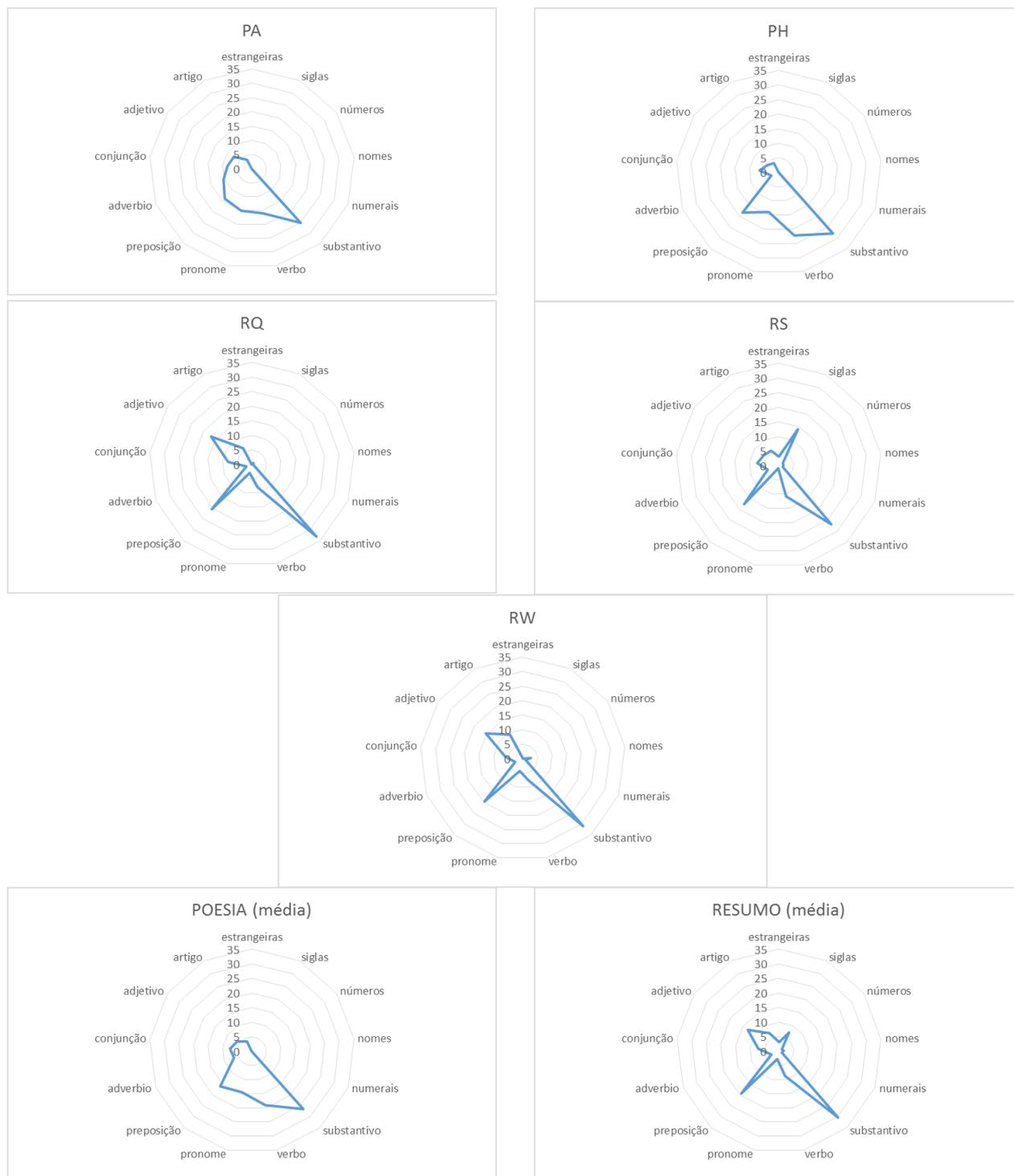


Figura 1. Gráficos da frequência relativa das classes de palavras (em porcentagem do total) em cada texto analisado: poemas de Alvares de Azevedo (PA) e de Hilda Hilst (PH); resumos de Quirino et al. (RQ), de Schultz et al. (RS) e de Watanabe et al. (RW); e nas médias dos textos de poesia e de resumos científicos.

## Conclusões

O uso de estatísticas simples, como médias, frações e proporções foi útil no estudo e na diferenciação de textos de diferentes tipos.

Em alguns casos, a visualização dos dados na forma de gráficos é ainda mais eficiente na compreensão dos padrões inerentes a cada tipo e auxilia no agrupamento de semelhantes e na separação dos diferentes.

As diferenças entre tipo de texto, poético ou científico, teve muito maior efeito nas características estudadas que a distância entre as épocas em que os textos foram produzidos.

Propomos ainda, a partir do exemplo, que o ensino de técnicas simples de estatística para alunos e pesquisadores que lidam mais frequentemente com dados textuais tem alto potencial, pois possibilita o trabalho com grande volume de dados em menor tempo.

## Referências Bibliográficas

ALEXA, M.; ZUELL, C. Text analysis software: commonalities, differences and limitations: the results of a review. *Quality & Quantity*, v. 34, n. 3, p. 299–32, 2000. DOI: 10.1023/A:1004740203542

AZEVEDO, A. *Lira dos Vinte Anos*. São Paulo: Martins Fontes, 1996. 346 p.

BORDAG, S.; HEYER, G. A. Structuralist Framework for Quantitative Linguistics. In: *Aspects of Automatic Text Analysis*. Berlin: Springer, p. 171-189, 2007. DOI: 10.1007/978-3-540-37522-7\_8

CAMARGO, B. V.; JUSTO, A. M. IRAMUTEQ: Um software gratuito para análise de dados textuais. *Temas em Psicologia*, v. 21, n. 2, p. 513-518, 2013. DOI: 10.9788/TP2013.2-16

EFRON, B.; THISTED, R. Estimating the number of unseen species: how many words did Shakespeare know? *Biometrika*, v. 63, n. 3, p. 435-447, 1976.

FRANZOSI, R.; DOYLE, S.; McCLELLAND, L. E.; RANKIN, C. P.; VICARI, S. Quantitative narrative analysis software options compared: PC-ACE and CAQDAS (ATLAS.ti, MAXqda, and NVivo). *Quality & Quantity*, v. 47, n. 6, Page 3219, 2013. DOI: 10.1007/s11135-012-9714-3

HILST, H. *Da poesia*. São Paulo: Companhia das Letras, 2017. 584 p.

LIMA, D. V.M.; MACHADO, E. J.; MACHADO, F. M. O ensino de estatística em curso de licenciatura em Ciências biológicas: o uso do laboratório de Informática. In: *12º CONGRESSO NACIONAL DE EDUCAÇÃO*, 2017, Curitiba. Anais. Curitiba: Champagnat, p. 22244-22252, 2017. Disponível em [http://educere.bruc.com.br/arquivo/pdf2017/24720\\_12503.pdf](http://educere.bruc.com.br/arquivo/pdf2017/24720_12503.pdf)

LUIZ, A. J. B. Importância de textos em português para a apropriação do conhecimento em ciência e tecnologia no Brasil. *Cadernos de Ciência & Tecnologia*, Brasília, v. 20, n. 3, p. 543-554, 2003.

**Sigmae**, Alfenas, v.8, n.2, p. 67-75, 2019.

64ª Reunião da Região Brasileira da Sociedade Internacional de Biometria (RBRAS).

18º Simpósio de Estatística Aplicada à Experimentação Agronômica (SEAGRO).

OLIVEIRA, A. J. Quantitative analysis in the study of language variation: notions of statistical and comparative analysis Varbrul and SPSS. *Revista de Estudos da Linguagem*, v. 17, n. 2, p. 93-119, jul./dez., 2009. DOI: 10.17851/2237-2083.17.2.93-119

OUSHIRO, L. Análise sociolinguística da posição variável do constituinte interrogativo. *Revista da ABRALIN*, v. 11, n. 2, p. 43-87, jul./dez., 2012

POPESCU, I. I.; LUPEA, M.; TATAR, D.; ALTMANN, G. *Quantitative Analysis of Poetic Texts*. Berlin/Boston: de Gruyter, 2015. 280 p.

QUIRINO, T. R.; LUIZ, A. J. B.; DIAS, E. C. Tecnologias agropecuárias e impacto ambiental: uma agenda para a pesquisa. *Cadernos de Ciência & Tecnologia*, Brasília, v. 16, n. 2, p. 69-96, 1999.

SCHULTZ, B.; FORMAGGIO, A. R.; EBERHARDT, I. D. R.; SANCHES, I. D.; OLIVEIRA, J. C. de; LUIZ, A. J. B. Classificação orientada a objetos em imagens multitemporais Landsat aplicada na identificação de cana-de-açúcar e soja. *Revista Brasileira de Cartografia*, Rio de Janeiro, v. 68, n. 1, p. 131-143, 2016.

STEEL, E. A.; LIERMANN, M.; GUTTORP, P. Beyond calculations: a course in statistical thinking. *The American Statistician*, v. 73, n.1, p. 392-401, 2019. DOI: 10.1080/00031305.2018.1505657

TUFTE, E. R. *The visual display of quantitative information*. 2 ed. Connecticut: Graphics Press, 2001. 200 p.

TURIK, C.; VIALI, L.; MORAES, J. F. D. Análise de atitudes de alunos universitários em relação à estatística por meio da teoria de resposta ao item. *Ciência & Educação*, v. 18, n. 1, p. 231-243, 2012. Disponível em: <http://www.scielo.br/pdf/ciedu/v18n1/14.pdf>

WATANABE, M. A.; LUIZ, A. J. B.; ABREU, L. S. de. Preços de hortifrúteis convencionais e orgânicos em feiras livres e supermercado de Barão Geraldo, Campinas, SP, Brasil. In: 56º CONGRESSO DA SOCIEDADE BRASILEIRA DE ECONOMIA, ADMINISTRAÇÃO E SOCIOLOGIA RURAL, 2018, Campinas. *Anais*. Campinas: SOBER, 2018. 11 p. Disponível em: <http://icongresso.itarget.com.br/tra/arquivos/ser.8/1/8675.pdf>

**Sigmae**, Alfenas, v.8, n.2, p. 67-75, 2019.

64ª Reunião da Região Brasileira da Sociedade Internacional de Biometria (RBRAS).

18º Simpósio de Estatística Aplicada à Experimentação Agrônoma (SEAGRO).