



DOSSIÊ: A ESCRITA ACADÊMICA A PARTIR DE DIFERENTES ABORDAGENS TEÓRICAS
ARTIGO ORIGINAL

Intricacia gramatical e densidade lexical: dois aspectos da escrita acadêmica

Grammatical intricacy and lexical density: two aspects of academic writing

Lucia Rottava¹ , Antonio Marcio Da Silva² 

1 Universidade Federal do Rio Grande do Sul - luciarottava@gmail.com

2 University of Essex - antonio.dasilva@essex.ac.uk

Como citar o artigo.

ROTTAVA, L.; DA SILVA, A. M. Intricacia gramatical e densidade lexical: dois aspectos da escrita acadêmica. *Revista Horizontes de Linguística Aplicada*, ano 23, n. 2, DT2, 2024.

Resumo

Este estudo investiga a relação entre a densidade lexical e a intricacia gramatical na escrita acadêmica, seguindo os princípios de Halliday (1985b). A análise compara duas versões de um texto universitário – a original e a reescrita –, visando compreender sua complexidade linguística. Utilizando uma abordagem qualitativa, examinamos um texto selecionado de um banco de dados de textos escritos por estudantes universitários. Os resultados revelam que a reescrita tende a aumentar o número de palavras, e confirmam que a densidade lexical e a intricacia gramatical são inversamente proporcionais, como sugerido por Halliday. Essas descobertas ressaltam a importância de considerar ambos os aspectos na escrita acadêmica, oferecendo contribuições valiosas para alunos que buscam melhorar sua produção e compreensão de textos complexos em contexto universitário.

Palavras-chave: Intricacia gramatical. Densidade lexical. Escrita acadêmica.

Abstract

This study investigates the relationship between lexical density and grammatical intricacy in academic writing, following Halliday's principles (1985b). The analysis compares two versions of a university text – the original and the rewritten –, aiming to understand their linguistic complexity. Using a qualitative approach, we examine a selected text from a database of texts written by university students. The results reveal that rewriting tends to increase the number of words and confirm that lexical density and grammatical intricacy are inversely proportional, as suggested by Halliday. These findings highlight the importance of considering both aspects in academic writing, offering valuable insights for students seeking to improve their production and understanding of complex texts in a university context.

Keywords: Grammatical intricacy. lexical density. Academic writing.

1 INTRODUÇÃO

A escrita que circula em contexto acadêmico situa um campo do conhecimento e requer de quem escreve um *continuum* entre o abstrato e o concreto (ROTTAVA; SANTOS, 2018; SANTORUM; ROTTAVA, 2024). Iniciantes em cursos de graduação têm se deparado com

Fonte de financiamento: Não há.

Conflito de interesse: Os autores declaram não haver.

Recebido em: 15 Mai 2024. Aprovado em: 16 Ago 2024.



Este é um artigo publicado em acesso aberto (Open Access) sob a licença Creative Commons Attribution Non-Commercial No Derivative, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que sem fins comerciais, sem alterações e que o trabalho original seja corretamente citado.

dificuldades na produção escrita por duas razões: as experiências prévias com reduzidas oportunidades para lidar com diferentes dimensões do texto ao longo da formação na educação básica (ROTTAVA; SANTOS, 2017) e a densidade dos textos que circulam no contexto acadêmico, que diferem quanto ao gênero textual normalmente solicitado para os estudantes produzirem na educação básica.

O recorte teórico para abordar a complexidade da escrita e seu impacto na escrita acadêmica alinha-se à Linguística Sistemico-Funcional (LSF) de viés hallidayano. Assim, o conceito de “intricacia gramatical” diz respeito à divisão do número total de orações no texto pelo número total de períodos (HALLIDAY, 1985b; 1987/2000). O conceito de “densidade lexical” refere-se à porcentagem de itens lexicais em um texto, calculado a partir do número de “palavras de conteúdo” (substantivos, adjetivos, verbos e alguns advérbios de modo) em relação ao número total de palavras de um determinado texto (HALLIDAY, 1985b).

Um levantamento inicial das características do texto escrito por iniciantes do curso de graduação, com base na LSF, verificou “construções metafóricas e seu efeito no fluxo de informação em textos produzidos por alunos ingressantes em contexto acadêmico” (ROTTAVA; SANTOS, 2018, p. 56). Os resultados indicaram baixo uso de construções nominalizadas¹, o que indicaria que esses alunos ainda precisam desenvolver o letramento acadêmico.

Essa dificuldade também foi verificada utilizando-se recursos de Inteligência Artificial (IA), em um estudo que objetivou analisar as escolhas linguísticas selecionados pelo *ChatGPT* para compor a versão reescrita de um texto (ROTTAVA; DA SILVA, 2023). O estudo analisou um texto produzido por iniciantes do curso de graduação, cujas versões reescritas pelo *ChatGPT* foram “diferentes de acordo com o comando dado, numa tentativa de se assemelhar à reescrita [por humanos], mas mostra dificuldade em produzir seguindo instruções extensas e reconhecer variáveis contidas no *feedback*” (ROTTAVA; DA SILVA, 2023, p. 1).

No que diz respeito à densidade lexical, o estudo de Rottava e Da Silva (2023) procedeu analiticamente versões de texto escrito em cinco diferentes línguas e em seis níveis de proficiência distintos, produzidos por IA, para medir a densidade lexical desses textos quanto à complexidade da escrita. Os resultados mostraram não haver um padrão entre as línguas que indicasse a complexidade textual referente à densidade lexical, especialmente devido ao fato de que a ferramenta de IA não parece diferenciar níveis de proficiência, mesmo com extensão de textos muito similares. A conclusão do referido estudo sugere que o *ChatGPT* precisa ser aprimorado para considerar melhor as características linguísticas específicas de cada língua adicional e produzir textos que reflitam com maior precisão os níveis de proficiência esperados. Para o uso pedagógico, a ferramenta deve ser utilizada com cautela, considerando suas limitações e indicando haver a necessidade de um tratamento qualitativo mais detalhado dos textos.

O estudo de Rottava e Da Silva (2023) deixou em aberto algumas questões relacionadas à proporcionalidade entre densidade lexical e intricacia gramatical. Para tanto, em um estudo ainda em andamento², analisou-se, nos textos desse mesmo *corpus*, a intricacia gramatical sob uma perspectiva lógico-semântica, para verificar se os resultados se conformam ao que Halliday (1985b) indicou: à medida que a densidade lexical aumenta, a intricacia gramatical diminui. Os resultados confirmaram a hipótese de Halliday. Sob a ótica lógico-semântica, os complexos oracionais mostraram: (a) relação de interdependência entre complexos oracionais e a intricacia gramatical; e (b) organização oracional com domínio de complexos oracionais hipotáticos.

¹ A nominalização é um recurso linguístico que, por meio da derivação de verbos e adjetivos, transforma-os em nomes. Por exemplo, um verbo pode se tornar um nome: distinguir /distinção; um adjetivo pode se transformar em um nome para qualificar: capaz/capacidade; um verbo pode se tornar um adjetivo: infectar/infeccioso; um advérbio pode se transformar em um adjetivo: rapidamente/rápido (ROTTAVA; SANTOS, 2018). Esses recursos permitem que conceitos sejam enfatizados, o que caracteriza a escrita acadêmica.

² O estudo em andamento aborda a intricacia gramatical em textos gerados pelo *ChatGPT*, cujo objetivo principal é investigar as características desses textos em termos de complexidade lógico-semântica.

Por fim, o estudo destaca que a compreensão de um texto, quanto à sua complexidade gramatical e semântica, precisa considerar o potencial de instanciação do texto (escrito, produzido), levando em conta diferentes contextos de uso e funcionamento da linguagem. Assim, o letramento acadêmico pode ser um campo vasto para pesquisas que investiguem a natureza dessa complexidade, tanto em textos produzidos por humanos quanto em textos gerados por IA.

Textos produzidos em contexto acadêmico por humanos, em mais de uma versão, ainda carecem de pesquisas para entender essa complexidade no funcionamento da linguagem. Diante dessa lacuna, este estudo congrega os conceitos de densidade lexical e de intricacia gramatical em *corpus* do português brasileiro, com textos produzidos em contexto acadêmico, visando ao desenvolvimento do letramento acadêmico. Para tanto, uma amostra de um texto e sua versão reescrita será objeto de análise detalhada. Portanto, este artigo busca contemplar essa lacuna, e o tratamento analítico pretende responder às seguintes perguntas de pesquisa:

1. Como a densidade lexical e a intricacia gramatical se manifestam na escrita (versão 1) e reescrita (versão 2) de um texto produzido por um ingressante universitário?
2. Existe uma relação inversamente proporcional entre a densidade lexical e a intricacia gramatical entre as versões do texto?
3. A reescrita do texto acadêmico, verificada na versão 2, resulta em mudanças significativas na densidade lexical e na intricacia gramatical?

Este estudo busca trazer duas contribuições: teórica e pedagógica. A contribuição teórica direciona-se para o entendimento do impacto da densidade lexical e da intricacia gramatical na clareza e sentido em textos acadêmicos; a contribuição pedagógica inclui estratégias que podem ser desenvolvidas para ajudar os alunos a melhorarem sua escrita acadêmica, considerando a densidade lexical e a intricacia gramatical.

O artigo está organizado em três principais seções, além desta introdução e das considerações finais. A primeira seção é dedicada à reflexão teórica que delinea os conceitos de densidade lexical e intricacia gramatical; a segunda seção resume os procedimentos metodológicos adotados para determinação da amostra de dados para fins analíticos; a terceira seção é destinada à apresentação dos dados e discussão dos resultados. Finaliza-se com as considerações finais, seguidas das referências.

2 REVISÃO DA LITERATURA

Em sua publicação seminal de 1985b, Halliday aborda os conceitos de densidade lexical e intricacia gramatical para discutir as diferenças entre a língua escrita e a falada. Inicialmente, Halliday define a densidade lexical na escrita, considerando a frequência, a configuração nas orações e a nominalização dos itens lexicais. Em seguida, ele descreve a intricacia gramatical, destacando que a fala e a escrita se diferenciam apenas em virtude do meio, mas que ambas podem ser analisadas quanto à sua configuração no complexo oracional e classificadas quanto aos tipos de complexidade: densidade lexical e intricacia gramatical.

No que diz respeito à densidade lexical, Halliday (1985b) ressalta que esse conceito está relacionado ao número de palavras de uma oração. Do total de palavras, ele classifica como itens lexicais as palavras de conteúdo e como itens gramaticais, as palavras funcionais. Segundo o autor, "Itens gramaticais são aqueles que funcionam em sistemas fechados na língua: em inglês, determinantes, pronomes, a maioria das preposições, conjunções, algumas classes de advérbios e verbos finitos³" (HALLIDAY, 1985b, p. 61, tradução nossa), destacando que a língua escrita contém, em média, o dobro de itens lexicais do total de palavras de uma oração ou texto.

Os itens lexicais são denominados de "palavras de conteúdo", visto que "são ITENS (de extensão variável), não palavras no sentido restrito do texto, porque podem consistir de mais

³ No original: "Grammatical items are those that function in closed systems in the language: in English, determiners, pronouns, most prepositions, conjunctions, some classes of adverb, and finite verbs" (HALLIDAY, 1985b, p. 61).

do que uma palavra [...] São LEXICAIS porque funcionam em conjuntos lexicais e não em sistemas gramaticais: isto é, entram em contrastes abertos, não fechados”⁴ (HALLIDAY, 1985b, p. 63, tradução nossa). A distinção entre item gramatical e lexical é o primeiro passo para medir a densidade lexical, requerendo que cada palavra seja considerada um item relevante para a contagem, proporcionalmente ao número total de palavras de uma oração ou texto (HALLIDAY, 1985b).

O segundo passo é a probabilidade de ocorrência desses itens lexicais inferida pela frequência no texto, ressaltando que os itens gramaticais tendem a ser mais frequentes em sua ocorrência, se comparados aos lexicais. Isso não significa que os itens gramaticais devam ser ignorados por se enquadrarem na frequência relativa, embora a frequência dos itens lexicais seja um fator importante na situação (HALLIDAY, 1985b).

Um terceiro passo é verificar se há repetição de um item lexical, visto que isso diminui a densidade lexical. Halliday destaca a repetição morfológica de um mesmo item lexical (diferir, diferença) como um fator que reduz o efeito da densidade. Como consequência, três categorias, não apenas duas, poderiam ser consideradas na análise da densidade lexical: itens gramaticais, itens lexicais de alta frequência e itens lexicais de baixa frequência (HALLIDAY, 1985b).

No entanto, a alta ou a baixa frequência de um item lexical (palavra) não é vista de maneira isolada para determinar a densidade lexical em uma oração. É o empacotamento de um item lexical em uma oração que determina a densidade informacional, como menciona Halliday (1985b). Assim, a densidade lexical é medida a partir da contagem do “número de itens lexicais por oração”⁵ (HALLIDAY, 1985b, p. 67, tradução nossa).

O conteúdo lexicalizado é, em sua grande maioria, constituído por grupos nominais, formados por substantivos e adjetivos pré-modificadores (HALLIDAY, 1985b). Nesse sentido, a nominalização impacta na densidade lexical e é também um indicador da escrita acadêmica (ROTTAVA; SANTOS, 2018). Assim como a densidade lexical, os processos de nominalização são recursos utilizados no texto para enfatizar conceitos, pois possuem maior carga semântica devido ao uso de itens lexicais.

No que diz respeito à intricacia gramatical, Halliday (1985b) traça um percurso para a definição desse tipo de complexidade na língua, delineando as características da língua falada e escrita. Para o autor, a língua falada não é menos estruturada do que a escrita, destacando que se trata de “aspectos inter-relacionados no que diz respeito ao meio, às funções e à forma exibidas: função, meio e forma”⁶ (HALLIDAY, 1985b, p. 78), conforme Figura 1:

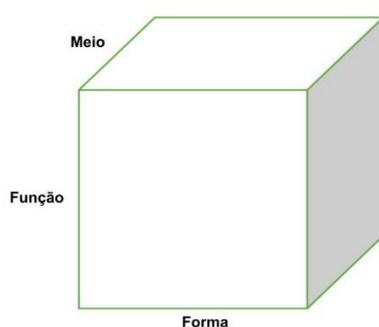


Figura 1: Aspectos inter-relacionados na escrita e na fala.
 Fonte: Adaptada pelos autores a partir de Halliday (1985b, p. 78).

⁴ No original: “[...] they are ITEMS (i.e. constituents of variable length) rather than words in the usual sense, because they may consist of more than one word: for example: stand up, take over, call off, and other phrasal verbs all function as single lexical items. They are LEXICAL because they function in lexical sets no grammatical systems: that is to say, they enter into open not closed contrasts” (HALLIDAY, 1985b, p. 63).

⁵ No original: “[...] will be measured as the number of lexical items per clause” (HALLIDAY, 1985b, p. 67).

⁶ No original: “One has to think of both written and spoken language in terms of the interrelated aspects: the nature of the medium, the functions served and form properties displayed – let us say function, medium, and form [...]” (HALLIDAY, 1985b, p. 78).

A Figura 1 mostra que a fala e a escrita são meios que dispõem de certas propriedades de organização e são apropriados para certas funções. São também manifestações de um mesmo sistema, pois possuem seu próprio tipo de complexidade decorrente da natureza do meio. Para entender essa complexidade, o objeto deste estudo é o complexo oracional realizado tanto na fala quanto na escrita; no entanto, em virtude da natureza dos dados deste artigo, o foco recai no texto escrito.

O complexo oracional é discutido por Halliday (1985b), que compara exemplos de fala e de escrita e destaca que os complexos oracionais falados tendem a ter um número maior de itens lexicais, muitos dos quais repetidos e com um certo número de itens gramaticais usados para manter o turno da interação. O exemplo trazido por Halliday (1985b, p. 80) é transcrito para ilustrar esse argumento, a partir do fraseado de três versões de um mesmo trecho, sendo o primeiro falado e os dois últimos versões escritas⁷:

1') || | or they could be in an aeroplane || and there was a great electrical storm || and they were blown off course || and the electricity made all the radio go dead || so there were no radio sounds || and nobody could hear them || |⁸

1'') || | Alternatively they might be in an aeroplane || which was blown off course by a violent electrical storm; || the electricity silenced the radio, || so that they could no longer be heard || |⁹

1''') || | As a possible alternative, the aeroplane || in which they were travelling || might have been deflected from its course by a violent electrical storm || which disrupted radio communication || and prevented them from being audible || |¹⁰

O trecho 1' está organizado em seis orações contendo 13 itens lexicais, resultando em uma densidade de 2. Para esse fraseado, Halliday (1985b) sugere duas versões escritas; o trecho 1'' contém 12 itens lexicais distribuídos em quatro orações, resultando em uma densidade de 3; por sua vez, o trecho 1''' contém 13 itens lexicais, mas organizados em três orações, excluindo-se as encaixadas¹¹, pois têm função dentro da oração a que se relacionam e não são contadas separadamente. Com esses exemplos, Halliday conclui que a fala é caracterizada por um índice de densidade lexical entre 1,5 e 2, enquanto a escrita pode ter uma densidade acima de 3, dependendo de sua formalidade, podendo alcançar até 6.

A diferença verificada nos três exemplos mostra que a fala tende a ter baixa densidade lexical, ao passo que, inversamente, o número de orações em cada complexo oracional tende a ser maior, indicando alta intricacia gramatical. Portanto, esse é um primeiro aspecto que Halliday destaca para mostrar a relação entre densidade lexical e intricacia gramatical.

O complexo oracional na língua escrita é o foco deste artigo, pois é uma maneira de mostrar "como os processos que ocorrem juntos em uma sequência estão relacionados entre si"¹² (HALLIDAY, 1985, p. 82, tradução nossa), com graus diferentes de dependências no nível oracional. Os complexos oracionais com processos que apresentam menor grau de dependência, com status semelhantes, são denominados parataxe (1, 2); por sua vez, os complexos oracionais cujos processos são interdependentes e com status desigual são denominados hipotaxe (α , β) (HALLIDAY; MATTHIESSEN, 2014).

⁷ Pelos exemplos, observa-se que a fala e a escrita são um mesmo fenômeno, porém a primeira representa o processo que envolve "fazer" e/ou "acontecer", enquanto a última representa o produto, na "ação" e/ou no "evento".

⁸ || | ou eles poderiam estar em um avião || e houve uma grande tempestade elétrica || e eles foram desviados da sua rota || e a eletricidade fez com que todos os rádios ficassem mudos || então não havia sons de rádio || e ninguém podia ouvi-los || |

⁹ || | Alternativamente, eles poderiam estar em um avião || que foi desviado da sua rota por uma violenta tempestade elétrica; || a eletricidade silenciou o rádio, || de modo que eles não puderam mais ser ouvidos || |

¹⁰ || | Como uma possível alternativa, o avião || no qual eles estavam viajando || pode ter sido desviado de sua rota por uma violenta tempestade elétrica || que interrompeu a comunicação via rádio || e impediu que eles fossem ouvidos || |

¹¹ Uma oração encaixada tem uma função complementar na oração. Entre os exemplos de orações encaixadas estão as orações relativas.

¹² No original: "[...] how the process going together in a sequence are related to each other" (HALLIDAY, 1985, p. 82).

A apropriação da escrita acadêmica pode ser compreendida por essa relação que Halliday faz entre a natureza da complexidade da língua falada e escrita, pois sempre são escolhas linguísticas que podem carregar características de uma ou de outra ou serem vistas como um *continuum*. As palavras de Halliday (1985, p. 87, tradução nossa) ilustram esse argumento:

A complexidade da língua escrita é estática e densa, enquanto a da língua falada é dinâmica e complexa. A complexidade gramatical toma o lugar da densidade lexical. As passagens de escrita altamente repletas de informações e lexicalmente densas muitas vezes tendem a ser extremamente simples em sua estrutura gramatical no que diz respeito à organização da oração (complexo oracional)¹³.

Portanto, o entendimento do texto acadêmico requer a descrição linguística de sua organização do ponto de vista sintático e semântico. A densidade lexical e a intricacia gramatical são possibilidades de implementar estratégias de ensino baseadas nessas duas variáveis para que os alunos possam melhorar sua escrita, reconhecendo que o texto escrito, por exemplo, é mais complexo pelo uso de construções nominalizadas.

3 PROCEDIMENTOS METODOLÓGICOS DO ESTUDO

Este estudo é de natureza qualitativa e interpretativa (DÖRNYEI, 2007), e o *corpus* de pesquisa é composto por um banco de dados¹⁴ com textos escritos (primeira versão) e reescritos (segunda versão), acompanhados de parecer descritivo e comentário, produzidos por alunos do primeiro semestre do curso de graduação em Letras de uma universidade brasileira ao longo de cinco anos (2014 a 2019).

O *corpus* contém a informação das tarefas, o gênero textual de acordo com a terminologia da LSF e o número de textos. Essa sistematização foi feita em estudo realizado por Silva *et al.* (2021), indicando que as temáticas abordadas nas cinco diferentes tarefas foram, respectivamente: *Apresentação Pessoal*, *Relato de uma Emoção*, *Descrição de um Processo*, *Memorial de Leitura* e, finalmente, *Opinião*. Os gêneros textuais contemplados foram relato autobiográfico, procedimento e protocolo, e discussão; essa categorização segue os princípios teóricos da LSF (MARTIN; ROSE, 2008). Cada uma das produções escritas contém um texto de primeira versão, um parecer avaliativo, uma versão de reescrita e um comentário, totalizando uma amostra de 2.140 textos¹⁵.

Esse banco de dados foi fonte de diversas pesquisas, incluindo, por exemplo, a metáfora gramatical (ROTTAVA; SANTOS, 2017; 2018; 2019), recursos de identificação do extrato semântico-discursivo (ROTTAVA; HAAG, 2019), subsistema de atitude do extrato semântico-discursivo (SILVA; ROTTAVA, 2020), subsistema de negociação do extrato semântico-discursivo (ROTTAVA; SILVA, 2022) e sistema lógico-semântico de expansão (ROTTAVA; DA SILVA, 2023). Em particular, nesse último estudo de 2023, verificou-se haver uma diferença significativa na extensão de textos em sua versão reescrita. Essa especificidade de alguns textos motivou a investigar em detalhes textos que são requeridos para escrever no curso de graduação. Do conjunto de dados, selecionou-se um texto (com duas versões), representativo dessas variáveis e que resultou da primeira tarefa¹⁶ proposta.

¹³ No original: "The complexity of the written language is static and dense. That of the spoken language is dynamic and intricate. Grammatical intricacy takes place of lexical density. The highly information-packed, lexically dense passages of writing often tend to the extremely simple in their grammatical structure, as far as the organization of the sentence (clause complex) is concerned" (HALLIDAY, 1985b, p. 87).

¹⁴ Disponível em: <https://textosletras1sem.blogspot.com/p/blog-page.html>. Acesso em: 10 abr. 2024.

¹⁵ Para um detalhamento do número de textos referentes a cada uma das propostas e versões, consultar Silva *et al.* (2021).

¹⁶ O detalhamento da proposta pode ser encontrado em Rottava e Da Silva (2023, p. 6), que consistia de: "A proposta de produção textual 01 consiste de uma apresentação pessoal. Para tanto, procure oferecer o máximo de detalhes que puder ao leitor. Procure pensar em uma particularidade sua que possa tornar sua apresentação interessante; persiga essa particularidade, visando delinear como essa especificidade mostra quem és. Ademais, para que teu leitor possa te conhecer melhor, dê exemplos e detalhe para que qualquer interlocutor, mesmo não te conhecendo pessoalmente, consiga entender o perfil apresentado."

Para tanto, a definição dos dados do estudo reportado neste artigo foi orientada pelos seguintes passos: (a) leitura dos dados presentes no *corpus*; (b) seleção de textos que apresentavam uma diferença significativa quanto à extensão entre a primeira e segunda versões; (c) cálculo da densidade lexical, identificando o número de itens lexicais do grupo nominal em relação ao número total de palavras; e (d) seleção de um texto ilustrativo da amostra com identificação de todas as orações e complexos oracionais, identificando as relações de interdependências presentes: parataxe (orações independentes com status igual em que há um iniciador 1 e um continuativo 2) ou hipotaxe (orações dependentes com status diferentes em que há um elemento dependente β e um dominante α) (HALLIDAY; MATTHIESSEN, 2014).

4 APRESENTAÇÃO DO CORPUS E ANÁLISE DOS RESULTADOS

O *corpus* de pesquisa resultou da aplicação dos critérios de seleção. Para compreender a organização do texto em termos de sua complexidade, partindo da análise da densidade lexical e, subsequentemente, da intricacia gramatical.

Inicialmente, procedeu-se à contagem das palavras em cada uma das versões do texto. Na Tabela 1 é apresentada a extensão das versões do texto e a identificação dos itens lexicais:

Tabela 1: Extensão das versões do texto.

| Texto | Nome | Adjetivo | Verbo ¹⁷ | Subtotal | Total de palavras do texto |
|----------|------|----------|---------------------|----------|----------------------------|
| Versão 1 | 66 | 20 | 34 | 120 | 289 |
| Versão 2 | 84 | 29 | 54 | 167 | 419 |
| | 66 | 20 | 34 | 120 | 289 |

Fonte: Elaboração dos autores.

Verifica-se que há uma diferença considerável no que diz respeito à extensão entre as versões do mesmo texto, ou seja, um aumento significativo de palavras, não apenas de itens lexicais, mas também de itens gramaticais. Dentre as palavras indicadoras de densidade lexical estão os itens lexicais, pois impactam semanticamente no sentido do texto, conforme é ilustrado no Gráfico 1:

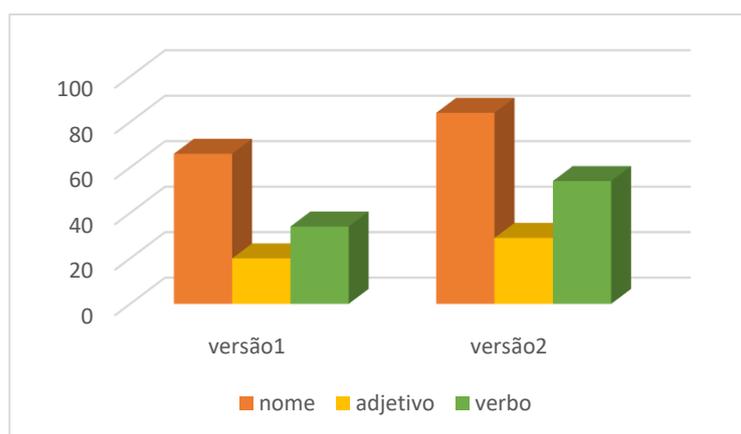


Gráfico 1: Ocorrências de itens lexicais em cada uma das versões do texto.

Fonte: Elaboração dos autores.

A tendência de aumento é confirmada pelo Gráfico 1, que mostra um crescimento nos itens lexicais do grupo nominal, particularmente nos nomes, e um leve aumento nos

¹⁷ Na contagem dos itens lexicais, não são considerados os verbos modais e os verbos de ligação no grupo verbal.

adjetivos. Da mesma forma, há um aumento de itens lexicais no grupo verbal. Esse percentual sugere que a *versão2*, que representa a reescrita do texto, pode ser um exemplo de texto acadêmico. Para entender melhor essas ocorrências, calculou-se a densidade lexical, conforme Tabela 2, que traz os percentuais:

Tabela 2. Percentuais de densidade lexical em cada uma das versões do texto.

| Texto | Nome (%) | Adjetivo (%) | Verbo (%) | Densidade lexical (%) |
|----------|----------|--------------|-----------|-----------------------|
| Versão 1 | 22,1 | 6,7 | 11,4 | 40,3 |
| Versão 2 | 20,0 | 6,9 | 12,9 | 39,9 |

Fonte: Elaboração dos autores.

Os resultados relativos à densidade lexical indicam que, na *versão2* do texto, o percentual é muito similar ao da *versão1*. A expectativa era que uma maior ocorrência de nomes e de nominalizações indicasse que a *versão2* atenderia aos critérios esperados para um texto acadêmico, em que as nominalizações enfatizam eventos, pessoas e conceitos.

Uma análise mais detalhada da densidade lexical foi realizada, verificando as ocorrências em cada uma das orações e/ou complexos oracionais das versões do texto, conforme as duas tabelas a seguir, identificadas como Tabela 3 e Tabela 4, respectivamente:

Tabela 3. Densidade lexical de cada uma das orações da *versão1* do texto.

| Oração | Nome | Adjetivo | Verbo | Total de itens lexicais | Total de palavras por oração | Densidade |
|--------------|-----------|-----------|-----------|-------------------------|------------------------------|-------------|
| 1 | 2 | 0 | 0 | 2 | 5 | 40 |
| 2 | 3 | 2 | 3 | 8 | 22 | 36,4 |
| 3 | 6 | 1 | 4 | 11 | 35 | 31,4 |
| 4 | 5 | 1 | 3 | 9 | 18 | 50,0 |
| 5 | 3 | 1 | 3 | 7 | 24 | 29,2 |
| 6 | 4 | 0 | 3 | 7 | 13 | 53,8 |
| 7 | 5 | 1 | 2 | 8 | 28 | 28,6 |
| 8 | 2 | 0 | 2 | 4 | 15 | 26,7 |
| 9 | 4 | 4 | 3 | 11 | 29 | 37,9 |
| 10 | 7 | 1 | 3 | 11 | 21 | 52,4 |
| 11 | 4 | 3 | 3 | 10 | 19 | 52,6 |
| 12 | 10 | 0 | 1 | 11 | 20 | 55,0 |
| 13 | 3 | 2 | 1 | 6 | 11 | 54,5 |
| 14 | 1 | 1 | 0 | 2 | 10 | 20,0 |
| 15 | 4 | 2 | 1 | 7 | 15 | 46,7 |
| 16 | 3 | 1 | 2 | 6 | 13 | 46,2 |
| Total | 66 | 20 | 34 | 120 | 298 | 40,3 |

Fonte: Elaboração dos autores.

Tabela 4. Densidade lexical de cada uma das orações da *versão2* do texto.

| Oração | Nome | Adjetivo | Verbo | Total de itens lexicais | Total de palavras por oração | Densidade | Oração |
|--------------|-----------|-----------|-----------|----------------------------|---------------------------------|-------------|--------------|
| 1 | 2 | 0 | 0 | 2 | 5 | 40,0 | 1 |
| 2 | 2 | 1 | 4 | 7 | 16 | 43,8 | 2 |
| 3 | 0 | 0 | 3 | 3 | 10 | 30,0 | 3 |
| 4 | 1 | 1 | 1 | 3 | 13 | 23,1 | 4 |
| 5 | 1 | 0 | 1 | 2 | 5 | 40,0 | 5 |
| 6 | 3 | 0 | 2 | 5 | 10 | 50,0 | 6 |
| 7 | 2 | 0 | 2 | 4 | 8 | 50,0 | 7 |
| 8 | 3 | 0 | 1 | 4 | 10 | 40,0 | 8 |
| 9 | 4 | 1 | 2 | 7 | 28 | 25,0 | 9 |
| 10 | 6 | 2 | 0 | 8 | 16 | 50,0 | 10 |
| 11 | 3 | 2 | 1 | 6 | 13 | 46,2 | 11 |
| 12 | 4 | 3 | 3 | 10 | 27 | 37,0 | 12 |
| 13 | 2 | 0 | 2 | 4 | 15 | 26,7 | 13 |
| 14 | 5 | 1 | 2 | 8 | 25 | 32,0 | 14 |
| 15 | 4 | 4 | 3 | 11 | 29 | 37,9 | 15 |
| 16 | 11 | 4 | 6 | 21 | 39 | 53,8 | 16 |
| 17 | 4 | 3 | 3 | 10 | 19 | 52,6 | 17 |
| 18 | 8 | 0 | 7 | 15 | 40 | 37,5 | 18 |
| 19 | 6 | 0 | 2 | 8 | 18 | 44,4 | 19 |
| 20 | 2 | 2 | 3 | 7 | 17 | 41,2 | 20 |
| 21 | 1 | 1 | 0 | 2 | 10 | 20,0 | 21 |
| 22 | 4 | 2 | 1 | 7 | 15 | 46,7 | 22 |
| 23 | 3 | 1 | 3 | 7 | 18 | 38,9 | 23 |
| 24 | 3 | 1 | 2 | 6 | 13 | 46,2 | 24 |
| Total | 84 | 29 | 54 | 167 | 419 | 39,9 | Total |

Fonte: Elaboração dos autores.

Quanto à organização do texto, as tabelas 3 e 4 revelam vários aspectos das escolhas linguísticas e da organização estrutural das versões do texto. Primeiramente, há uma confirmação do aumento significativo no número de palavras da *versão2* em relação à *versão1*. Em segundo lugar, esse aumento é também proporcional ao número de orações e complexos oracionais, passando de 16 na *versão1* para 24 na *versão2*. Em terceiro lugar, há uma maior quantidade de complexos oracionais na *versão2*, formados por um maior número de palavras.

Identificada a manifestação da densidade lexical nas versões do texto, passa-se em seguida à configuração oracional da intricacia gramatical nos dados sob análise. Para tanto, foram identificadas e contabilizadas as orações e os complexos oracionais nas duas versões do texto, como ilustram os resultados na Tabela 5:

Tabela 5. Composição oracional nas versões do texto.

| Texto | Orações simples | Complexos oracionais |
|----------|-----------------|----------------------|
| Versão 1 | 04 | 12 |
| Versão 2 | 06 | 18 |

Fonte: Elaboração dos autores.

A ocorrência oracional nas versões do texto revela uma elevação significativa de complexos oracionais. Além disso, esses complexos oracionais aumentam em relação ao número de palavras, indicando que a *versão2*, a reescrita do texto, apresenta uma maior densidade lexical e uma menor intricacia gramatical, conformando-se ao que Halliday (1985b) discute sobre as características da língua escrita. No entanto, a configuração do texto sob análise é formada também por orações e complexos oracionais que poderiam indicar características de uma linguagem falada, informal.

Com relação à inter-relação dos complexos oracionais, a Tabela 6 mostra as ocorrências de orações que estabelecem relações de parataxe e hipotaxe, respectivamente:

Tabela 6. Relações de interdependência das orações.

| Texto | Simples | Parataxe | Hipotaxe |
|----------|---------|----------|----------|
| Versão 1 | 4 | 13 | 25 |
| Versão 2 | 6 | 29 | 35 |

Fonte: Elaboração dos autores.

Os resultados, inferidos na Tabela 6, mostram que na reescrita verifica-se uma relação entre a extensão do texto e o número de orações, com um aumento significativo nas relações de parataxe, cujas orações mantêm status igual no fluxo da informação. Esse aumento também é verificado nos fraseados em que as relações de dependências entre as orações são mais elaboradas e extensas, ou seja, complexos oracionais hipotáticos.

Esse resultado pode ser melhor ilustrado a partir de um trecho da *versão2* do texto em que o escritor produz fraseados bastante complexos, incluindo relações paratáticas e hipotáticas. Para tanto, apresenta-se o trecho referente à *versão1*:

Versão 1

Desenvolvi certo fascínio – quase como Narciso ao ver sua imagem no lago – em observar meu reflexo nas páginas impressas dos livros.

1. Desenvolvi certo fascínio α
2. – quase como Narciso ao ver sua imagem no lago β
3. [[- em observar meu reflexo nas páginas impressas dos livros.]]

Versão 2

Desenvolvi certo fascínio – quase como Narciso, personagem da mitologia grega que apaixonou-se por sua própria imagem ao vê-la refletida no lago – em observar meu reflexo nas páginas impressas dos livros, e apaixonei-me pelo sentimento de me reconhecer nas histórias.

1. Desenvolvi **1 α**
- 2a. certo fascínio**1 β**
3. – quase como Narciso, personagem da mitologia grega que apaixonou-se por sua própria imagem **2**
4. ao vê-la refletida no lago –
- 2b. [[em observar meu reflexo nas páginas impressas dos livros]],**1 β**
5. e apaixonei-me pelo sentimento [[de me reconhecer nas histórias]].**3**

Os fraseados ilustrados nos dois trechos acima mostram as dificuldades em reelaborar o texto buscando dar mais densidade semântica para diminuir a intricacia gramatical, mas é exatamente o desafio que se apresenta. Esse resultado é explicado pelo fato de a reescrita aumentar a ocorrência de nominalizações, pois passou de 11 ocorrências na *versão1* para 21 na *versão2* do texto, como sendo um indicador de conhecimento pelo autor de texto acadêmico.

No entanto, esse resultado não se confirma no que diz respeito ao número de orações que compõem o complexo oracional, visto que elas passam de três na *versão1* para cinco na *versão2* do texto, incluindo alternância de orações com diferentes relações: paratáticas e hipotáticas. Ou seja, verificou-se apenas três ocorrências de parataxe na *versão2*, ao passo que a hipotaxe é recorrente nas duas versões com dois complexos oracionais em cada uma delas. Além disso, as ocorrências de itens lexicais, que carregam o sentido semântico, são similares às ocorrências de itens gramaticais.

4 CONSIDERAÇÕES FINAIS

Este estudo examinou a relação entre densidade lexical e intricacia gramatical em duas versões de um texto: a primeira versão original e a segunda, reescrita, baseando-se nos princípios de Halliday (1985b). Esses conceitos refletem aspectos inversamente proporcionais na escrita: maior intricacia gramatical em um texto sugere escolhas linguísticas mais marcadas pela oralidade, enquanto maior densidade lexical indica afastamento da oralidade.

Com relação à primeira pergunta de pesquisa que orientou a análise, os resultados indicaram que tanto a densidade lexical quanto a intricacia gramatical foram verificadas nas duas versões do texto: na escrita (*versão1*) e na reescrita (*versão2*). No entanto, as ocorrências tiveram impacto distinto no texto. De acordo com as discussões de Halliday (1985b) sobre a complexidade na escrita, ao reescrever, a densidade lexical deveria aumentar à medida que a intricacia diminuísse. No entanto, esse resultado não se confirmou e não se pode afirmar com clareza se a *versão2* do texto se aproxima mais do texto acadêmico em comparação com a primeira versão.

No que diz respeito à segunda pergunta de pesquisa, verificamos que a relação inversamente proporcional entre a densidade lexical e a intricacia gramatical nas versões do texto não se confirmou. O que se observou foi que, na reescrita, não houve um aumento da densidade lexical, mas, sim, uma diminuição. Por fim, quanto à terceira pergunta de pesquisa, que buscava analisar se a reescrita do texto, verificada na *versão2*, resulta em mudanças significativas na densidade lexical e na intricacia gramatical, verificamos que houve mudanças, mas não no sentido de torná-la mais densa. O que se constatou foi que a reescrita representou um aumento na extensão do texto, refletindo mais na intricacia gramatical do que o esperado.

Algumas hipóteses podem ser levantadas diante dos resultados verificados no *corpus* analisado neste artigo. Duas delas podem estar ligadas à concepção de escrita dos alunos construída ao longo de sua formação: a primeira diz respeito ao fato de que se pode inferir que a solicitação para a reescrita de um texto é sinônimo de falta de informações e que estas precisam ser inseridas no texto. A segunda poderia ter relação com a visão de que a complexidade da escrita é caracterizada pelo uso de complexos oracionais mais longos, alternando diversas inter-relações entre as orações.

Outro fator pode estar relacionado à prática pedagógica que não prioriza o ensino explícito da escrita e reescrita em contexto acadêmico. Nesse sentido, propostas de ensino têm sido sugeridas por linguistas sistemicistas, como o Programa Ler para Aprender, desenvolvido por pesquisadores australianos (ROSE; MARTIN, 2012; ROSE, 2020a; 2020b), com possíveis orientações para outros contextos.

Por fim, um exemplo de material destinado a alunos de cursos de graduação em contexto brasileiro foi proposto por Rottava *et al.* (2023a), o qual propõe o ensino explícito da escrita em contexto acadêmico, levando em consideração as características dos gêneros textuais requeridos nesse ambiente. Pedagogicamente, foram sugeridas tarefas que explicitam quais recursos linguísticos são necessários para planejar e propor a produção de textos, acompanhadas de um processo de avaliação baseado em critérios orientadores. Esses critérios permitiram verificar quais aspectos poderiam ser abordados e ofereceram atividades complementares para compreender o funcionamento da língua em textos que circulam no contexto acadêmico (ROTTAVA *et al.*, 2023b).

REFERÊNCIAS

- DA SILVA, A. M.; ROTTAVA, L. Densidade lexical em textos gerados pelo *ChatGPT*: implicações da Inteligência Artificial para a escrita em línguas adicionais. *Texto Livre*, v. 17, p. 1-19, 2024. Disponível em: <https://periodicos.ufmg.br/index.php/textolivres/article/view/47836>. Acesso em: 30 abr. 2024.
- DÖRNYEI, Z. *Research methods in Applied Linguistics*. New York: Oxford University Press, 2007.
- HALLIDAY, M. A. K. *An introduction to functional grammar*. London: Edward Arnold, 1985a. [3. ed. Rev. by C. M. I. M. Matthiessen, 2004; 4. ed. Rev. by C. M. I. M. Matthiessen, 2014].
- HALLIDAY, M. A. K. *Spoken and Written Language*. New York: Oxford University Press, 1985b.
- HALLIDAY, M. A. K. Spoken and written modes of meaning. In: HALLIDAY, M. A. K.; WEBSTER, J. (Ed.). *On Grammar*. Collected Works of M. Halliday (vol. 1). Edited by Jonathan Webster. London: Continuum, 1987/2002. p. 323-351.
- HALLIDAY, M. A. K.; MATTHIESSEN, C. M. I. M. *An introduction to functional grammar*. 4th ed. rev. London: Edward Arnold, 2014.
- MARTIN, J. R.; ROSE, D. *Genre relations*. Mapping culture. London/Oakville: Equinox, 2008.
- ROSE, D. Building a pedagogic metalanguage I: curriculum genres. In: MATON, K.; MARTIN, J. R.; DORAN, Y. J. (Ed.). *Studying science: knowledge, language and pedagogy*. Chapter 9. London: Taylor & Francis, 2020a. p. 236-267.
- ROSE, D. Building a pedagogic metalanguage II: knowledge genres. In: MATON, K.; MARTIN, J. R.; DORAN, Y. J. (Ed.). *Studying science: knowledge, language and pedagogy*. Chapter 10. London: Taylor & Francis, 2020b. p. 268-302.
- ROSE, D.; MARTIN, J. R. *Learning to write, reading to learn: genre, knowledge and pedagogy in the Sydney School*. London/ Oakville: Equinox, 2012.
- ROTTAVA, L.; HAAG, D. P. A metafunção textual e os recursos de identificação em memorial de leitura. *Cadernos do IL (Instituto de Letras)*, Porto Alegre, v. 1, p. 201-219, 2019.
- ROTTAVA, L.; SANTOS, S. S. dos. Grammatical Metaphor – a window to understand rewriting in academic contexts. In: NEUMANN, S. et al. (Ed.). *Challenging boundaries in linguistics: systemic functional perspectives*. Frankfurt: Peter Lang Verlag., 2017. v. 1. p. 239-253.
- ROTTAVA, L., SANTOS, S. S. dos. Os efeitos de construções metafóricas em textos produzidos em contexto acadêmico. *DELTA: Documentação e Estudos em Linguística Teórica e Aplicada*, v. 34, n. 1, p. 55-79, 2018. Disponível em: <https://revistas.pucsp.br/index.php/delta/article/view/38988>. Acesso em: 22 mar. 2024.
- ROTTAVA, L.; SANTOS, S. S. dos. The effects of metaphorical constructions in texts written in academic context. In: PÉREZ-VENEROS, M.; IZASKUN, E. (Org.). *Systemic functional linguistics at the crossroads: intercultural and contrastive descriptions of language*. Salamanca: Ediciones Universidad de Salamanca, 2019. v. 1. p. 91-98.
- ROTTAVA, L.; DA SILVA, A. M. Sistema lógico-semântico de expansão na reescrita de textos acadêmicos: escolhas linguísticas de uma estudante versus as escolhas do *ChatGPT*. *Diálogo das Letras*, v. 12, p. 1-18, 2023.
- ROTTAVA, L. et al. *Leitura e escrita na graduação – pedagogia de gêneros*. Porto Alegre: Zouk, 2023a.
- ROTTAVA, L. et al. Critérios orientadores para tarefas de leitura e de escrita de textos acadêmicos fundamentados na pedagogia com base em gêneros textuais. *ReVEL*, v. 21, n. 40, 2023b. [www.revel.inf.br].
- SANTORUM, K. A. T.; ROTTAVA, L. A metáfora gramatical ideacional e seus efeitos na escrita acadêmica. *Caderno Seminal*, n. 48, 2024. DOI: 10.12957/seminal.2024.77817. Disponível em: <https://www.e-publicacoes.uerj.br/cadernoseminal/article/view/77817>. Acesso em: 02 set. 2024.

SILVA, L. B. M. G. *et al.* O texto e a linguagem em perspectiva: as investigações científicas sob a luz da LSF na Universidade Federal do Rio Grande do Sul (UFRGS). *Organon* (UFRGS), v. 36, p. 320-337, 2021.

SILVA, L. B. M. G.; ROTTAVA, L. A escrita acadêmica e os recursos do subsistema de atitude. *Revista de Estudos Híbridos na Área da Linguagem (REHAL)*, v. 1, p. 14-35, 2020.

SILVA, L. B. M. G.; ROTTAVA, L. O sistema de negociação na escrita acadêmica: um estudo semântico-discursivo à luz da linguística sistêmico-funcional. In: OLIVEIRA, E. dos S. (Org.). *Linguística em foco: estudos reunidos sobre linguística, análise do discurso e ensino de línguas*. Itapiranga, SC: Schreiben, 2022. v. 1. p. 75-92.

Contribuição dos autores.

O artigo foi escrito colaborativamente.