

FOCO NO ENVOLVIMENTO OU NA INFORMAÇÃO: ANÁLISE DOS PRONOMES EM UM CORPUS DE CIÊNCIAS AGRÁRIAS

Shirlene Bemfica de Oliveira^{1*}

¹Instituto Federal de Minas Gerais, Ouro Preto, MG, Brasil

Deise Prina Dutra^{2**}

²Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brasil

Carolina Godoi de Faria Marques^{2***}

Resumo

Este estudo analisa o uso de pronomes pessoais enquanto marcadores do discurso de envolvimento em comparação ao uso da voz passiva como marcador do discurso informacional. O objetivo é compreender essas escolhas linguísticas na língua inglesa de autores da área de Ciências Agrárias, uma vez que são escassos os estudos sobre as características acadêmicas desta área do conhecimento (Shi; Wannaruk, 2014; Teixeira *et al.*, 2022). Este trabalho, de base quantitativa, foi desenvolvido com o apoio de ferramentas computacionais e do “*Corpus of Agrarian Sciences Scientific Articles*” (CorAgrarian) (Teixeira *et al.*, 2022; Marques 2024), com aproximadamente 2.500.000 palavras. Os resultados indicam que os artigos científicos dessa área apresentam tanto marcas do discurso de envolvimento quanto do discurso informacional, refletindo um equilíbrio entre a presença autoral e a impessoalidade.

Palavras-chave: CorAgrarian; Ciências Agrárias; pronomes; voz passiva; escrita acadêmica.

* Professora Titular do Instituto Federal de Minas Gerais, campus Ouro Preto onde leciona língua inglesa para o Ensino Médio técnico. Doutora em Estudos Linguísticos e pós-doutorado em Linguística de Corpus. E-mail: shirlene.o@ifmg.edu.br. ORCID: <https://orcid.org/0000-0003-1435-1728>.

** Professora Titular da Universidade Federal de Minas Gerais no Brasil onde ministra aulas de Inglês na graduação e de Linguística / Linguística Aplicada na pós-graduação. Seus interesses de pesquisa são linguística de corpus, corpora de aprendizes e especializados, escrita acadêmica em inglês e aprendizagem movida a dados. E-mail: dpdutra@ufmg.br. ORCID: <https://orcid.org/0000-0001-5799-5174>.

*** Mestre em Estudos Linguísticos pela UFMG, doutoranda em Estudos Linguísticos pela mesma instituição. E-mail: carol.godoi350@gmail.com. ORCID: <https://orcid.org/0009-0004-3927-4428>.



EMPHASIS ON ENGAGEMENT OR INFORMATION: AN ANALYSIS OF PRONOUNS IN AN AGRARIAN SCIENCE CORPUS

Abstract

This study analyses the use of personal pronouns as markers of involved discourse, with the use of passive voice as a marker of informational discourse. The goal is to understand these linguistic choices in English made by authors in the field of Agrarian Sciences regarding these markers, since there are few studies on the academic features of this area (Shi; Wannaruk, 2014; Teixeira *et al.*, 2022). This work, from a quantitative basis, was developed with the support of computational tools and the “*Corpus of Agrarian Sciences Scientific Articles*” (CorAgrarian) (Teixeira *et al.* 2022; Marques 2024), with approximately 2,500,000 words. The results indicate that research articles in this field display both involved discourse and informational discourse markers, reflecting a balance between authorial presence and impersonality.

Keywords: CorAgrarian; Agrarian Sciences; pronouns; passive voice; academic writing.

Introdução

A escrita acadêmica é influenciada pela língua-alvo, pelas crenças e práticas de diferentes grupos sociais e filiações disciplinares e, desta forma, a construção textual reflete as preferências atitudinais e epistemológicas desses grupos (Almeida; Miranda, 2009). Os artigos científicos, como práticas discursivas epistemológicas, exigem não apenas o domínio da gramática e do vocabulário técnico, mas também o conhecimento das características retóricas específicas dos registros, da área disciplinar e das diferentes seções do texto. Esses se constituem em uma linguagem formal que prioriza a objetividade e a impessoalidade que são marcados pela co-ocorrência de itens léxico-gramaticais, entre eles, a voz verbal (ativa e passiva) e os pronomes (Biber, 1998; Almeida; Miranda, 2009; Biber; Gray, 2016; Dutra *et al.*, 2024).

O uso dos verbos carrega, entre outras características, a voz verbal que indica a relação entre o sujeito e a ação expressa pelo verbo, informando se o sujeito é quem realiza a ação ou quem a recebe (Aquino, 2013). Em textos acadêmicos, a utilização da voz verbal e dos pronomes pessoais pode garantir clareza e objetividade por meio das vozes ativa e passiva (Aquino, 2013). Essa estratégia gramatical também é adequada para enfatizar a neutralidade dos pensamentos ou para mudar o elemento central em uma frase (Lubis *et al.*, 2024). Os autores apontam que recursos léxico-gramaticais, tais como sujeitos nominalizados, marcas desinenciais e pronomes pessoais, gerenciam a interação no contexto situacional e isso ocorre porque o grau de formalidade e a atenuação da impessoalidade no registro acadêmico são influenciados pelas marcas de autoria.

Para que o foco seja dado no conteúdo e na apresentação do estudo, pronomes pessoais de primeira e segunda pessoa são geralmente evitados diminuindo a subjetividade. No entanto, em alguns contextos, há uma variação de uso desses pronomes que se relaciona ao tipo de artigo científico (experimental ou teórico), à área de estudo e a forma que os autores usam para relatar a pesquisa ou se posicionarem no texto para expressar opiniões (Almeida; Miranda, 2009).

Almeida e Miranda (2009), em um estudo sobre artigos escritos em português, apontam que a presença de formas de primeira pessoa ocorre na Linguística e na Linguística Aplicada, sendo mais acentuada na segunda área, uma vez que a valorização e a explicitação do sujeito são encorajadas e, frequentemente, constituem o próprio foco da investigação. Segundo as autoras, em contraste, o uso do pronome “nós” na introdução e em notas de rodapé confere visibilidade e autoridade ao autor do artigo científico. No entanto, a construção impessoal pode servir como uma estratégia retórica para transferir a responsabilidade dos resultados do autor para os dados, corroborando as abordagens de Bazerman (1988) e Swales (1990). Dessa forma, enquanto a primeira pessoa do singular pode ser empregada na introdução para contextualizar o histórico do pesquisador com o objeto de estudo, indicando que a escolha da voz verbal não é uma mera questão de neutralidade, mas uma decisão retórica consciente, o uso da terceira pessoa é frequentemente associado à impessoalidade, e pode garantir a continuidade do raciocínio entre múltiplos autores e debates pré-existentes na área (Almeida; Miranda, 2009). Um outro ponto a ser ressaltado sobre o uso de construções impessoais, por exemplo, por muitos anos, houve a recomendação de se evitar o uso das passivas devido à avaliação negativa que os manuais de escritas faziam desse uso, como se essas estruturas ocorressem indistintamente em qualquer trecho de um texto acadêmico (Biber; Gray 2016). As passivas são 25% dos sintagmas verbais flexionados em inglês tendo funções muito específicas, por exemplo, na seção de metodologia de um artigo quando seria redundante a menção do autor do artigo (Biber *et al.*, 1999).

A utilização de marcadores linguísticos em artigos que revelam ora maior envolvimento ora a valorização de informações é especialmente relevante para essa pesquisa que trabalha com pesquisadores que publicam em inglês e atuam em diversas áreas do conhecimento. Hyland (2001, p. 121) apontou que “algumas áreas permitem uma maior presença autoral que outras”, tendo identificado que os textos de Engenharia tinham muito menos referências ao autor do texto do que textos da Biologia, Sociologia e Física. Neste artigo, interessa-nos entender certas características do discurso acadêmico das Ciências Agrárias por serem escassos os estudos sobre as especificidades linguísticas desta área (Shi; Wannaruk, 2014), havendo uma lacuna quanto a pesquisas que aprofundem a compreensão sobre como a autoria e a impessoalidade são gerenciadas nesse campo de estudo. Diante do exposto, este artigo visa investigar e analisar o uso de pronomes pessoais enquanto marcadores do discurso de envolvimento (*involved discourse*) em comparação ao uso da voz passiva como marcador do discurso informacional (*informational discourse*) (Biber, 1988)

para compreender as escolhas linguísticas na língua inglesa de pesquisadores da área de ciências agrárias. Sendo assim, as perguntas de pesquisa que nortearam este trabalho foram: 1. Como são utilizados os pronomes pessoais e possessivos nos artigos científicos de ciências agrárias? 2. Como se dá o uso de estruturas passivas nos artigos científicos de ciências agrárias? 3. Há variação no uso dos itens linguísticos investigados nas subáreas das ciências agrárias (agronomia, engenharia agrícola, engenharia de alimentos, engenharia florestal e zootecnia)? 4. Há variação no uso dos itens linguísticos investigados nas seções dos artigos científicos (resumo, introdução, metodologia, resultados, discussão e conclusão) das ciências agrárias?

Referencial Teórico

Nesta seção são apresentados os pressupostos teóricos que embasam o estudo: a referência na língua inglesa e o uso das estruturas passivas na língua inglesa. Além disso, exemplos do *corpus* de estudo são trazidos para ilustrar a discussão.

A referência na língua inglesa

A referência é um dos tipos de coesão textual em inglês que ocorre quando o significado de um item lexical ou frase depende de outra palavra ou frase no texto (Halliday; Hasan, 1976). Segundo os autores, a referência na língua inglesa pode ser classificada em três tipos: pessoal, demonstrativa e comparativa. Dentro desse quadro, a referência pessoal é estabelecida por pronomes pessoais que remetem a entidades (pessoas ou objetos) no domínio da realidade ou a outros elementos linguísticos denominados “antecedentes” (Celce-Murcia e Larsen-Freeman, 1999, p. 297). Para as autoras, os antecedentes na língua inglesa englobam as diversas categorias de pronomes pessoais, os demonstrativos, bem como outras formas de referência. A posse, por sua vez, pode ser expressa por três vias: por flexão (*genitive case* através da marca ‘s), de modo perifrástico (na estrutura ‘x de y’), ou referencialmente (por meio de determinantes e pronomes possessivos).

No que diz respeito aos pronomes pessoais, distinguem-se as formas de sujeito e de objeto. As primeiras atuam como sintagmas nominais de sujeito, enquanto as formas de objeto podem desempenhar funções de objeto direto, indireto ou preposicional. Ambas as formas podem atuar como predicativos do sujeito. Historicamente, a forma de sujeito é mais antiga e prescritiva. No entanto, o uso contemporâneo, especialmente na linguagem informal, demonstra uma maior frequência da forma de objeto (Celce-Murcia e Larsen-Freeman, 1999). Essa observação é corroborada por Biber *et al.* (1999), que mostram a alta frequência e relevância dos pronomes pessoais reforçando seu papel na coesão textual nos registros acadêmicos e seu papel como marcador do discurso de envolvimento. Os exemplos de 1 e 2 ilustram o seu uso no *corpus* deste estudo (CorAgrarian¹).

Exemplo 1: *“Thinking up new ideas and being creative is important to **him / her**. **He / she** likes to do things in his / her own original way”. The respondents were asked to respond through a scale ranging from 1 to 6, where 1 “not like **me** at all”, 2 “not like **me**”, 3 “a little like **me**”, 4 “somewhat like **me**”, 5 “like **me**” and 6 “very much like **me**”. (CorAgrarian.FEg.meat.12_methods)*

Exemplo 2: *A different participant in this group of anglers, who also liked to catch fish with nets in mountain lakes, followed up on the previous commented on visiting anglers’ assumed preference of not killing the fish: When **I** go angling and bring back fish home for dinner, **I** think about it like harvesting the resources nature gives to **us** (.). Harvesting from nature, **you** take what is sound, and nothing more. When there is little fish left in the lake, **we** just leave these lakes for a couple of years (.). (CorAgrarian.AnS.anhusbandry.08_results)*

Os determinantes e os pronomes possessivos também integram o sistema de referência pessoal na língua inglesa. De acordo com Celce-Murcia e Larsen-Freeman (1999), eles desempenham duas funções sintáticas primárias: podem atuar como determinantes possessivos antes de um substantivo ou substituir um sintagma nominal possessivo completo. No primeiro caso, são tradicionalmente classificados como adjetivos possessivos, conforme mostram os exemplos 3 e 4 extraídos dos dados da pesquisa.

Exemplo 3: *Ochratoxin A (OTA) is the most frequent mycotoxin in dry cured meat products, and **its** presence could be a consequence of the growth of *Aspergillus westerdijkiae*. (CorAgrarian.FEg.meat.26_introduction)*

Exemplo 4: *Suppose individual n belongs to class $c \{1, \dots, C\}$; then, **his** utility function associated with the preferred alternative $i \in J$ can be written as: where x_{nit} is a vector of observable attributes associated with alternative i and individual n observed making a choice on occasion t , and c is a class specific vector of preference parameters. (CorAgrarian.FEg.meat.12_methods)*

A forma dessas categorias varia ligeiramente dependendo de sua função, com duas formas distintas coexistindo em todos os casos, exceto para a terceira pessoa do singular masculino, cuja forma não sofre alteração (Celce-Murcia e Larsen-Freeman, 1999, p. 298). Biber *et al.* (1999), a partir da análise de *corpus*, demonstram que os determinantes possessivos são uma categoria gramatical

central, assim como os artigos definidos e os demonstrativos. Eles podem, assim, ser precedidos por um pré-determinante ou seguidos por um pós-determinante. Por outro lado, os pronomes possessivos substituem o sintagma nominal por completo e podem atuar como sujeitos ou objetos na oração, conforme mostram os exemplos 5 e 6.

Exemplo 5: *COREC and COREN values from both OR and MG (Table 2) were almost half of those reported by Rumpel et al. (2007), obtained on charcoal particles larger than 2 mm with an oxidation time equal to **ours**, for a burned agricultural soil cultivated with Job's tears.* (CorAgrarian.For.woodtec.21_discussion)

Exemplo 6: *The four size fractions we focused on in this study were chosen a priori, based on the work by Nocentini et al. (2010), to enable direct comparison of samples from two Italian pine forest sites, OR in this study, and MI in **theirs**.* (CorAgrarian.For.woodtec.21_discussion)

Estruturas passivas na língua inglesa

A voz passiva representa uma reestruturação sintática significativa da oração, que vai além de uma simples variação na ordem dos constituintes (Biber *et al.*, 1999, p. 935). Conforme os autores, essa construção altera o foco informacional da sentença ao transformar o objeto da voz ativa em sujeito da passiva, deslocando o agente para uma posição periférica ou suprimindo-o completamente. Essa transformação pode ocorrer em duas formas principais: a passiva longa, na qual o agente pode ser explicitamente expresso por meio de uma *by-phrase* ou por *by phrases* seguidas de sintagmas preposicionados, e a passiva curta, em que o agente permanece não expresso, sendo esta última especialmente recorrente em contextos formais e acadêmicos. Os exemplos 7 e 8 do *CorAgrarian* ilustram essas construções verbais.

Exemplo 7: *An increase in the dependence proximity between fuzzy sets under unfavourable growth conditions **was** predominantly **caused by** a long term decrease in pollution.* (CorAgrarian.For.woodtec.16_discussion)

Exemplo 8: *No correlations **were recorded** between 1 1 km grid Fpred and the point field as the predictors in the former generalise a broad range of abiotic conditions, whereas the point field only records unique local arrangements of the abiotic conditions.* (CorAgrarian.For.woodtec.16_discussion)

Do ponto de vista semântico, o agente da passiva geralmente corresponde ao agente da ação verbal, mas pode também assumir outros papéis, como o de experienciador, dependendo da natureza do verbo e da estrutura argumental envolvida. A disponibilidade da voz passiva está sujeita a restrições gramaticais e lexicais, sendo particularmente sensível ao tipo de verbo utilizado. Verbos transitivos diretos tendem a permitir a passivização com maior frequência, enquanto verbos intransitivos ou com complementos preposicionados apresentam limitações nesse processo (Biber *et al.*, 1999, p. 935).

Segundo Biber *et al.* (1999, p. 936), a voz passiva não se limita às duas construções apresentadas (verbo auxiliar “to be” seguido de um particípio passado). Em registros acadêmicos, é comum encontrar estruturas em que o particípio passado funciona como modificador pós-nominal dentro de um sintagma nominal, formando o que se denomina voz passiva com pós-modificadores. Essas construções consistem em orações reduzidas (*ed-clauses*) que modificam diretamente um substantivo, expressando uma ação passiva de forma condensada (substantivo + particípio passado passivo). Observe os exemplos 9 e 10 do CorAgrarian:

Exemplo 9: *The modelling platform **used** consisted of a 1 1 km grid comprised of generalised forest ecosystem properties for the CR, and a point field for the validation of the fuzzy sets.* (CorAgrarian.For.woodtec.16_methods)

Exemplo 10: *Table 3 presents mean values and standard deviations for all variables **used** to fit the full model for the four species groups.* (CorAgrarian.For.woodtec.23_results)

Nos exemplos acima, “*used*” é uma oração reduzida passiva que modifica o substantivo “*method*”. A ação é atribuída ao método, mas sem o uso explícito do verbo auxiliar. Essa estrutura, conforme descrita por Biber *et al.* (1999, p. 937), contribui para a densidade informacional do texto acadêmico e permite uma apresentação mais concisa dos elementos temáticos.

Além de sua função sintática, a voz passiva desempenha papéis discursivos relevantes. Ela contribui para a coesão textual e o ajuste contextual, permitindo a ordenação estratégica da informação e a omissão de elementos considerados irrelevantes ou já conhecidos pelo interlocutor — especialmente na forma curta (Biber *et al.*, 1999, p. 935). Essa omissão, segundo os autores, favorece a progressão temática e a manutenção do foco discursivo. Por outro lado, a passiva longa pode ser empregada para gerenciar o peso informacional da sentença, deslocando elementos mais pesados para o final da oração, o que facilita o processamento cognitivo do leitor ou ouvinte.

Metodologia

Nesta seção são apresentados procedimentos metodológicos adotados e o *corpus* utilizado para a realização deste estudo.

O Corpus: CorAgrarian

Para o desenvolvimento desta pesquisa foi utilizado o “*Corpus of Agrarian Sciences Scientific Articles*” (CorAgrarian) (Teixeira *et al.*, 2022; Marques, 2024). Trata-se de um *corpus* especializado de artigos científicos da área de Ciências Agrárias publicados em inglês em revistas de alto impacto (*Qualis A*), compilado no âmbito do projeto no qual essa pesquisa está vinculada².

De forma a ser representativo dessa área, o *corpus* é composto por cinco subcorpora, definidos a partir dos cursos ofertados pelo Instituto de Ciências Agrárias da Universidade Federal de Minas Gerais, quais sejam: Agronomia, Engenharia Agrícola, Engenharia de Alimentos, Engenharia Florestal e Zootecnia. Além disso, como critério de seleção dos artigos que comporiam o *corpus* foi estabelecido que estes deveriam obrigatoriamente ser subdivididos nas seções: resumo, introdução, metodologia, resultados, discussão e conclusão.

Ademais, considerando-se a intenção de selecionar textos representativos da escrita acadêmica contemporânea, foi adotado um critério temporal: os artigos deveriam ter sido publicados entre 2014 e 2023. Por fim, para possibilitar a realização de análises linguísticas mais aprofundadas, o *corpus* foi etiquetado em *Part of Speech (POS)* com o *Biber Tagger* (Biber, 1988) que etiqueta o *corpus* em mais de 130 traços léxico-gramaticais. A configuração do *corpus* está disposta na Tabela 1:

Tabela 1 - Composição do CorAgrarian por subárea

Subcorpus	Nº de artigos	Nº de palavras	Média
Agronomia	87	526639	6053,32
Engenharia Agrícola	90	523245	5813,83
Engenharia de Alimentos	90	459535	5105,94
Engenharia Florestal	90	505454	5616,16
Zootecnia	90	487159	5412,88
Total	447	2502032	5597,39

Fonte: Marques (2024)

Conforme a Tabela 1, o CorAgrarian (Teixeira *et al.*, 2022; Marques, 2024) é constituído por 447 textos, totalizando aproximadamente 2 milhões e meio de palavras. No que se refere ao balanceamento do *corpus*, foi definido que cada *subcorpus*, ou seja, cada subárea das Ciências Agrárias, seria composto por 90 textos. No caso da subárea de Agronomia houve a necessidade de excluir 3 artigos em razão de problemas na compilação, como a presença de textos duplicados.

Após coletados, os artigos foram segmentados e organizados de acordo com suas seções (resumo, introdução, métodos, resultados, discussão e conclusão) com auxílio de um script em *Python* para viabilizar investigações específicas por seção. A Tabela 2 apresenta a composição do *corpus* após a segmentação em seções.

Tabela 2 - Composição do CorAgrarian por seção

Seção	Nº de textos	Nº de palavras	Média
Resumo	447	346477	775,12
Introdução	447	368222	823,76
Métodos	447	440752	986,02
Resultados	444	501947	1130,51
Discussão	441	479004	1086,18
Conclusão	444	365630	823,49
Total	2670	2502032	937,09

Fonte: Marques (2024)

Como indicado na Tabela 2, os 447 artigos selecionados originaram 2670 textos.

Critérios de análise

Para possibilitar a análise, foram contabilizadas as ocorrências dos pronomes pessoais e possessivos (*pro1*, *pro2*, *pro3* e *it*)³ e as estruturas passivas sem agentes (*agls pasv*), com agentes marcados por *by* (*by pasv*) e com modificadores pós nominais (*pasv postnom mod*). Essas ocorrências foram contabilizadas com o *Tag Count* contido no *Biber Tagger* (Biber, 1988) que conta todas as ocorrências dos traços léxico-gramaticais etiquetados pelo *tagger* e as normaliza por 1000, gerando uma tabela em *.tsv* com esses dados.

A partir dessa tabela, foram extraídas as contagens dos traços a serem analisados, pelas respectivas subáreas das Ciências Agrárias e pelas seções dos artigos científicos de forma a verificar sua frequência e as diferenças de uso desses pronomes e estruturas passivas por subárea e seção.

O teste estatístico *Analysis of variance* (ANOVA) foi aplicado para verificar as diferenças estatísticas entre o uso desses traços em cada subárea e em cada seção. Esse teste fornece o valor de F e o *p-value*. O primeiro mede a razão entre a variância do uso dos pronomes entre os grupos analisados, no caso, as subáreas de Ciências Agrárias e as seções dos artigos científicos. Da realização do teste são calculados o valor de F crítico e o valor de F para os grupos analisados. Se o valor de F para os elementos analisados for superior ao F crítico as diferenças observadas são consideradas estatisticamente significativas. Já o valor p indica a probabilidade dessas diferenças serem significativas ou não. No caso, um valor de corte de $p < 0,05$ foi utilizado, isso significa que se o valor de p for inferior ao valor de corte, as diferenças observadas são estatisticamente significativas (Brezina,

2018). Além do ANOVA, foi realizado o teste R^2 . Conforme Brezina (2018), o R^2 é um teste estatístico que mede o quanto da variação total nos dados é explicada pelas diferenças entre os grupos analisados, indicando o efeito dessa variação na prática. Esses testes, ANOVA e R^2 , são independentes, e ainda que o ANOVA demonstre que não há diferenças significativas entre as áreas, o teste de R^2 ainda pode ser relevante para demonstrar a variação percebida. Da mesma forma, um valor de R^2 baixo não invalida a significância estatística dos resultados, mas indica que o nível de variação se dá dentro das subáreas e / ou seções, não entre elas.

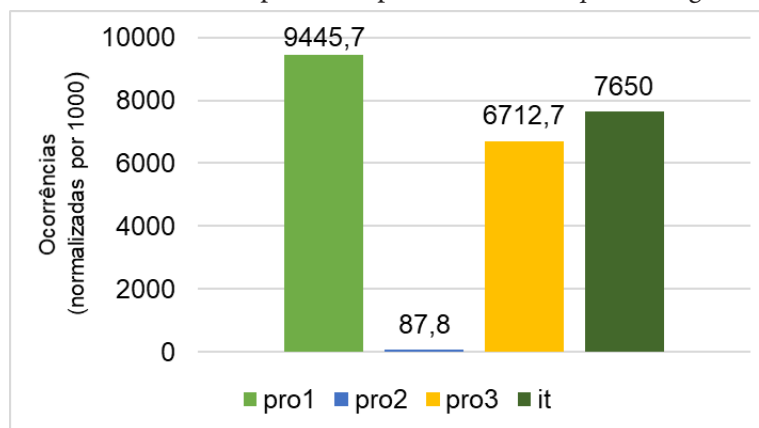
Resultados e Discussão

Os resultados a seguir trazem as análises do uso de pronomes pessoais enquanto marcadores do discurso de envolvimento em comparação ao uso da voz passiva como marcador do discurso informacional nos artigos científicos da área de Ciências Agrárias (Biber, 1998; Biber *et al.*, 1999; Almeida; Miranda, 2009; Banks, 2021; Aquino, 2013; Wang *et al.*, 2021; Lubis *et al.*, 2024). Em concordância com os objetivos da pesquisa, a investigação quantitativa teoricamente embasada concentra-se na frequência, uso e distribuição desses itens linguísticos no CorAgrarian (Teixeira *et al.*, 2022; Marques, 2024).

Frequência dos Pronomes pessoais e possessivos no CorAgrarian

Neste estudo, a frequência dos pronomes pessoais e possessivos foram extraídos do *corpus* CorAgrarian (Teixeira *et al.*, 2022; Marques, 2024) e os resultados abaixo se referem respectivamente aos pro1, pro2, pro3 e *it*, que foi contabilizado separadamente por ter uma etiqueta dedicada a ele no *tagger*. No Gráfico 1 é apresentada a ocorrência desses pronomes no *corpus* como um todo.

Gráfico 1 - Pronomes pessoais e possessivos no *corpus* CorAgrarian



Fonte: dados da pesquisa (2025)

Conforme demonstrado no Gráfico 1, os pronomes de primeira pessoa (pro1) exibem a maior frequência no *corpus* (9445.7 ocorrências por mil palavras). Esse grupo é seguido pelo pronome “*it*” e os pronomes de terceira pessoa (pro3).

Em contraste, o pronome de segunda pessoa (pro2) apresenta menor número de ocorrências no *corpus*. Essa baixa frequência de uso do pronome de segunda pessoa é coerente com as convenções do estilo acadêmico, que evita interlocução direta (Biber, 1988). Os exemplos de 11 a 13 abaixo ilustram o uso desses pronomes no *corpus*:

Exemplo 11: *I* also strongly recommend authors to provide results of all soil organic P species rather than only functional groups. (CorAgrarian.Agr.APhD.Conclusion.2016.005)

Exemplo 12: *We* did not observe any meteorological abnormalities during the study period at the site in Panama. (CorAgrarian.For.silviculture.06_results)

Exemplo 13: *Our* study shows that sowing of rare arable plants can increase niche diversity and therefore resource availability for wild bees, and it can also conserve diversity of arable plants in degraded agricultural landscapes. (CorAgrarian.Agr.ecology.07_abstract)

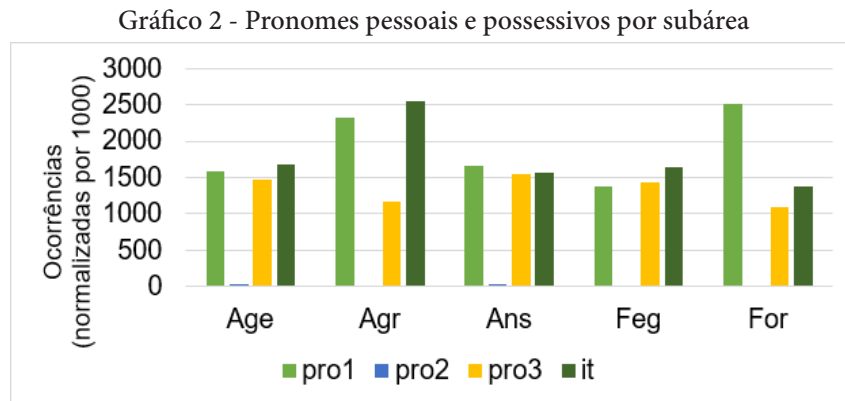
Esses resultados demonstram que o posicionamento dos escritores das Ciências Agrárias por meio dos pronomes pessoais pode ser relacionado a fatores como autopromoção, demonstração de expertise (Exemplo 11) e tentativa de equilíbrio entre as afirmações do autor com a modéstia apropriada (Exemplo 12) (Hyland, 2001). Além disso, pode sinalizar a responsabilidade da voz autoral na construção argumentativa (exemplo 13), ressaltar o papel dos pesquisadores na realização do estudo e sua contribuição direta no processo de investigação (Hyland, 2001; Banks, 2021; Wang *et al.*, 2021; Berber Sardinha, *et al.*, 2024; Dutra *et al.*, 2024). Ademais, os pronomes pessoais podem ter sido uma escolha para marcar uma interação acadêmica entre o autor e sua audiência, pois explicitam a presença do autor e compartilham a informação “como sendo de consenso geral e dessa forma, em uma perspectiva interpessoal, promovem no texto a marcação da pessoa gramatical” (Aquino, 2013, p. 70).

É importante ressaltar que o *Biber Tagger* (Biber, 1988) não diferencia o pronome pessoal *I* do numeral *I* e diante disso, pode haver um discreto enviesamento dos dados do pro1. O exemplo 14 abaixo mostra esta possível discrepância nos dados:

Exemplo 14: *Similarly, the Pp curve shape in L-DS trees changed from state **III** to state **II** and **I** when irrigation was resumed as Ψ stem recovered to similar values to the L-WW treatment. When data from all the Yara ZIM-probes of each treatment were analyzed, the Pp curve in L-WW trees was in*

the state I during the whole experiment. (CorAgrarian.AgE.irrigation.14_results)

O Gráfico 2 apresenta a frequência dos pronomes (pro1, pro2, pro3 e it) ranqueados por subárea.



Fonte: dados da pesquisa (2025)

Conforme observado no Gráfico 2, os artigos da subárea de Engenharia Florestal (For) apresentam maior uso de pro1, sugerindo maior presença do autor na escrita e da sua participação nos processos descritos. Em contraste, nos artigos de Agronomia (Agr) predominam o uso de *it*, indicando a predominância de referências a objetos, fenômenos, conforme ilustram os exemplos 15 e 16 do corpus. Esse resultado remete a um estilo acadêmico mais descritivo no qual os autores buscam por objetividade, impessoalidade e neutralidade argumentativa (Hyland, 2001; Banks, 2021).

Exemplo 15: ***Soil Grids** produces maps of soil properties for the entire globe at medium spatial resolution (250 m cell size) using state of the art machine learning methods to generate the necessary models. **It** takes as inputs soil observations from about 240 000 locations worldwide and over 400 global environmental covariates describing vegetation, terrain morphology, climate, geology and hydrology.* (CorAgrarian.Agr.soil.07_abstract)

Exemplo 16: *Notwithstanding the recognized high water storage capacity of volcanic ash soils in TMCF, **it** remains unclear which vegetation characteristics modulate the entry and movement of water in the soil.* (CorAgrarian.For.woodtec.18_introduction)

Os artigos de Engenharia Agrícola (Age), Engenharia de Alimentos (Feg) e Zootecnia (Ans), por sua vez, apresentam uma distribuição mais aproximada de pro1, *it* e pro3, sendo o pro1 mais frequente nos artigos de Zootecnia e *it* nas demais subáreas, marcando um estilo analítico, neutro, e focado nos dados.

Por fim, todas as subáreas apresentam baixa ocorrência de pro2, usados para instruções diretas para facilitar a aplicação de técnicas do campo ou para obter informações acerca dos processos, conforme ilustram os exemplos 17 e 18. Do ponto de vista discursivo, esses pronomes criam um efeito de imersão, fomentando a interação centrada no interlocutor.

Exemplo 17: *When the researcher asked “What do you think about these safety figures, and what could make them more comprehensible among migrant trainees?” (CorAgrarian.AgE.agrmachine.07_results)*

Exemplo 18: *First, ‘How likely is it that, if you diversify your agricultural production on your farm in the next five years, you would [outcome i] (unlikely e likely)’. Second, ‘How important is it that, if you diversify your agricultural production on your farm in the next five years, you would [outcome i] (unimportant important)’. For each outcome i, these two questions elicited bi and ei as shown in Fig. 1. (CorAgrarian.AnS.anhusbandry.06_methods)*

A Tabela 3, a seguir, apresenta os resultados dos testes estatísticos ANOVA e R^2 para avaliar a significância das diferenças no uso de pronomes entre as subáreas das Ciências Agrárias.

Tabela 3 - Significância estatística dos resultados por subárea

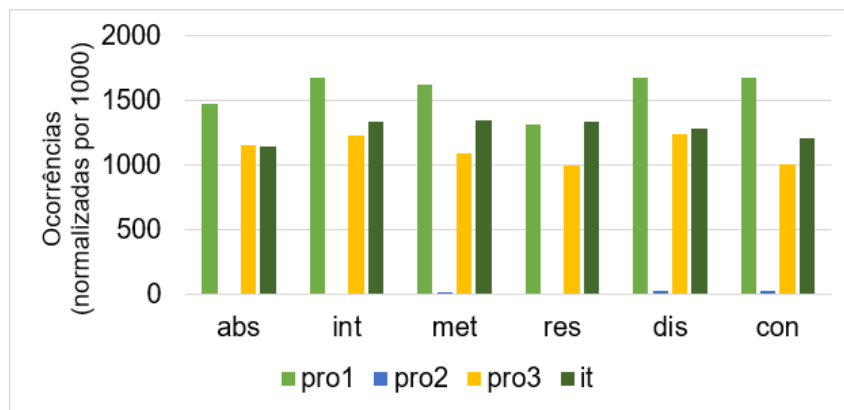
Pronome	F	Valor-p	R² (%)
pro1	19.57	0	2.85
pro2	3.74	0.0048	0.56
pro3	5.35	0.0003	0.8
it	2.54	0.038	0.38

Fonte: dados da pesquisa (2025)

Os resultados indicam diferenças significativas no uso de todos os pronomes nas subáreas investigadas, uma vez que o $p < 0,05$ e o F é superior ao crítico em todos os casos, mas o efeito prático (R^2) é baixo, exceto para o pronome de 1ª pessoa (pro1). Logo, embora a variação seja significativa, seu efeito é pequeno em termos de variação representada.

Em relação ao uso desses pronomes nas seções dos artigos científicos há uma prevalência de pronomes de primeira pessoa em todas as seções, com exceção dos resultados, como pode ser observado no Gráfico 3.

Gráfico 3 - Pronomes pessoais e possessivos por seção



Fonte: dados da pesquisa (2025)

Os pronomes pessoais e possessivos de primeira pessoa (pro1) são mais frequentes nas seções de resumo, introdução, metodologia, discussão e conclusão. Isso sugere maior envolvimento do autor nessas seções, onde há contextualização, interpretação e fechamento do texto. Esses pronomes não são tão frequentes na seção dos resultados, o que reforça o caráter impessoal, objetivo e descritivo desta seção (Wang *et al.*, 2021). Por outro lado, os pronomes pessoais e possessivos de segunda pessoa (pro2) são raros em todas as seções, com algumas ocorrências nas seções de metodologia, discussão (exemplos 19 e 20) e conclusão. Este dado é característico da escrita acadêmica, que evita interlocução direta com o leitor, privilegiando a objetividade e a impessoalidade, conforme mencionado anteriormente (Berber Sardinha *et al.*, 2024; Dutra *et al.*, 2024).

Exemplo 19: *“The canned fruits were all sweet with sugar. Really, none of it was really healthy for us, but of course we loved that government cheese. We loved that cheese; everybody did. Of course, our people had a real, real lactose intolerance. They couldn’t tolerate the milk products which they gave out, and mostly it was dry skim milk, and they used it for various things; baking, I think most of them, but they got a lot of flour and a few pasta products. It was basically those items that the government felt were surplus, and I will say somebody made a decision, oh, just give it to the Indians; they’ll eat it”* (Charles “Red” Gates, Interview, 10 7 2016). (CorAgrarian.AnS.anhusbandry.04_results)

Exemplo 20: *The welfare system makes people lazy, you just have to go show up, you just have to maybe check a few boxes and do a few things and you continually get these resources, there’s no incentive for you to progress or to teach your children the values that make a healthy lifestyle, that contributes to a healthy, positive mental health capacity”* (Cultural practitioner)

2, *Interview*, 5 3 16). (CorAgrarian.AnS.anhusbandry.04_results)

Os pronomes pessoais e possessivos de terceira pessoa (pro3), por sua vez, são mais frequentes nas seções de introdução e discussão, pois nelas são indicadas as referências a outros estudos ou autores, e, portanto, a necessidade de uma linguagem analítica, neutra, técnica e focada nos fatos (Hyland, 2001; Banks, 2021; Wang *et al.*, 2021). Por fim, os resultados mostram que o pronome “it” (etiquetado separadamente) é recorrente nas seções de introdução, metodologia, resultados e discussão, refletindo a natureza técnica, descritiva dos procedimentos, objetos de estudo e fenômenos. O exemplo 21, ilustra a presença autoral e o envolvimento técnico dentro do discurso científico na área de ciências florestais.

Exemplo 21: *The goal of this study was to evaluate the available aboveground biomass (AGB, including trunk, branches and foliage) of the widespread woody plant invader Pittosporum undulatum in the exotic woodland in Sao Miguel, Terceira and Graciosa islands (Azores archipelago), in order to assess **its** potential for energetic valorization. We used different modeling approaches in combination with forest inventory data to estimate total AGB. We sampled 127 stands dominated by P. undulatum, estimated stand density, measured diameter at breast height, basal area, tree height, and number of branches at breast height in a total of 5872 trees, and determined the AGB of 674 trees. Allometric equations were used to estimate AGB from dendrometric traits (R² 0.88).*

(CorAgrarian.For.woodtec.26_abstract)

Nesse exemplo, o pronome de terceira pessoa mantém a coesão anafórica que sustenta a impessoalidade ao manter o foco nas propriedades do objeto mencionado anteriormente – AGB (Celce-Murcia e Larsen-Freeman, 1999). Além disso, ele promove o envolvimento textual ao estabelecer uma estrutura de causa e efeito, guiando o leitor pelo fluxo argumentativo sem a necessidade de repetir termos complexos. Essa estrutura demonstra como a escrita acadêmica substitui a interação interpessoal direta por uma densidade coesiva que assegura a objetividade científica.

A Tabela 4 apresenta os resultados dos testes estatísticos empreendidos para avaliar a significância das diferenças no uso de pronomes entre as seções dos artigos científicos.

Tabela 4 - Significância estatística dos resultados por seção

Pronome	F	Valor-p	R ² (%)
pro1	2.82	0.015	0.53
pro2	1.73	0.124	0.32
pro3	2.07	0.066	0.39
it	1.25	0.285	0.23

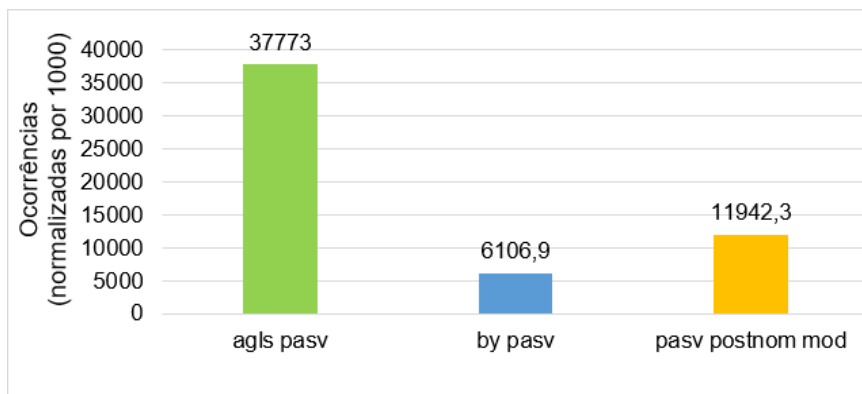
Fonte: dados da pesquisa (2025)

Os resultados dos testes para avaliar a significância das diferenças no uso de pronomes entre as seções dos artigos científicos apontam que, com exceção do pronome de 1ª pessoas, todos os demais pronomes não possuem diferenças estatisticamente significativas, uma vez que o $p > 0,05$ e o F é inferior ao crítico. O resultado do teste R², por sua vez, foi muito baixo para todos os pronomes, indicando que o critério “seção” implica em pouca variação no uso dos pronomes.

Frequência das estruturas passivas no CorAgrarian

A distribuição das estruturas passivas no CorAgrarian sugere uma tendência ao uso de estruturas passivas sem agente (*agls pasv*), com baixa incidência da passiva longa (*by passive*), o que reflete uma preferência por construções mais impessoais. O Gráfico 4 apresenta os resultados da distribuição das estruturas passivas no *corpus*.

Gráfico 4 - Ocorrências de passivas no CorAgrarian



Fonte: dados da pesquisa (2025)

A predominância da estrutura passiva sem agente (*agls pasv*) no CorAgrarian indica uma preferência por construções que promovem a impessoalidade e a objetividade, características valorizadas na escrita científica. Ao omitir o agente da ação, essa estrutura permite que o foco recaia sobre os processos, resultados ou fenômenos descritos, sem enfatizar no pesquisador que os realizou (Lubis *et al.*, 2024). Tal escolha discursiva está alinhada às convenções de áreas como as Ciências Agrárias, nas quais a neutralidade e o distanciamento autoral são frequentemente associados à credibilidade e ao rigor metodológico (Banks, 2021; Hyland, 2001). Além disso, o uso da passiva sem o agente contribui para a

construção de um discurso informacional, em que o conhecimento é apresentado como resultado de procedimentos sistemáticos e verificáveis, e não como fruto da ação individual, conforme mostra o exemplo 22 (Wang *et al.*, 2021).

Exemplo 22: *A total of 187 isolates and 21 fungal species **were recovered** from red alder and Douglas fir lap joints (Table 1). Eleven Ascomycetes, eight Basidiomycetes, and two Mucoromycetes **were identified**. Basidiomycete fungi were more frequently isolated (64.7) compared to Ascomycetes (31.6), and Mucoromycetes (3.7).* (CorAgrarian.For.woodtec.01_results)

As análises mostram baixa frequência da estrutura passiva com agente explícito precedido pela preposição *by* (*by passive*) como apresentado nos exemplos 5, 6, 7, 21 e 23 abaixo.

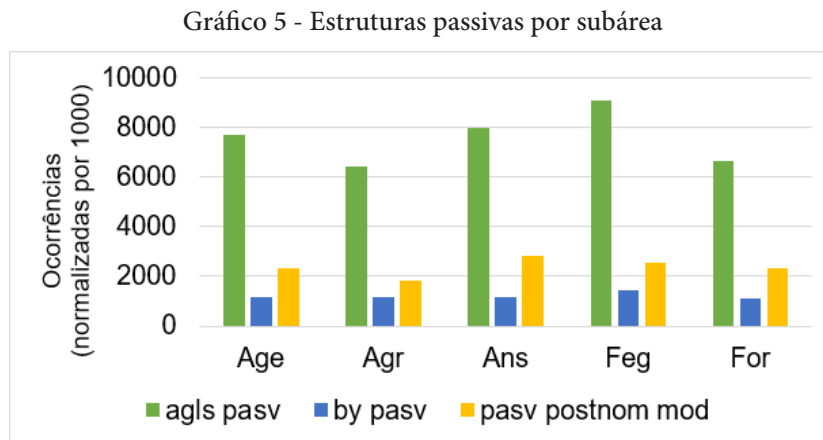
Exemplo 23: *The noise that is most apparent at the highest thermal powers **is caused by the variations** in oxygen and carbon dioxide concentrations because of the intermittent aeration.* (CorAgrarian.For.woodtec.06_results)

Este dado corrobora a tendência da escrita acadêmica de minimizar a presença do autor, favorecendo a uma construção textual mais impessoal e objetiva, conforme mencionado anteriormente. Segundo Pires (2024), esse tipo de estrutura contribui para deslocar o foco da ação para o processo ou resultado, reforçando o caráter informacional do discurso científico. A autora observa que, no discurso acadêmico, a omissão do agente é uma estratégia recorrente para apagar a autoria explícita e construir uma imagem de neutralidade e distanciamento epistemológico, conforme discutido anteriormente sobre o uso da terceira pessoa que tem a mesma função.

Por outro lado, a voz passiva com modificadores pós-nominais (*pasv postnom mod*) são frequentes sugerindo que os autores usam a estratégia de condensação informacional aumentando a densidade lexical e semântica do texto, como apresentado anteriormente nos exemplos 9, 10 e no exemplo 24 abaixo (Pires, 2024). Essa estratégia, segundo a autora, permite incorporar elementos descritivos ou qualificativos ao núcleo nominal, otimizando o espaço textual e contribuindo para a fluidez argumentativa.

Exemplo 24: ***The results obtained** from active air sampling **are reported** in Table 1. [...] **No significant differences** in total amount of fungal spores **were found** within the first 5 m from the piles, while at a distance of 300 m the level of fungal spore concentration declined to 614 CFU m³, due to atmospheric dispersion and dilution.* (CorAgrarian.For.woodtec.10_results)

A seguir, o Gráfico 5 traz o uso dessas estruturas por subárea das Ciências Agrárias.



Fonte: dados da pesquisa (2025)

As passivas sem agente (*agls pasv*) são as mais frequentes em todas as subáreas, sendo a Engenharia de Alimentos (Feg) a subárea que mais utiliza esse recurso linguístico, sugerindo um estilo mais técnico e descritivo. A Zootecnia (Ans), apesar de se valer predominantemente do uso da passiva sem agente, apresenta relativa frequência de uso da voz passiva com modificadores pós-nominais (*pasv postnom mod*), o que pode refletir uma preferência por construções mais densas e informativas. Em contraste, a voz passiva com a preposição *by* (*by pasv*) é pouco utilizada em todas as subáreas.

Nos testes estatísticos empreendidos, cujo resultado se encontra na Tabela 5, todas as estruturas passivas apresentaram um $p < 0.05$ e F superior ao crítico, indicando uma diferença estatisticamente significativa no seu uso em todas as subáreas analisadas. Em contraste, o valor do teste de R^2 tem efeito moderado nas passivas sem agente (3,15%), enquanto nas demais seu valor é pequeno.

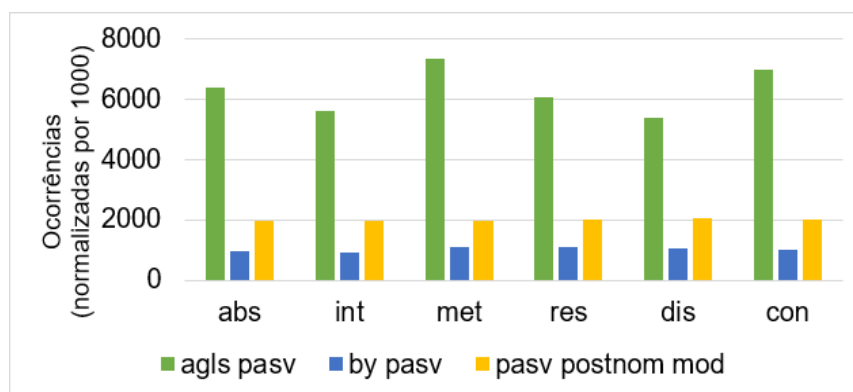
Tabela 5 - Significância estatística dos resultados por subárea

Tipo	F	Valor-p	R^2 (%)
agls pasv	21,68	0,000	3,15
by pasv	4,38	0,000	0,65
pasv postnom mod	10,61	0,000	1,57

Fonte: dados da pesquisa (2025)

O uso das estruturas passivas contrastadas por seção pode ser observado no Gráfico 6.

Gráfico 6 - Ocorrências de passivas por seção



Fonte: dados da pesquisa (2025)

No que diz respeito às seções, há também predominância do uso das passivas sem agente (*agls pasv*), sendo especialmente utilizadas nas seções de metodologia e conclusão. Na metodologia em virtude da natureza descritiva e objetiva dos procedimentos técnicos e operacionais, deslocando o foco do agente para a ação realizada enfatizando a replicabilidade e neutralidade científica (Pires, 2024). Já na seção de conclusão, segundo a autora, o uso da passiva sem o agente permite apresentar os resultados com distanciamento autoral, evitando a personalização das interpretações e mantendo o tom impessoal.

A baixa frequência das estruturas passivas com agente explícito precedido de *by* (*by pasv*) embora constante em todas as seções, reflete a escolha dos autores pelo item linguístico em momentos em que a explicitação do agente pode ser relevante para a compreensão da ação. Para Hyland (2001), a escolha pela voz passiva com agente explícito pode servir à construção de um discurso mais transparente, sem comprometer a impessoalidade predominante no registro acadêmico.

A distribuição uniforme da voz passiva com modificadores pós-nominais (*pasv postnom mod*) ao longo das seções do *corpus* revela uma preferência dos autores por estruturas que mantêm o foco no conteúdo e nos objetos e que favorecem a condensação informacional e o aumento da densidade argumentativa (Hyland, 2001). Para o autor, esses modificadores, ao expandirem o grupo nominal com qualificações, especificações ou complementos, permitem que as informações sejam integradas e compactadas valorizando a escrita científica.

Na Tabela 6 são apresentados os resultados dos testes de significância estatística por seção.

Tabela 6 - Significância estatística dos resultados por seção

Tipo	F	Valor-p	R ² (%)
agls pasv	14.88	0	2.72
by pasv	2.02	0.07	0.38
pasv postnom mod	0.15	0.981	0.03

Fonte: dados da pesquisa (2025)

No cálculo da significância estatística no uso das estruturas passivas entre as seções, os resultados demonstram que apenas a voz passiva sem o agente apresenta um resultado relevante, tendo $p < 0,05$ e F superior ao crítico indicando que há evidências para rejeitar a hipótese nula de igualdade entre as seções quanto ao uso dessas estruturas. No entanto, embora o efeito da sua variação seja estatisticamente relevante, é considerado pequeno a moderado (2,72%), explicando apenas uma pequena parte da variabilidade observada.

Nas demais estruturas passivas analisadas, os resultados não atingiram significância estatística e o resultado do teste de R^2 foi muito baixo para representar qualquer variação entre as seções pelos modelos testados. Isso reforça a ideia de que, embora a passiva sem o agente tenha um papel discursivo mais marcado em determinadas seções, como metodologia e conclusão, as demais estruturas mantêm distribuição mais homogênea ou aleatória, sem padrão estatístico consistente.

Conclusão

Esta pesquisa buscou primeiramente investigar o uso dos pronomes e da voz verbal em artigos científicos da área de Ciências Agrárias publicados em inglês enquanto marcadores do discurso de envolvimento e do discurso informacional de forma a caracterizar linguisticamente a escrita acadêmica dessa área do conhecimento. Para tanto, foram analisados os pronomes pessoais, possessivos e a voz passiva (passiva sem agente, passiva com *by* e com modificadores pós-nominais) no CorAgrarian (Teixeira *et al.* 2022; Marques 2024).

Em relação ao uso dos pronomes pessoais e possessivos nos artigos científicos das ciências agrárias, a análise do *corpus* revelou marcas tanto do discurso de envolvimento quanto do discurso informacional. Observou-se a adoção de estratégias híbridas, em que a impessoalidade convive com as marcas de autoria, mantendo o rigor metodológico e, ao mesmo tempo, enfatizando a presença do pesquisador, propiciando ainda um discurso descritivo e técnico. Os testes estatísticos empreendidos (ANOVA e do R^2) revelam que, apesar das especificidades linguísticas e discursivas de cada subárea e seção, prevalece uma relativa homogeneidade estilística nos artigos científicos de Ciências Agrárias.

Os resultados indicam que os pronomes pessoais e possessivos de primeira pessoa e o pronome *it* são os mais recorrentes, seguidos pelos pronomes pessoais e possessivos de terceira pessoa. Esse padrão evidencia que, por um lado, a presença discursiva dos pesquisadores, em especial nos artigos de Engenharia Florestal e nas seções resumo, introdução, metodologia, discussão e conclusão – independentemente da subárea. Por outro lado, há a predominância de referências a objetos, fenômenos e procedimentos científicos principalmente nos artigos de Agronomia e nas seções de resultados. O uso reduzido de pronomes de segunda pessoa, por subáreas ou por seções, confirma a tendência à objetividade e o afastamento da interlocução direta na escrita acadêmica. Portanto, o uso dos pronomes nas Ciências Agrárias parece trazer traços híbridos tanto de modelos

textuais tradicionais (pouco uso de pronomes pessoais e mais uso de voz passiva) quanto modelos progressistas (mais uso de pronomes pessoais e menos uso da voz passiva) combinando características de ambos (Banks, 2021). Em relação à significância estatística dos resultados, os pronomes pessoais e possessivos de primeira pessoa apresentam uma variação significativa ($p < 0.05$), mas o teste R^2 baixo, explica apenas uma parte dos resultados, seja por subáreas ou por seções. Já os pronomes pessoais e possessivos de segunda e terceira pessoa apresentaram variação significativa entre as subáreas. Entretanto, o teste R^2 é baixo, sugere que, embora haja diferenças, ela explica uma pequena parte da variabilidade total. Esses pronomes não apresentam variação significativa entre seções e seu R^2 é baixo, o que pode indicar estabilidade e consistência estilística ao longo das seções dos artigos.

No que se refere ao uso da voz passiva, verificou-se uma preferência marcante pelas passivas sem o agente em todas as subáreas e seções, sendo mais frequentes nos textos de Engenharia de Alimentos e Zootecnia e nas seções de metodologia e conclusão, o que reforça a impessoalidade e a ênfase nos procedimentos em detrimento do autor da ação e um estilo mais denso e técnico. As construções com modificadores pós nominais também são utilizadas em todas as subáreas, sendo mais frequentes nos artigos de Zootecnia, e tendo uma frequência similar ao longo de todas as seções. Seu uso sugere uma estratégia de condensação informacional que contribui para a densidade textual típica da escrita científica (Pires, 2024). Em contraste, a passiva com *by* apresenta baixa ocorrência e frequência similar em todas as subáreas e seções, alinhando-se às práticas acadêmicas que privilegiam a ocultação do agente.

Todas as estruturas passivas analisadas apresentaram diferenças estatisticamente significativas entre as subáreas. Enquanto o R^2 tem efeito moderado nas passivas sem o agente e pequeno nas demais, tendo efeitos práticos reduzidos. Por outro lado, a análise por seções revelou que, com exceção da passiva sem agente, as demais estruturas não têm variação estatisticamente relevante e seu R^2 é baixo, logo o efeito prático dessas variações é pequeno.

A análise do CorAgrarian corrobora os resultados de outros estudos (e.g. Biber *et al.*, 1999; Hyland, 2001; Banks, 2021; Teixeira, 2022; Berber Sardinha *et al.*, 2024; Marques, 2024) que evidenciam o uso dos pronomes pessoais e possessivos e da voz passiva na escrita acadêmica. Os primeiros são usados para autopromoção e marcação da voz autoral (pronomes de primeira pessoa), assim como para referência a outros estudos, objetos, métodos e procedimentos e generalizações (prônimo *it* e pronomes de terceira pessoa). A voz passiva reflete a preferência por estruturas impessoais e objetivas nos textos. Essa variação confirma o papel funcional da voz passiva como ferramenta de adequação estilística e pragmática, adaptando-se às exigências comunicativas de cada contexto (Biber *et al.*, 1999; Pires, 2024). Essas percepções são valiosas para a compreensão das convenções da escrita acadêmica na área de Ciências Agrárias e podem orientar abordagens pedagógicas voltadas para o aprimoramento das

habilidades de escrita de pesquisadores brasileiros, com vistas à divulgação internacional de suas pesquisas.

DECLARAÇÃO DE DISPONIBILIDADE DE DADOS: Dados disponíveis mediante solicitação para os autores.

Notas finais

1. O corpus será descrito detalhadamente na seção de metodologia.
2. Este estudo faz parte do projeto de pesquisa: “Escrita acadêmica em língua inglesa nas ciências agrárias: necessidades e insumos para aplicações pedagógicas” financiado pela FAPEMIG sob o número APQ-01173-22 que visa contribuir para o aprimoramento da escrita acadêmica produzida em língua inglesa por pesquisadores das Ciências Agrárias e auxiliar na divulgação internacional da produção científica brasileira.
3. Pro1 são os pronomes pessoais e possessivos de 1ª pessoa singular e plural (I, we, me, us, my, our, mine, ours). Pro2 são os pronomes pessoais e possessivos de 2ª pessoa singular e plural (you, your, yours). Pro3 são os pronomes pessoais e possessivos de 3ª pessoa singular e plural (he, she, him, her, his, hers, they, them).

Referências

- ALMEIDA, J. R. C. F.; MIRANDA, M. A. O uso de pronomes de primeira pessoa em artigos acadêmicos: uma abordagem baseada em *corpus*. *Veredas On-Line: Linguística de Corpus e Computacional*. PPG Linguística/UFJF: Juiz de Fora, v. 2, p. 68-83, 2009. ISSN: 1982-2243.
- AQUINO, L. D. Interação, autoria e envolvimento no gênero editorial: uma análise sistêmico-funcional. *Working Papers em Linguística*. Florianópolis, v. 14, n. 3, p. 64-81, ago./dez. 2013. DOI: <https://doi.org/10.5007/1984-8420.2013v14n3p64>.
- BANKS, D. Passive voice, first person pronouns and mental process verbs in the physical sciences research article. *Revista de Lingüística y Lenguas Aplicadas*, v. 16, p. 37-48, 2021. DOI: <https://doi.org/10.4995/rlyla.2021.14434>.
- BAZERMAN, C. *Shaping Written Knowledge: The Genre and Activity of the Experimental Article in Science*. Wisconsin: The University of Wisconsin Press, 1988. Disponível em: https://monoskop.org/images/9/9d/Bazerman,_Charles_-_Shaping_Written_Knowledge._The_Genre_and_Activity_of_the_Experimental_Article_in_Science.pdf Acesso em 01/08/2025.
- BIBER, D. *Variations across speech and writing*. Cambridge: Cambridge University Press, 1988.
- BIBER, D.; CONRAD, S.; REPPEN, R. *Corpus linguistics: investigating language structure and use*. Cambridge: Cambridge University Press, 1998.
- BIBER, D.; JOHANSSON, S.; LEECH, G.; CONRAD, S.; FINEGAN, E. *Longman Grammar of Spoken and Written English*. Longman: London, 1999.
- BIBER, D; CONRAD, S. *Register, genre, and style*. Cambridge: Cambridge University Press, 2009.
- BIBER, D.; GRAY, B. *Grammatical Complexity in Academic English: Linguistic Change in Writing*. Cambridge: Cambridge University Press, 2016.

- BERBER SARDINHA, T.; DUTRA, D. P.; DELFINO, M. C. N.; SHIMAZUMI, M. KAUFFMANN, C. H.; MACEDO, L. D.; VICTOR, A. C. Key feature analysis of academic writing in the humanities. In: AMERICAN ASSOCIATION FOR CORPUS LINGUISTICS CONFERENCE (AACL), 16., 2024, Eugene. *Anais [...]*. Eugene: University of Oregon, 2024.
- BREZINA, V. *Statistics in Corpus Linguistics. A practical guide*. Cambridge: Cambridge University Press, 2018.
- BOCORNY, A. E. P.; DUTRA, D. P.; TEIXEIRA, G. L. *Mapping the language needs of the Brazilian researchers in the Agrarian Sciences: a corpus-based study of abstracts*. In: AMERICAN ASSOCIATION FOR CORPUS LINGUISTICS CONFERENCE (AACL), 15., 2022, Flagstaff. *Anais [...]*. Flagstaff: Northern Arizona University, 2022.
- CELCE-MURCIA, M.; LARSEN-FREEMAN, D. *The Grammar Book: An ESL/EFL Teacher's Course*. 2nd ed. Boston, MA: Heinle & Heinle, 1999.
- DUTRA, D. P.; QUEIROZ, J. M. S.; MACEDO, L. D.; COSTA, D. D.; MATTOS, E. Adjectives as nominal pre-modifiers in chemistry and applied linguistics research articles. In: Ute Römer-Barron; Viviana Cortes; Eric Friginal (Ed.) *Advances in Corpus-based Research on Academic Writing: Effects of discipline, register, and writer expertise*. Amsterdam: John Benjamins Publishing Company. p. 206-226. 2020. <https://doi.org/10.1075/scl.95>
- DUTRA, D. P.; BOCORNY, A. E. P.; COSTA, D. D.; TEIXEIRA, G. L. MARQUES, C. G. F. Desafios metodológicos na compilação do *corpus* de textos acadêmicos das Ciências Agrárias. In: *Anais eletrônicos do XVI Encontro de Linguística de Corpus e da XII Escola Brasileira de Linguística Computacional*. 21-24 de outubro de 2024. *Anais...* Brasília-DF, 2024.
- DUTRA, D. P.; BERBER SARDINHA, T.; DELFINO, C. N.; BOCORNY, A. E. P.; SHIMAZUMI, M.; KAUFFMANN, C.; MACEDO, L. D.; TABORDA, A. C. Key Feature Analysis of Academic Writing in the Humanities. In: *Semana de Eventos da Faculdade de Letras (SEvFALE) 2024*, Universidade Federal de Minas Gerais, Belo Horizonte, 2024.
- HALLIDAY, M. A. K. HASAN, R. *Cohesion in English*. London: Longman, 1976.
- HYLAND, K. Humble servants of the discipline? Self-mention in research articles. *English for Specific Purposes*, 2001, v. 20, n. 3, 2001, p. 207-226. DOI: [https://doi.org/10.1016/S0889-4906\(00\)00012-0](https://doi.org/10.1016/S0889-4906(00)00012-0).
- LUBIS, M. Y. A.; MIRANTI, R.; LUBIS, Y. Passive Voice and Active Voice in Sentence Structure. *Journal of Psychology, Counseling and Education*, 2024, v. 2, n. 1, p. 59-64. DOI: <https://doi.org/10.58355/psy.v2i1.20>.
- MARQUES, C. G. F. *CorAgrarian: um corpus de artigos científicos de ciências agrárias*. In: *Semana de Eventos da Faculdade de Letras (SEvFALE)*, Universidade Federal de Minas Gerais, Belo Horizonte, 2024.
- PIRES, A. C. Ensino de voz ativa e voz passiva: estrutura, efeitos de sentido e variação. Rio de Janeiro, 2024. *Dissertação de Mestrado apresentada ao programa de Mestrado Profissional em Letras – PROFLETRAS*, da Faculdade de Letras da Universidade Federal do Rio de Janeiro – UFRJ, 2024.
- SHI, H.; WANNARUK, A. Rhetorical Structure of Research Articles in Agricultural Science. *English Language Teaching*, 2014, v. 7, n. 8, p. 1-13.
- SWALES, J. M. *Genre analysis: English in academic and research settings*. Cambridge: Cambridge University Press, 1990.

TEIXEIRA, G.; BOCORNY, A. E. P.; DUTRA, D. P. Mapping the language needs of the Brazilian researchers in the Agrarian Sciences: a corpus-based study of abstracts. In: AMERICAN ASSOCIATION FOR CORPUS LINGUISTICS CONFERENCE (AACL), 15., 2022, Flagstaff. *Anais [...]*. Flagstaff: Northern Arizona University, 2022.

WANG, S.; TSENG, W.-T.; JOHANSON, R. To We or Not to We: *Corpus*-Based Research on First-Person Pronoun Use in Abstracts and Conclusions. *SAGE Open*, 2021, v. 11, n. 2. <https://doi.org/10.1177/21582440211008893>.

Data de submissão: 02/10/2025

Data de aceite: 27/04/2026

Editora: Roberta Pires de Oliveira