



O sistema da transitividade no vocabulário acadêmico de artigos científicos em Português Brasileiro

Monique Vieira Miranda* e Ana Larissa Adorno Marciotto Oliveira

Universidade Federal de Minas Gerais, Av. Presidente Antônio Carlos, 6627, 31270-901, Belo Horizonte, Minas Gerais, Brasil. *Autor para correspondência. E-mail: nk.miranda@gmail.com

RESUMO. O presente artigo se propôs a analisar o sistema da TRANSITIVIDADE instanciado pelos verbos acadêmicos mais frequentes em artigos científicos do português brasileiro, baseando-se na proposta de divisão do vocabulário de Webb e Nation (2017) e na Linguística Sistêmico-Funcional (Halliday & Matthiessen, 2014). Para atingir tal propósito, compilou-se um *corpus* de 12,3 milhões de palavras, contendo artigos científicos de todas as oito Grandes Áreas do conhecimento do CNPq, representativos das diferentes Áreas de Avaliação, de modo a haver variedade de amostras linguísticas suficientes para identificar os lemas verbais acadêmicos. A partir desses resultados preliminares, seguiu-se a análise desses itens sob a perspectiva da Linguística Sistêmico-Funcional. Nessa análise, observou-se que as escolhas dos tipos de processos, associados à configuração da sentença e aos tipos de participantes integrantes, refletem diretamente as características do gênero acadêmico, especificamente a escrita de artigos científicos, principalmente quanto à redação impessoal, ao maior foco nas atividades e nos procedimentos de pesquisa e à objetividade ao reportar análises e resultados. Desse modo, por meio da descrição do sistema da transitividade em verbos acadêmicos frequentes, esperou-se contribuir para uma compreensão mais aprofundada sobre a escrita acadêmica, além de avançar estudos teóricos sob diferentes perspectivas de análise com base em amostras reais da língua em uso.

Palavras-chave: vocabulário acadêmico; artigos científicos; sistema da transitividade.

The transitivity system in the academic vocabulary of research articles in Brazilian Portuguese

ABSTRACT. This article analyses the transitivity system instantiated by the most frequent academic verbs in research papers in Brazilian Portuguese, based on the proposed division of vocabulary by Webb and Nation (2017) and Systemic-Functional Linguistics (Halliday & Matthiessen, 2014). To achieve this purpose, a *corpus* of 12.3 million words was compiled, containing research papers from all eight Major Areas of knowledge, representing the different disciplines according to CNPq, so that there was sufficient variety of linguistic evidence to identify academic vocabulary. From preliminary results, these items were analysed from the perspective of Systemic-Functional Linguistics, in which we observed that the choices of types of processes, the configuration of the sentence and the participants, directly reflect the characteristics of the academic genre, specifically writing of research articles, such as impersonal writing, focus on research activities and procedures, and objectivity when reporting results. Thus, through the description of the transitivity system in frequent academic verbs, we expect to contribute to a deeper understanding of academic writing, in addition to advancing theoretical studies under different perspectives of analysis, based on real expositions of the language in use.

Keywords: academic vocabulary; research papers, transitivity system.

Received on March 29, 2023.

Accepted on March 19, 2024.

Introdução

Há uma ampla literatura em língua inglesa voltada ao estudo do vocabulário acadêmico, dividindo-se em duas grandes vertentes de pesquisa. Uma defende a perspectiva em que existe um vocabulário acadêmico único e comum às diferentes áreas do conhecimento, sendo, portanto, relevante o desenvolvimento de listas de palavras acadêmicas, assim como o ensino-aprendizagem por meio desse recurso (Xue & Nation, 1984; Coxhead, 2000; Nation, 2001; Gardner & Davies, 2014). Do outro lado, teóricos debatem a relevância dessa

classificação, argumentando que parâmetros como frequência, alcance, significado e colocações de uma palavra mudam ao longo das diferentes áreas de conhecimento, em função das variações de uso em cada disciplina, tornando ineficazes as listas de palavras acadêmicas (Martínez, 2001; Hyland & Tse, 2007; Durrant, 2016). Nesta vertente, a lista de palavras é vista como um recurso ideal desde que desenvolvida para uma determinada área de conhecimento específica.

Estudos mais recentes (Hyland, 2016; Dang, Coxhead, & Webb, 2017) questionam tal dicotomia no estudo do vocabulário, priorizando a importância de interpretá-lo como um contínuo. Assim, listas de palavras deveriam considerar não apenas o campo de estudo com base na qual foram construídas, mas também as circunstâncias das habilidades pré-existentes dos aprendizes e seu conhecimento e envolvimento prévio com a área de conhecimento e a língua em questão.

Até o momento, trabalhos que exploram essa temática de vocabulário acadêmico em língua portuguesa são escassos (Santos, 2006), especialmente aqueles focados em explorá-lo por meio de procedimentos metodológicos mais rigorosos (conforme feito em língua inglesa por Gardner & Davies, 2014) ou que verifique se os termos considerados acadêmicos são de fato centrais às diferentes áreas do conhecimento ou específicos.

Diante desse panorama, a presente pesquisa propôs a suprir algumas das lacunas de ambas as vertentes. Para atingir esse objetivo, foi compilado um *corpus* de artigos científicos representante de todas as áreas do conhecimento, com um total aproximado de 12,3 milhões de palavras, a fim de analisar os verbos acadêmicos mais frequentes. Neste ponto da pesquisa, de modo também a ampliar os estudos da área e avançar as pesquisas feitas em PB, empregou-se a perspectiva teórico-metodológica da Linguística Sistêmico-Funcional (LSF), especificamente o sistema de transitividade (Halliday & Matthiessen, 2014). Com base nesta perspectiva, ao analisar como tal sistema foi instanciado pelos verbos acadêmicos mais frequentes nas diferentes disciplinas, pretendeu-se verificar se seus usos são similares, corroborando com a vertente que defende a existência de um vocabulário acadêmico central, ou se os usos são particulares em cada disciplina, o que justificaria a existência de um vocabulário específico e/ou técnico em cada área.

No âmbito da LSF, o sistema de transitividade corresponde a uma forma de o usuário da língua representar e organizar suas ideias por meio da linguagem. Esse sistema relaciona-se à organização sentencial, tendo como elementos ‘processos’ (verbos), ‘participantes’ e ‘circunstâncias’. Centrais à oração, os processos, alvos desta pesquisa, estabelecem relações que codificam ações e eventos, dividindo-se em ‘materiais’, ‘mentais’, ‘relacionais’, ‘verbais’, ‘comportamentais’ e ‘existenciais’.

Com base no breve contexto exposto nesta seção, no qual se insere e justifica-se a presente pesquisa, as próximas duas seções apresentam sucintamente os conceitos teóricos utilizados nesta pesquisa. Em seguida, descrevem-se os procedimentos metodológicos deste trabalho e, na seção posterior, os dados analisados junto à discussão do tema. Por último, a Conclusão retoma o caminho percorrido para responder as questões levantadas na Introdução.

Os tipos de vocabulário na língua

Com relação ao ensino-aprendizagem de uma língua estrangeira, é relevante perguntar-se de que tipo de vocabulário precisam os alunos. Segundo os estudos de *English for Specific Purposes* (ESP), acredita-se que diferentes vocabulários atendem a necessidades específicas dos alunos (Coxhead, 2013). Assim, aprendizes interessados em comunicar-se em um contexto profissional precisariam dominar um determinado tipo de vocabulário que um aprendiz com fins turísticos, por exemplo, não necessitaria.

Segundo Coxhead (2013), o ensino do vocabulário específico de uma área do conhecimento é importante para a aprendizagem de uma língua, pois tem-se a certeza de que o conteúdo administrado dentro da sala de aula será proveitoso para o aprendiz. Desse modo, parte-se do princípio de que o conhecimento de uma área está diretamente ligado à linguagem daquela disciplina, sendo seu domínio um indicador de pertencimento àquele grupo, pois é por meio da língua que se demonstra o conhecimento sobre o conteúdo (Woodward-Kron, 2008).

Partindo-se dessa premissa, e com a finalidade de facilitar estudos posteriores voltados para a descrição da língua e também para o ensino, Nation (2001) categorizou o vocabulário de uma língua em quatro grupos de palavras: as ‘palavras de alta frequência’ (linguagem básica) e as ‘palavras de baixa frequência’, que são as mais utilizadas em geral; e os termos pertencentes ao ‘vocabulário técnico’ e ao ‘vocabulário acadêmico’, que são frequentes em contextos específicos.

As palavras de alta frequência são as mais prováveis de serem utilizadas em situações comunicativas cotidianas e, portanto, tendem a ter maior valor no ensino-aprendizagem, em comparação com as menos

frequentes. Assim, ao lado dos propósitos do aprendiz, a frequência também é um dos critérios que podem ser utilizados para identificar palavras relevantes para a construção do vocabulário do aluno. De acordo com o gênero, tais palavras cobrem cerca de 71,5 a 89,6% de um texto (Coxhead & Hirsh, 2007).

Para a identificação desses itens, listas de palavras tornaram-se recursos valiosos. Dentre as mais conhecidas estão a *General Service List*, a *New General Service List*, a lista do *British National Corpus* e a do *Corpus of Contemporary American English*. Em geral, palavras de alta frequência são incorporadas ao vocabulário do aprendiz ao se deparar com ela no cotidiano; entretanto, seu ensino deliberado por ser útil, uma vez que nem sempre pode-se estar imerso na cultura e na língua em estudo. Em PB, apesar de não haver lista específica criada e publicada, é possível sua realização com base em dois grandes *corpora* disponíveis: o *PtTenTen11*, na versão PB, que conta com 3,6 bilhões de *tokens* coletados de textos da internet; e o 'Corpus Brasileiro' (Berber-Sardinha, 2014), formado por textos representantes de diferentes gêneros da língua escrita, totalizando 1 bilhão de palavras.

Por sua vez, as palavras de baixa frequência correspondem a todas as demais da língua, podendo ser divididas em 'média' e 'baixa frequência', segundo Schmitt e Schmitt (2014). Para os autores, as palavras de média frequência bastam para a compreensão de interações escritas e faladas. Já as palavras de baixa frequência são usadas esporadicamente na língua e, geralmente, em função de contextos específicos.

Quanto às palavras técnicas, esses termos se distinguem por serem relativamente pouco utilizados na linguagem cotidiana, mas muito frequentes em uma determinada área do conhecimento. Tais palavras ocupam um lugar privilegiado em certas áreas, sendo seu conhecimento fundamental para os usuários da língua inseridos naquele contexto. Vale ressaltar que o vocabulário técnico é aprendido concomitantemente com o conhecimento sobre o tema (para usuários de língua materna ou estrangeira) e, assim, costuma-se haver explicações, glossários e outros recursos de aprendizagem relacionados a tais termos, uma vez que seus significados se restringem a tal área (Nation, 2001). Nem sempre esses termos são transparentes, isto é, exclusivos de uma disciplina, podendo ocorrer em outras áreas, ou na linguagem cotidiana, com significado similar ou distinto nos diferentes contextos de uso.

Por fim, em relação às palavras acadêmicas, essas referem-se àquelas usadas com frequência em textos acadêmicos de diferentes disciplinas, podendo ser infrequentes em textos que não são do meio acadêmico. Segundo Coxhead (2013) e Nation (2001), enquanto as palavras técnicas possuem significado relevante para a compreensão de um determinado tópico, as acadêmicas são utilizadas a fim de auxiliar o seu uso. Embora possam ocasionalmente se relacionar a uma área (ou conjunto de áreas), os termos acadêmicos contribuem para a construção do trabalho dos pesquisadores/autores. Isto é, são palavras que permitem referenciar outras pesquisas, assim como trabalhar com dados de forma científica (Nation, 2001).

Destarte, o estudo desse vocabulário se justifica por ser comum a uma gama de áreas do conhecimento, nas quais ocorre com frequência, e nem tão comum em textos não-acadêmicos. Nesse âmbito, voltando-se a uma perspectiva de ensino e aprendizagem, as listas de palavras retornam como um recurso que facilita esse processo. Em língua inglesa estão disponíveis a *Academic Word List* (AWL) (Coxhead, 2000), a *Academic Vocabulary List* (AVL) (Gardner & Davies, 2014). Em PB, dois trabalhos procuram descrever esse vocabulário em maior e menor grau: Santos (2006) e XXX, a cuja tese este artigo está associado.

Pode-se ressaltar que essas listas são ainda mais úteis quando levam em consideração possíveis combinações de palavras, uma vez que conhecê-las (e dominá-las) contribui para tornar o discurso mais natural, empregando palavras e termos conforme são utilizados na língua. Baker (1988) afirma que maior ênfase deve ser dada ao ensino desse vocabulário, pois possui função retórica e organizacional no gênero. Além disso, ele não é tão conhecido ou explicitamente ensinado como o técnico. Concomitantemente, uma lista de palavras acadêmicas pode ser ensinada mais facilmente pelo professor de línguas, pois atende a uma gama maior de alunos e não há necessidade de conhecimento prévio especializado para seu ensino.

Entretanto, vale reconhecer algumas das limitações do vocabulário acadêmico. Entre elas, está o fato de que pode ser mais difícil generalizar seus significados ao longo das diferentes disciplinas. Durrant (2016) critica essa abordagem ao investigar a diferença no uso de palavras acadêmicas entre diferentes áreas do conhecimento. O autor observou que a variação existente poderia ser diminuída ao agrupar determinadas áreas, uma vez que seus resultados apontaram para uma homogeneidade inerente às áreas. Assim, para o autor, é necessário levar em consideração as áreas (Grandes Áreas, no caso dessa pesquisa) no ensino de gêneros acadêmicos.

Hyland e Tse (2007) criticam os procedimentos metodológicos na elaboração de listas (e na identificação do vocabulário acadêmico), afirmando que normalmente não são levadas em consideração as diferenças no alcance das palavras ao longo das disciplinas, nem o favorecimento de determinadas formas e significados em cada uma delas, tampouco as colocações específicas de uma área ou um conjunto de disciplinas.

Retomando as questões levantadas nesta seção, conhecer os itens lexicais de alta frequência tende a ser mais útil do que os de baixa, uma vez que aqueles correspondem a um número pequeno de itens altamente frequentes. Caso o aprendiz tenha fins específicos, analisar as palavras frequentes dentro do discurso da área contribui para sua aprendizagem. No contexto científico, dominar o vocabulário acadêmico e técnico da disciplina em que se insere é fundamental para o sucesso em interações neste contexto. Assim, para auxiliar o ensino dos diferentes tipos de vocabulário, as listas de palavras são uma interessante estratégia, embora sejam voltadas em grande parte para a descrição da língua inglesa até o momento.

O presente artigo procurou avançar nesse quadro teórico, fornecendo uma lista de verbos acadêmicos em artigos científicos, frequentes em diferentes áreas do conhecimento, seguindo critérios inéditos em PB. Para além disso, analisou-se a representação de uma amostra do sistema da transitividade nos seus empregos mais frequentes, relacionando-os às particularidades de cada Grande Área do conhecimento. Assim, a próxima seção apresenta brevemente a LSF, com foco no sistema em questão.

O sistema da transitividade sob a perspectiva da LSF

A LSF compreende a linguagem como um sistema multidimensional semiótico, na qual a organização reflete as descrições das línguas, com potencial para construção de significados em uma abordagem semiótica holística, centrando-se no contexto e na proeminência da organização sistêmica (Halliday & Matthiessen, 2014). Ou seja, a linguagem é construída ao redor de diferentes dimensões semióticas, interseccionadas entre si e com a língua em contexto. Nessa perspectiva, a linguagem é vista como um recurso de fazer significados que residem em padrões sistêmicos de escolha. Desse modo, a escolha de cada sistema contribui para a formação da estrutura, compreendida como ‘realizações’. Assim, na análise de um texto, como um artigo científico neste trabalho, procura-se demonstrar a organização funcional da estrutura e quais as escolhas significativas realizadas, cada uma em relação ao contexto do que poderia ter sido.

Segundo Halliday e Matthiessen (2014), determinadas configurações linguísticas são tipicamente empregadas em certos contextos, podendo variar conforme o uso. Tal característica diz respeito ao ‘registro’, que corresponde a padrões de instanciação do sistema associados a um tipo de contexto – tipo de situação – que aparecem quantitativamente como ajustes no sistema probabilístico da língua. Isto é, a linguagem materializa as escolhas do usuário de acordo com o conjunto de variáveis que condicionam a comunicação. Essas variáveis correspondem a dois níveis de organização contextual: de cultura e de situação.

O ‘contexto de cultura’ relaciona-se ao contexto sociocultural mais amplo, que inclui ideologia, convenções sociais e instituições. Enquadram-se aqui não apenas a linguagem, mas também paralinguagem, dança, desenhos e arquitetura – o potencial cultural de uma comunidade. Além disso, o contexto de cultura associa-se à noção de propósito social, desenvolvendo tipos comuns de textos – gêneros – que alcançam objetivos similares.

Por sua vez, o ‘contexto de situação’ apresenta variáveis que se constituem ao redor mais imediato em que o texto se insere. Tais variáveis são ‘campo’, ‘relação’ e ‘modo’, que podem caracterizar qualquer situação, tendo impacto direto e significativo na linguagem utilizada (Eggins, 2004). O campo diz respeito à atividade que acontece na situação, reconhecida na cultura pelos participantes, e à natureza dessa ação social. A relação corresponde aos participantes da situação e à natureza dos papéis sociais desempenhados por eles, assim como o grau de controle entre eles, a possível hierarquia e o nível de formalidade. O modo remete à função exercida pela linguagem, ao veículo comunicativo e ao que se espera alcançar por meio dela (Eggins, 2004).

Conforme Halliday e Matthiessen (2014), essas três variáveis se relacionam com os três tipos de significados estruturados pela língua, exercendo funções diferentes, e que podem ser associados ‘para cima’ – com o contexto – ou ‘para baixo’ com a léxico-gramática. Na relação para cima, está o tipo de registro relacionado a cada um desses significados: ‘campo’ pode ser expresso por meio de padrões ideacionais no texto, ‘modo’ pode ser expresso através de significados textuais e a ‘relação’ através de significados interpessoais. Já na associação para baixo, a ligação está nos tipos de significados sendo realizados por meio da combinação de padrões léxico-gramaticais.

Por conseguinte, cada significado relaciona-se de forma previsível a uma variável do contexto situacional, correspondendo a um significado da linguagem: ideacional, textual e interpessoal. O campo ressoa com

significado ideacional, o modo com o significado textual e a relação com o significado interpessoal. Assim, a correspondência entre o contexto e a linguagem baseia-se na organização funcional dos significados, em que tais funções são denominadas ‘metafunções’ (Halliday & Matthiessen, 2014), instanciadas de forma entrelaçada e simultânea da linguagem.

Assim, cada sistema descrito acima possui um foco de análise léxico-gramatical distinto, de modo que uma mesma oração pode ser analisada por diferentes perspectivas, embora estejam, ao mesmo tempo, sistematicamente relacionadas. Para fins deste artigo, é relevante detalhar um pouco mais a metafunção ideacional, um dos enfoques da presente análise.

Na abordagem sistêmico-funcional, a metafunção ideacional fornece recursos léxico-gramaticais para construir a experiência sobre o mundo, interno ou externo, enquanto significado por meio da linguagem (Thompson, 2014) por meio de dois construtos: lógico e experiencial. A função experiencial compreende os processos que são instanciados, seus participantes e circunstâncias associadas, construindo um modelo de representação do mundo ao nível da oração. Já a função lógica diz respeito à relação entre um processo e outro, ou entre participantes que compartilham a mesma função no texto, sendo responsável por combinar grupos lexicais e oracionais. Desse modo, ao analisar a oração, o sistema que se emprega é o de transitividade, foco deste trabalho.

A fim de denominar as representações de experiências gramaticalizadas em figuras de acontecimentos, são utilizadas etiquetas funcionais que indicam o papel de cada elemento: ‘processos’ que envolvem ‘participantes’ em certas ‘circunstâncias’. Os processos são o núcleo da oração, uma vez que seu significado gira ao redor de uma ação, estado ou evento, sendo realizados tipicamente por um grupo verbal. Os participantes, normalmente realizados por um grupo nominal, podem ou não ser impactados/envolvidos no processo. Por fim, as circunstâncias costumam ser realizadas por grupos adverbiais ou preposicionados. Em geral, apesar de serem opcionais, esses componentes são fundamentais para a completude do significado da oração (Thompson, 2014).

De acordo com Halliday e Matthiessen (2014), os processos são construídos em um conjunto de Tipos de Processos, em que cada tipo configura um modelo distinto da construção de um determinado domínio da experiência, sendo três primários (materiais, mentais e relacionais) e três secundários (verbais, comportamentais e existenciais).

Os processos primários podem ser resumidos em: ‘materiais’, que representam a experiência externa (ações e eventos), realizados por verbos como ‘fazer’, ‘construir’ e ‘agir’; ‘mentais’, que representam a experiência interna (lembranças, reações e sentimentos), realizados por verbos como ‘pensar’, ‘imaginar’, ‘gostar’ e ‘querer’; e ‘relacionais’, que representam relações (identificação e caracterização), realizadas por verbos como ‘ser’, ‘estar’, ‘parecer’ e ‘ter’.

Nas áreas limítrofes entre os processos primários estão os secundários, cujas características podem ser resumidas em: ‘verbais’, que representam dizeres através da atividade linguística dos participantes e podem ser exemplificados por ‘dizer’, ‘afirmar’ e ‘responder’; ‘comportamentais’, que representam comportamentos, manifestando atividades psicológicas ou fisiológicas do ser humano como ‘dormir’, ‘tossir’ e ‘sonhar’; e ‘existenciais’, que representam a existência de um participante, como ‘haver’ e ‘existir’.

Cada processo possui particularidades que implicam a interação de certas entidades cuja natureza se dá em função do tipo de processo com os quais ocorrem. Assim, cada processo é executado por participantes distintos, ilustrados na Figura 1.

Por fim, este capítulo procurou descrever brevemente a abordagem da LSF sobre a linguagem, desde seus contextos mais amplos até sua realização na sentença por meio do sistema de transitividade. Delineada essa parte da revisão de literatura, a próxima seção descreve os procedimentos metodológicos empregados na construção do *corpus* e na análise dos dados.

Metodologia

Neste trabalho procurou-se observar os dados em seu contexto, analisando-os minuciosamente enquanto uma pesquisa de descrição da língua em uso. Para tanto, tomou-se como base as premissas da divisão do vocabulário e a perspectiva da LSF. Com o objetivo de analisar a realização do sistema de transitividade nos verbos relevantes em artigos acadêmicos (vocabulário acadêmico), foi compilado pelas autoras um *corpus* de aproximadamente 12,3 milhões de palavras com textos representantes das diferentes áreas de conhecimento, de modo a evitar que fosse tendencioso a uma única área e, conseqüentemente, tivesse maior proporção de vocabulário técnico.

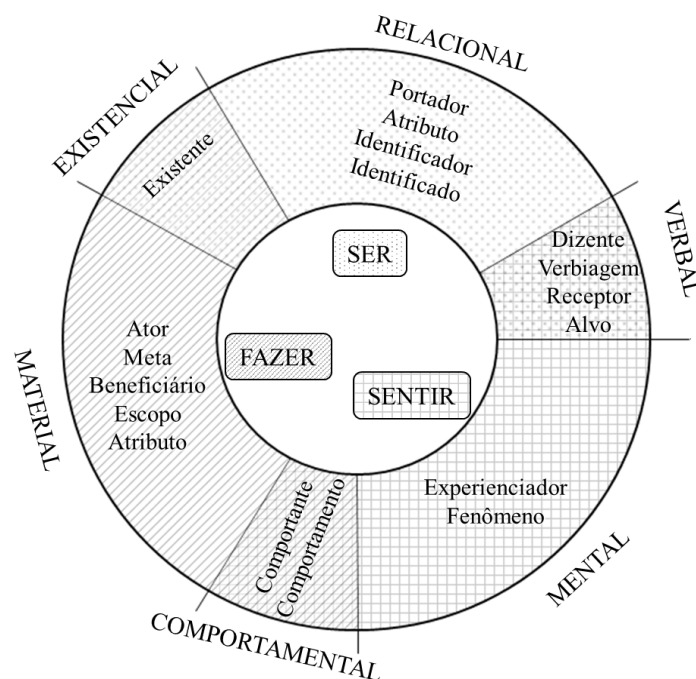


Figura 1. Tipos de processos e participantes no sistema da transitividade. Fonte: Adaptado de Fuzer e Cabral (2014).

Em relação às áreas do conhecimento, de acordo com o CNPq, as várias disciplinas estão organizadas segundo a estrutura da árvore do conhecimento, em que as Grandes Áreas são a hierarquia mais abrangente, seguidas pelas Áreas, Subáreas e, por último, pelas Especialidades. Os três primeiros níveis são mais estáticos enquanto as Especialidades costumam sofrer modificações mais radicais conforme a necessidade da disciplina. Atualmente a lista do CNPq possui 8 Grandes Áreas, 76 Áreas, 340 Subáreas e diversas Especialidades. Desse modo, justifica-se a atual organização dos textos analisados por esta pesquisa, que pretendeu compilar um *corpus* que não estivesse sujeito a muitas variações, como de nomenclatura, por exemplo. Assim, subdividiu-se o *corpus* em oito *subcorpora* que correspondem às oito Grandes Áreas: Ciências Agrárias, Ciências Biológicas, Ciências da Saúde, Ciências Exatas e da Terra, Engenharias, Ciências Humanas, Ciências Sociais Aplicadas e Linguística, Letras e Artes.

Durante a coleta, os periódicos foram escolhidos de modo que contemplassem, se não todas, a maioria das Subáreas para que os *subcorpora* não fossem tendenciosos para certo tópico. Para a seleção dos artigos, observou-se uma série de critérios tais quais estarem disponíveis no portal SciELO, serem escritos em PB e publicados posteriormente a 2010. Desse modo, procurou-se ser o mais representativo possível de todas as Subáreas, contemplando diferentes periódicos e tópicos pesquisados. Obedecendo a tais critérios e após a limpeza de cada texto, obteve-se um *corpus* com 2.898 artigos científicos, contendo aproximadamente 12,3 milhões de *tokens* – que compõe o ‘Corpus Acadêmico do Português Brasileiro’, disponível online¹.

Após a compilação e lematização do *corpus*, para a elaboração de uma lista de verbos acadêmicos em artigos científicos, o presente trabalho baseou-se em métodos anteriores de elaboração de listas acadêmicas (mencionadas na seção 2), como AWL e AVL. Para a criação da presente lista, cinco critérios foram adotados: frequência mínima, razão da frequência, alcance, dispersão e medida entre disciplinas. O primeiro estabeleceu uma ‘frequência mínima’ de 100 ocorrências para os verbos do *corpus* Geral, garantindo que os verbos resultantes sejam de alta frequência. A ‘razão’ referiu-se à etapa empregada a fim de eliminar as palavras de alta frequência da lista por meio da especificação de que um verbo deve ser pelo menos 50% (razão 1.5) mais frequente neste *corpus* do que no *corpus* de referência (PtTenTen11). O terceiro critério, a razão do ‘alcance’, determinou que uma palavra deva ocorrer com 20% da frequência esperada em seis ou mais dos oito *subcorpora*. A ‘dispersão’ garantiu que o verbo estivesse distribuído igualmente ao longo dos *subcorpora*, sendo calculado o índice Juilland’s D com valor 0.8. Por último, a ‘medida de disciplina’ procurou eliminar as palavras técnicas de uma área ao determinar que um verbo não poderia ocorrer com mais de três vezes a frequência esperada em nenhum dos *subcorpora*. Com base nesses critérios, elaborou-se uma lista dos 160 verbos acadêmicos em artigos científicos, cujas palavras mais frequentes estão representadas na Tabela 1.

¹Recuperado de <https://sites.google.com/view/corpusacademico/página-inicial>.

Tabela 1. Lista dos principais verbos acadêmicos.

Verbo	Agrárias		Biológicas		Exatas		Engenharias		Humanas		Letras		Saúde		Sociais		Geral	
	FR	FA	FR	FA	FR	FA	FR	FA	FR	FA	FR	FA	FR	FA	FR	FA	FR	FA
1 apresentar	379,6	4.520	337,0	2.370	379,5	5.807	358,4	7.783	93,1	1.992	146,8	2.112	304,5	4.976	142,5	2.465	255,4	32.025
2 utilizar	226,5	2.697	155,0	1.090	303,5	4.644	241,4	5.242	49,1	1.051	64,9	934	180,7	2.954	83,5	1.444	159,9	20.056
3 realizar	202,7	2.414	153,7	1.081	166,1	2.542	178,1	3.868	68,6	1.467	85,2	1.226	257,9	4.215	84,1	1.456	145,7	18.269
4 considerar	100,4	1.195	110,5	777	123,9	1.896	147,3	3.199	92,1	1.970	98,4	1.416	150,1	2.454	109,5	1.894	118,0	14.801
5 observar	194,1	2.311	134,1	943	146,8	2.246	107,7	2.339	39,2	839	58,0	835	108,4	1.772	51,3	887	97,1	12.172
6 encontrar	137,6	1.639	154,5	1.087	96,8	1.481	80,0	1.738	60,6	1.297	76,0	1.094	118,3	1.933	62,6	1.084	90,5	11.353
7 ocorrer	112,0	1.334	126,5	890	106,4	1.629	97,3	2.112	49,8	1.065	55,5	799	71,5	1.169	56,6	980	79,6	9.978
8 mostrar	46,7	556	47,5	334	140,0	2.143	101,7	2.209	47,8	1.022	60,9	876	95,7	1.564	49,0	848	76,2	9.552
9 analisar	74,5	887	63,1	444	67,0	1.025	77,5	1.683	42,1	901	48,6	699	78,9	1.290	56,9	985	63,1	7.914
10 determinar	80,2	955	49,3	347	75,5	1.156	69,0	1.499	50,5	1.081	47,3	680	46,3	757	64,5	1.116	60,5	7.591
11 representar	41,2	490	64,0	450	90,4	1.383	74,7	1.622	45,8	979	53,1	764	41,2	673	56,3	975	58,5	7.336

Legenda: (FR) Frequência relativa (refere-se à frequência normalizada utilizada na comparação de dois ou mais *corpora*); (FA) Frequência absoluta (valor bruto proveniente das estatísticas do *corpus*).

Posteriormente à elaboração da lista, os verbos mais frequentes foram analisados em seu contexto de uso, levando-se em consideração o contexto do texto e da Grande Área em que ocorreu. Para tanto, selecionou-se uma amostra aleatória de cinquenta linhas de concordância de cada *subcorpus* da pesquisa para cada um dos dez verbos acadêmicos mais frequentes. Assim, analisou-se quatro mil linhas de concordância focando na compreensão de como o processo foi empregado no texto, seguindo uma série de parâmetros referentes à sua instanciação, tais quais a realização do grupo verbal quanto à efetividade (sistema de agência), o tempo em que decorre a oração, a presença de modalizador, assim como as características próprias de cada tipo de processo e as circunstâncias acompanhantes.

Cabe ressaltar ainda que alguns verbos podem ser classificados como um ou mais tipos de processos, em função do seu contexto de uso. Halliday e Matthiessen (2014) discorrem sobre essa possibilidade com os verbos *suggest* e *indicate*, definindo-os como processos relacionais quando os elementos nominais na estrutura transitiva da oração são abstrações, ou como processos mentais quando o sentido do verbo utilizado é cognitivo. Como pode ser observado nos dados expostos na próxima seção, o mesmo ocorreu nos dados desta pesquisa.

Resumidos os critérios que embasaram a análise, a seção seguinte apresenta os resultados encontrados sobre a instanciação do sistema de transitividade em amostras dos verbos acadêmicos mais frequentes em artigos científicos.

Análise dos tipos de processos entre verbos acadêmicos mais frequentes

Para a presente análise, foram estudadas quatro mil amostras aleatórias dos dez verbos considerados acadêmicos mais frequentes: ‘apresentar’, ‘utilizar’, ‘realizar’, ‘considerar’, ‘observar’, ‘encontrar’, ‘ocorrer’, ‘mostrar’, ‘analisar’ e ‘determinar’. Tais verbos realizaram majoritariamente processos materiais e relacionais, seguidos por mentais e existenciais e por algumas poucas ocorrências de verbais. Conforme pode ser observado na Tabela 2, que apresenta a distribuição dos tipos de processos analisados, apenas quatro verbos (‘utilizar’, ‘ocorrer’, ‘mostrar’ e ‘analisar’) instanciaram um único tipo de processo enquanto os demais instanciaram dois ou mais processos distintos.

Tabela 2. Distribuição dos tipos de processos por verbos analisados.

	Processos materiais	Processos relacionais	Processos mentais	Processos existenciais	Processos verbais
apresentar	55	344	-	1	-
utilizar	400	-	-	-	-
realizar	389	-	-	11	-
considerar	-	311	89	-	-
observar	-	-	396	-	4
encontrar	215	185	-	-	-
ocorrer	-	-	-	400	-
mostrar	-	400	-	-	-
analisar	400	-	-	-	-
determinar	259	123	-	-	18
Total	1718	1363	485	412	22

Em âmbito geral, os resultados encontrados mostram-se convergentes com os demais estudos voltados à investigação do sistema de transitividade em diferentes gêneros acadêmicos (Mendes, 2013; Zhen, Yang, & Ge, 2014; Vathanalaoha & Tangkiengsirisin, 2018), nos quais os processos materiais, relacionais e mentais (nessa ordem) são os mais frequentes. A fim de comprovar essa mesma tendência, os dados foram analisados ao longo das diferentes Grandes Áreas, demonstrando que tal padrão também foi verificado nas mesmas proporções em cada um dos oito *subcorpora* e confirmando, desse modo, que esta é uma característica do gênero. A seguir, a Figura 2 apresenta essa tendência ao longo das Grandes Áreas do conhecimento.

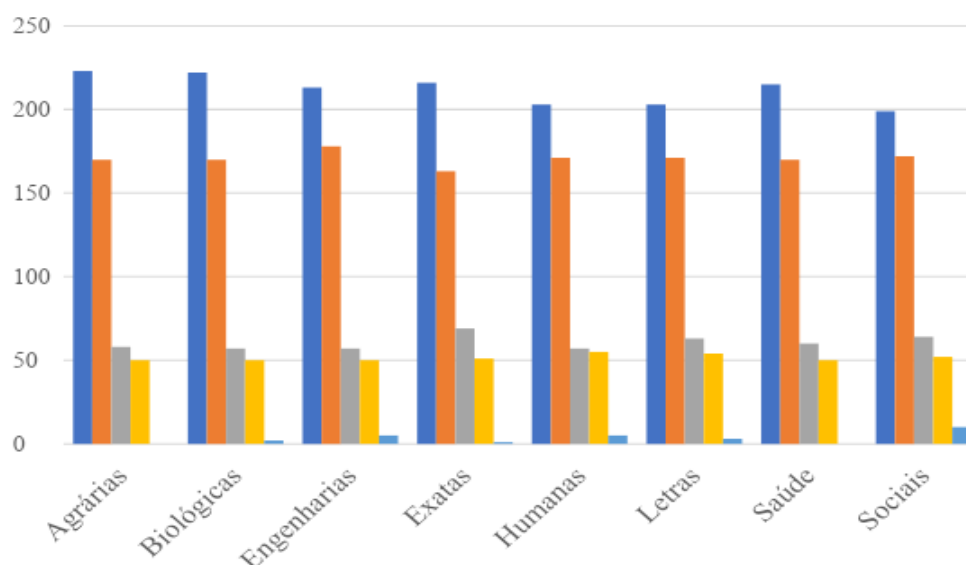


Figura 2. Contagem dos tipos de processos por Grande Área do conhecimento.

Cabe observar que além dos processos analisados estarem presentes em todos os *subcorpora*, também ocorreram com distribuição similar de frequência nas diferentes Grandes Áreas. Essa característica encontrada nos dados ajuda a corroborar a tese de que esses verbos são relevantes e constantes em todas as subáreas e, à primeira vista, seus usos estão distribuídos similarmente ao longo das disciplinas.

Dentre os tipos que ocorreram nas amostras analisadas, os mais frequentes foram os processos materiais, que expressam a noção de que alguma entidade ‘fez’ algo, incluindo eventos e acontecimentos físicos ou abstratos. As orações materiais correspondem a 43% das orações totais, sendo as mais frequentes em todos os *subcorpora*. Essa porcentagem elevada reflete uma propriedade do gênero que é o uso dos processos materiais pelos autores dos textos a fim de descrever e relatar, majoritariamente, procedimentos metodológicos, análises e contribuições suas ou de outros pesquisadores no texto.

Nos dados, os verbos ‘analisar’, ‘apresentar’, ‘determinar’, ‘encontrar’, ‘realizar’ e ‘utilizar’ instanciaram, no total, 1.694 orações materiais. Todos esses verbos realizaram ao menos um outro tipo de processo, com exceção de ‘analisar’ e ‘utilizar’, cujas ocorrências instanciariam processos materiais.

Com relação ao tempo, as orações materiais foram classificadas conforme sua realização no pretérito, presente ou futuro, sem adentrar em classificações mais detalhadas, e sua relação com a construção da oração dentro do texto. Quando analisados os resultados do *corpus* como um todo, os processos materiais foram majoritariamente empregados no pretérito, sendo frequentemente associados à descrição das ações dos pesquisadores (o próprio autor ou outros), havendo um número relevante de usos no presente e, com uma frequência menor, no futuro.

Quando esses dados são analisados filtrando-os por *subcorpora*, conforme demonstrado na Figura 3 a seguir, é possível perceber que as Grandes Áreas de Ciências Agrárias, Ciências Biológicas, Engenharias, Ciências Exatas e Ciências da Saúde tenderam a empregar os processos materiais no pretérito com frequência bastante superior aos usos nas Grandes Áreas de Ciências Humanas, Ciências Sociais e Letras, Linguística e Artes.

Quanto à realização do grupo verbal em relação a cada processo, a maior parte das orações foram efetivas receptivas (voz passiva), seguidas por efetivas operativas (voz ativa) e por orações médias, à exceção de ‘encontrar’ em que a maior parte das orações materiais foram empregadas em construções efetivas receptivas com Ator omitido da oração. Nesses casos, as omissões puderam ser facilmente recuperadas tendo como referentes ao/s próprio/s autor/es do texto ou a outros pesquisadores expressos por meio de circunstâncias ou em orações adjacentes.

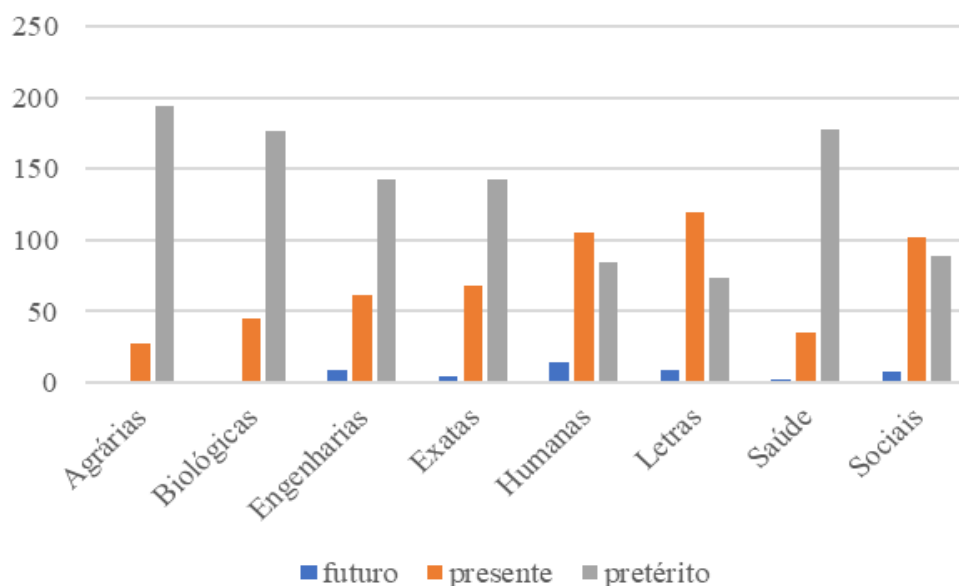


Figura 3. Tempo verbal em orações materiais por Grande Área.

Essa omissão do Ator é uma estratégia que visa a desfocar a atenção do autor do texto (quem realizou a ação/Ator) para dar destaque à mensagem/Meta, provocando um efeito de objetividade na construção textual. Assim, por meio da lacuna no preenchimento dos Atores nas orações materiais (que frequentemente fazem referência ao/s próprio/s autor/es), os artigos apresentam ao leitor a informação de modo conciso, direcionando sua atenção à Meta, ou seja, àquilo que foi realizado pelo Ator não-expresso.

Retomando a questão do modo de realização das orações materiais – como efetivas operativas, receptivas ou construções com clítico *se* –, cabe ainda observar que a finalidade de seus usos foi similar ao longo dos diferentes *subcorpora*, embora a frequência dos tipos varie conforme a Grande Área. Seguindo o mesmo padrão da Figura 3, Ciências Humanas, Ciências Sociais Aplicadas e Letras tendem a empregar os processos materiais em orações efetivas operativas, enquanto as demais Grandes Áreas tendem a optar por orações efetivas receptivas. Quando esmiuçados os participantes dessas orações, observa-se que os dados ou os objetos de estudo das áreas de Ciências Humanas, Ciências Sociais e Letras, Linguística e Artes costumam ‘realizar’ mais ações – o que é congruente com o fato de que essas disciplinas são voltadas ao estudo da sociedade como um todo e, portanto, lidam com seres humanos, tornando necessário explicitá-los no decorrer das ações realizadas em orações materiais. Esse achado sobre o tipo de Ator é coerente com os resultados sobre os tipos de construções, demonstrando que em tais Grandes Áreas a frequência de orações materiais com o Ator marcado é consideravelmente – o dobro – maior do que nas demais.

Por sua vez, os processos relacionais foram o segundo tipo mais frequente, com 1.365 ocorrências realizadas pelos verbos ‘apresentar’, ‘considerar’, ‘determinar’, ‘encontrar’ e ‘mostrar’ – destes, apenas ‘mostrar’ teve todos os usos analisados enquanto processo relacional. As orações relacionais realizam uma figura de ‘ser’ que expressa a relação entre duas coisas ou conceitos abstratos, de modo a caracterizá-los (‘atributivos’) ou identificá-los (‘identificativos’). Relacionados aos textos acadêmicos, que são textos descritivos, argumentativos e expositivos, esses processos tendem a ser altamente informativos (Halliday, 1985) e, portanto, utilizados com frequência na construção dos textos a fim de contribuir com descrições e especificidades do conteúdo.

É interessante notar a alta presença de nominalizações associadas aos processos relacionais, conforme pode ser observada no trecho (1). Segundo Halliday e Matthiessen (2014) e Martínez (2001), os processos relacionais podem renegar a possibilidade de expressar ações por meio do uso de nominalizações, de modo a diminuir a presença da agência humana – seja do próprio autor ou de outros – em relação à proposição na oração. Por conseguinte, ao empregar tais processos, os autores renunciam expressar sua participação na oração, mudando o foco para a relação entre as duas entidades e apresentando informações como contribuições factuais, por serem semanticamente impessoais e por mostrarem apenas a relação entre dois conceitos que não envolvem a presença de participantes agentivos em sua estrutura transitiva. Essa ausência de Agente torna a informação apresentada mais objetiva, impessoal e factual, contribuindo para deixar os textos mais precisos e as descrições materiais, menos mecânicas.

(1) A comparação no Ciclo II ‘mostrou’ desempenho inferior do GP II em relação ao GC II em todas as variáveis estudadas. (Sa_CODA_2014_0_AA_004)

Em geral, ao longo dos diferentes *subcorpora*, os usos das orações relacionais associaram-se majoritariamente a construções efetivas operativas no presente, com intuito de introduzir conceitos ou resultados da pesquisa, caracterizando-os ou identificando-os, de modo descritivo e preciso, e contribuindo para a exposição de informações, teorias e descobertas (próprias ou de outros) de modo impessoal. Além disso, também foram empregadas a fim de apresentar partes do trabalho, como exemplos, tabelas e figuras.

Quanto aos processos mentais, esses foram os terceiros mais frequentes, com 485 ocorrências, instanciados pelos verbos ‘considerar’ e ‘observar’ – ambos também realizaram processos relacionais e verbais, respectivamente. As orações mentais consistem em um Experienciador consciente (ou ao qual é atribuído consciência) e um Fenômeno, que pode também projetar orações. Por meio dessas orações, os autores tenderam a reduzir sua participação no texto, omitindo-se em construções efetivas receptivas ou construções médias, como exemplificado em (2) e (3).

(2) Segundo Cecílio et al. (2009), a aplicação da Krigagem necessita de uma distribuição mais ou menos uniforme das estações pluviométricas (pontos amostrais), fato que não ‘foi observado’ neste trabalho, o que pode indicar os altos valores de IDE e a baixa dependência espacial entre os coeficientes interpolados. (B_RAA_2018_1_AA_002)

(3) Para que não haja influência das decisões de longo prazo nas análises, ‘considera-se’ que a energia proveniente de leilões de longo prazo (A-5 e A-3) será recontratada, caso esses contratos terminem dentro do período de análise considerado. (E_SBA_2012_0_AA_001)

Entretanto, apesar dessas ocorrências, os processos mentais foram majoritariamente empregados nas Grandes Áreas com o objetivo de trazer as ideias de outros pesquisadores para o texto, contribuindo para a construção da resenha de literatura. No geral, as orações mentais forneceram um espaço para que fossem inseridas análises e observações dos autores (do texto ou outros pesquisadores), sendo seu modo de realização uma forma de priorizar a mensagem exposta, por meio do Fenômeno/oração projetada, em vez de o Experienciador. Quando relacionado ao próprio autor, optou-se por construções na voz efetiva receptiva ou operativas com clítico ‘se’, omitindo o Experienciador do processo mental. Tais estratégias afastam o autor de suas observações, em uma tentativa de tornar o texto objetivo e impessoal – mesmo ao lidar com questões subjetivas, como a interpretação dos dados.

No que concerne as orações existenciais, foram encontrados 412 usos instanciados pelos verbos ‘ocorrer’, ‘apresentar’ e ‘realizar’, dos quais apenas o primeiro foi empregado exclusivamente. Os processos existenciais se caracterizam por expressarem a existência de um único participante, o Existente. Enquanto os demais preconizam a relação entre dois ou mais participantes e/ou circunstâncias, os processos existenciais reduzem a agência humana ao evitar a relação entre entidades. Por meio de seu uso, o Existente pode ser apresentado como participante único, que ocorre independente da agência humana – ou do autor do texto, nos casos analisados. Essa propriedade acentua o caráter de objetividade atribuído ao texto por seu uso, em especial quando associado aos resultados da pesquisa. Além disso, contribui para que o foco do leitor permaneça na informação expressa pelo Existente

Quanto à realização do grupo verbal, a preferência foi pelo uso do tempo presente. Em relação às ocorrências existenciais de ‘apresentar’ e ‘realizar’, essas associaram-se frequentemente ao emprego da partícula ‘se’. Por outro lado, as orações existenciais instanciadas pelo verbo ‘ocorrer’ foram majoritariamente (98%) efetivas operativas. Entretanto, ao analisar esses dados por *subcorpora*, pode-se perceber a preferência de Ciências Agrárias por empregar tais processos no tempo pretérito. Por meio da análise minuciosa das ocorrências, notou-se que isso aconteceu devido à própria função do processo nas orações. Nessa Grande Área, ‘ocorrer’ foi amplamente utilizado para apresentar resultados/dados encontrados, atribuindo a eles uma noção de completude.

Com relação às orações verbais, foram as menos frequentes nos dados analisados, com poucas ocorrências sendo instanciadas pelos verbos ‘determinar’ e ‘observar’, totalizando 22 orações. Em geral, os processos verbais podem envolver os participantes Dizente, Receptor, Alvo e Verbiagem, além da possibilidade de projetarem orações. Nos dados da pesquisa, quanto aos participantes, todas as ocorrências de orações verbais tinham o Dizente exposto, referindo-se majoritariamente a outros estudos. Nesses casos, trazer a voz de outros pesquisadores para o texto contribuiu para aumentar a legitimidade das teorias selecionadas para a pesquisa, assim como ajuda a validar as observações e os posicionamentos expressos pelo autor, atribuindo um traço confiável e persuasivo ao texto. Além de trazer para o texto a voz de outras pesquisas, os processos

verbais também foram empregados a fim de relatar regulamentações de procedimentos, como normas governamentais e decretos.

Voltando a uma perspectiva mais ampla de análise, foi possível observar que as ocorrências de um mesmo tipo de processo tenderam a ser empregadas de maneira similar – tanto entre si quanto ao longo dos diferentes *subcorpora*. Também se observou que, apesar da função dentro do texto acadêmico ser parecida, o modo de realização das orações foi diferente, de acordo com a Grande Área em que ocorreu. Por exemplo, os processos materiais foram mais frequentemente empregados em orações efetivas receptivas no pretérito nos *subcorpora* de Ciências Agrárias, Ciências Biológicas, Ciências Exatas, Engenharias e Ciências da Saúde, enquanto nos demais – Ciências Humanas, Ciências Sociais e Letras, Linguística e Artes – a escolha preferida foram as orações efetivas operativas no tempo presente.

Por meio da análise detalhada de cada ocorrência, pode-se notar que tais diferenças são motivadas, em parte, pela natureza dos estudos de cada área. Apesar desse elemento distintivo, os processos foram majoritariamente empregados com a mesma função ao longo dos textos. Tais resultados justificam e verificam a validade da lista de verbos acadêmicos apresentada introdutoriamente neste artigo. Apesar de o modo de realização não ser o mesmo, a finalidade com que foram empregados os processos foi constante ao longo dos *subcorpora* (guardadas as devidas limitações de tipo). Pode-se, enfim, comprovar que a lista é de fato relevante para os estudos do gênero artigo científico, assim como para a construção das etapas e procedimentos envolvidos nesse tipo de texto. Por fim, a próxima seção apresenta as considerações finais da pesquisa, a fim de responder as perguntas e objetivos propostos na Introdução.

Conclusão

A presente pesquisa teve como objetivo verificar a existência de um vocabulário acadêmico em artigos científicos de diferentes áreas do conhecimento, por meio de uma análise minuciosa de suas ocorrências seguindo a perspectiva da LSF. Aliado a esse propósito, procurou-se observar também a validade e a relevância de uma lista com esse vocabulário para os estudos com fins específicos de uma determinada área de conhecimento.

Conforme exposto na seção 4, apesar de pontuais diferenças entre determinados usos ao longo dos *subcorpora*, particularmente relacionadas ao modo como a oração foi construída, a lista acadêmica criada mostrou-se relevante uma vez que foi possível perceber generalizações sobre os significados, tais quais os grupos verbais que realizaram tais processos, a regularidade na realização das escolhas sistêmicas, como/quais os participantes acompanharam os processos, assim como as consequências de agnação e pressuposição a respeito do léxico empregado.

Contudo, percebeu-se a necessidade de um estudo mais aprofundado e com fins específicos de cada disciplina, conforme sugerem os trabalhos de Martínez (2001), Hyland e Tse (2007) e Durrant (2016), a fim de verificar outras possíveis preferências de cada área que ultrapassem o limite da análise do verbo. Porém, para fins de ensino desse gênero acadêmico, a presente lista se torna útil e válida, especialmente se aliada à perspectiva da LSF, podendo, também, fornecer bases teóricas para aperfeiçoar a descrição de artigos científicos, indo além dos objetivos propostos e realizados neste estudo.

Por fim, este estudo procurou contribuir para um campo de pesquisa centrado no papel do léxico e da gramática na construção do texto, particularmente no que diz respeito à influência mútua entre textualidade e situação comunicativa no domínio acadêmico.

Referências

- Baker, M. (1988). Sub-technical vocabulary and the ESP teacher: an analysis of some rhetorical items in medical journal articles. *Reading in a Foreign Language*, 4(2), 91-105.
- Berber-Sardinha, T. (2014). Looking at collocations in Brazilian Portuguese through the Brazilian Corpus. In T. B. Sardinha, & T. L. S. B. Ferreira (Eds.), *Working with Portuguese corpora* (p. 9-32). Londres, GB: Bloomsbury Academic.
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34(2), 213-238.
DOI: <https://doi.org/10.2307/3587951>
- Coxhead, A. (2013). Vocabulary and ESP. In B. Paltridge, & S. Starfield (Eds.), *The Handbook of English for Specific Purposes* (p. 115-132). Boston, MA: Wiley-Blackwell.

- Coxhead, A., & Hirsh, D. (2007). A pilot science word list for EAP. *Revue Française de Linguistique Appliquée*, XII(2), 65-78.
- Dang, T. N. Y., Coxhead, A., & Webb, S. (2017). The academic spoken word list. *Language Learning*, 67(4), 959-997. DOI: <https://doi.org/10.1111/lang.12253>
- Durrant, P. (2016). To what extent is the Academic Vocabulary List relevant to university student writing? *English for Specific Purposes*, 43, 49-61. DOI: <https://doi.org/10.1016/j.esp.2016.01.004>
- Egins, S. (2004). *An introduction to systemic functional linguistics* (2nd ed.). Nova York, NY: Continuum.
- Fuzer, C., & Cabral, S. R. S. (2014). *Introdução à gramática sistêmico-funcional em língua portuguesa*. Campinas, SP: Mercado de Letras.
- Gardner, D., & Davies, M. (2014). A new academic vocabulary list. *Applied Linguistics*, 35(3), 305-327. DOI: <https://doi.org/10.1093/applin/amt015>
- Halliday, M. A. K. (1985). *An introduction to functional grammar*. London: Edward Arnold.
- Halliday, M. A. K., & Matthiessen, C. M. I. M. (2014). *An introduction to functional grammar* (3a ed., rev.). London, GB: Edward Arnold.
- Hyland, K. (2016). General and specific EAP. In K. Hyland, & P. Shaw (Eds.), *The Routledge handbook of English for academic purposes* (p. 17-29). London, GB: Routledge.
- Hyland, K., & Tse, P. (2007). Is there an “academic vocabulary”? *TESOL Quarterly*, 41(2), 235-253.
- Martínez, I. A. (2001). Impersonality in the research article as revealed by analysis of the transitivity structure. *English for Specific Purposes*, 20(3), 227-247. DOI: [https://doi.org/10.1016/S0889-4906\(00\)00013-2](https://doi.org/10.1016/S0889-4906(00)00013-2)
- Mendes, J. M. O. (2013). *Processos verbais em artigos científicos da administração de empresas e engenharia: perspectivas sistêmico-funcionais* (Dissertação de Mestrado). Pontifícia Universidade Católica de São Paulo, São Paulo.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge, GB: Cambridge University Press.
- Santos, V. (2006). *Vocabulário e leitura: a elaboração de uma lista de palavras de uso acadêmico em Português do Brasil* (Dissertação de Mestrado). Universidade Estadual de Campinas, Campinas.
- Schmitt, N., & Schmitt, D. (2014). A reassessment of frequency and vocabulary size in L2 vocabulary teaching. *Language Teaching*, 4(47), 484-503.
- Thompson, G. (2014). *Introducing functional grammar* (3rd ed.). Abingdon, GB: Routledge.
- Vathanalaotha, K., & Tangkiengsirisin, S. (2018). Genre analysis of experiment-based dental research article abstracts: thai and international journals. *3L: Language, Linguistics, Literature*, 24(3), 1-14. DOI: <https://doi.org/10.17576/3L-2018-2403-01>
- Webb, S., & Nation, P. (2017). *How vocabulary is learned*. Oxford, GB: Oxford University Press.
- Woodward-Kron, R. (2008). More than just jargon – the nature and role of specialist language in learning disciplinary knowledge. *Journal of English for Academic Purposes*, 7(4), 234-249. DOI: <https://doi.org/10.1016/j.jeap.2008.10.004>
- Xue, G., & Nation, I. S. P. (1984). A university word list. *Language Learning and Communication*, 3(2), 215-229.
- Zheng, S., Yang, A., & Ge, G. (2014). Functional stylistic analysis: transitivity in English-medium medical research articles. *International Journal of English Linguistics*, 4(2), 12-25. DOI: <https://doi.org/10.5539/ijel.v4n2p12>