



UNIVERSIDADE ESTADUAL PAULISTA
“JÚLIO DE MESQUITA FILHO”
Campus de São José do Rio Preto

Larissa Souza da Silva

**AUTOCITAÇÃO EM ARTIGOS CIENTÍFICOS EM
CIÊNCIAS HUMANAS E EM CIÊNCIAS EXATAS**

São José do Rio Preto

2023

Larissa Souza da Silva

AUTOCITAÇÃO EM ARTIGOS CIENTÍFICOS EM CIÊNCIAS HUMANAS E EM CIÊNCIAS EXATAS

Dissertação apresentada como requisito parcial para a obtenção do Título de Mestre em Estudos Linguísticos, junto ao Programa de Pós-Graduação em Estudos Linguísticos do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho” (Unesp), Câmpus de São José do Rio Preto.

Financiadora: CAPES (código de financiamento 001)

Orientadora: Prof^ª. Dr^ª. Fabiana Komesu – Unesp

Coorientadora: Prof^ª. Dr^ª. Adriana Fischer –Fundação Universidade Regional de Blumenau (FURB)

São José do Rio Preto

2023

S586a Silva, Larissa Souza da
Autocitação em artigos científicos em Ciências Humanas e em Ciências Exatas /
Larissa Souza da Silva. -- São José do Rio Preto, 2023
88 p. : il., tabs.

Dissertação (mestrado) - Universidade Estadual Paulista (Unesp), Instituto de
Biociências Letras e Ciências Exatas, São José do Rio Preto
Orientadora: Fabiana Cristina Komesu
Coorientadora: Adriana Fischer

1. Letramentos Acadêmicos. 2. Autocitação. 3. Discurso Científico. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca do Instituto de Biociências Letras e Ciências
Exatas, São José do Rio Preto. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

Larissa Souza da Silva

AUTOCITAÇÃO EM ARTIGOS CIENTÍFICOS EM CIÊNCIAS HUMANAS E EM CIÊNCIAS EXATAS

Dissertação apresentada como requisito parcial para a obtenção do Título de Mestre em Estudos Linguísticos, junto ao Programa de Pós-Graduação em Estudos Linguísticos do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho” (Unesp), Câmpus de São José do Rio Preto.

Financiadora: CAPES (código de financiamento 001)

Comissão Examinadora

Prof^ª. Dr^a. Fabiana Komesu
Universidade Estadual Paulista (Unesp) – Câmpus de São José do Rio Preto
Orientadora

Prof^ª. Dr^a. Adriana Fischer
Fundação Universidade de Blumenau (FURB)
Coorientadora

Prof^ª. Dr^a. Maria Angela Paulino Teixeira Lopes
Pontifícia Universidade Católica de Minas Gerais (PUC Minas)
Membro titular

Prof^ª. Dr^a. Maria Raquel Miotto Morelatti
Universidade Estadual Paulista (Unesp) – Câmpus de Presidente Prudente
Membro titular

Prof^ª. Dr^a. Fernanda Correa Silveira Galli
Universidade Federal de Pernambuco (UFPE)/ PPG Estudos Linguísticos – Unesp
Membro suplente

Prof^ª. Dr^a. Juliana Alves Assis
Pontifícia Universidade Católica de Minas Gerais (PUC Minas)
Membro suplente

São José do Rio Preto

17 de abril de 2023

À memória de minha avó, Olga.

Agradecimentos

À minha avó Olga, fonte de inspiração, amor e aprendizado em sua vivência de mulher preta. Minha gratidão por seu apoio constante em minha vida acadêmica, por ser colo e abraço, com palavras de sabedoria.

Aos meus pais Clarice e José, para quem sempre posso voltar e contar o vivido e o sentido. Minha gratidão por seu amor e apoio incondicionais, sobretudo, na realização deste trabalho. Eu sou porque vocês são.

Ao meu irmão Brendo, com quem choro e rio, compartilhando a vida. Minha gratidão por estar ao meu lado, me encorajando a ir além.

Aos meus amigos da Unesp (da vida), que me acolheram com sorrisos, especialmente, Isabela, Miller e Helen. Aos igualmente queridos: Aline, Carla, Daniel, Isadora, João Paulo, Lucas, Marcela e Nathália: que felicidade ter vocês em minha vida!

Às professoras doutoras Maria Angela Paulino Teixeira Lopes e Maria Raquel Miotto Morelatti, pelas importantes contribuições desde o exame de qualificação geral da Dissertação. Também agradeço às professoras doutoras Fernanda Correa Silveira Galli e Juliana Alves Assis por aceitarem participar, na qualidade de membros suplentes, dos exames de qualificação e de defesa.

Aos professores do Programa de Pós-graduação em Estudo Linguísticos (PPGEL) da Unesp, meus agradecimentos por minha formação acadêmica. Aos servidores técnico-administrativos da Seção Técnica de Pós-graduação da Unesp, câmpus de São José do Rio Preto, meus agradecimentos pelos inúmeros serviços prestados com eficiência e gentileza.

Aos colegas do grupo de pesquisa “Práticas de leitura e escrita em contexto digital” (Unesp/CNPq), meus agradecimentos pela oportunidade de diálogo e aprendizado conjunto. Também sou grata aos colegas dos grupos de pesquisa da FURB e da PUC Minas (“Núcleo de Estudos em Linguagens, Letramentos e Formação” – NELLF/CNPq/PUC Minas) pelo trabalho coletivo realizado por ocasião da coleta dos dados.

À minha orientadora, Fabiana Komesu, por sua orientação humanizadora, pela competência, disponibilidade e encorajamento neste percurso acadêmico. Minha gratidão também se estende à minha coorientadora, professora doutora Adriana Fischer.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

Resumo

Com base em pressupostos advindos dos estudos de letramentos, de maneira particular, de letramentos acadêmicos, este trabalho tem como objetivo principal investigar os padrões de autocitação em artigos científicos das grandes áreas de Ciências Humanas (subárea de Linguística) e Ciências Exatas (subárea de Matemática Aplicada). Parte-se da assunção de que a autocitação é uma forma retórica de manter a credibilidade profissional, podendo variar interdisciplinarmente. Os objetivos específicos são: (i) estudar padrões de autocitação nos artigos científicos dessas duas diferentes grandes áreas de conhecimento, considerando-se o modo como esses artigos se autorreferenciam no processo de autocitação; (ii) estudar as funções da autocitação nessas duas diferentes grandes áreas de conhecimento, considerando suas atuações na construção da dimensão argumentativa do discurso científico. O conjunto do material é formado de 40 (quarenta) artigos científicos publicados em língua portuguesa, 20 (vinte) de cada subárea, 10 (dez) de cada uma das plataformas escolhidas (*Web of Science* e *SciELO*), com alto índice de citação segundo indicadores bibliométricos, no período de 2015 a 2019. Parte-se da hipótese de que, por um lado, há uma prevalência das demais áreas de conhecimento em relação às Ciências Humanas, segundo uma dinâmica de relações de poder e autoridade em práticas letradas acadêmico-científicas; por outro, o processo de autocitação corrobora a contagem de citações em indicadores bibliométricos, com promoção de visibilidade em comunidades discursivas. Interessa, pois, a este trabalho investigar o processo de produção de textos científicos em relações de poder e autoridade, como as derivadas de indicadores bibliométricos que promovem a identificação e o reconhecimento de determinados autores, segundo quantificação, com consequências para a visibilidade em plataformas de bases de dados.

Palavras-chave: letramentos acadêmicos; autocitação; artigo científico; Ciências Humanas; Ciências Exatas; discurso científico.

Abstract

Based on assumptions arising from Academic Literacy Studies, the main objective of this work is to investigate the patterns of self-citation in scientific articles in the broad areas of Soft Sciences (sub-area of Linguistics) and Hard Sciences (sub-area of Applied Mathematics). The specific objectives are: (i) to study patterns of self-citation in scientific articles of these two different areas of knowledge, taking into account the way in which these articles use self-referencing in the self-citation process; (ii) to study the functions of self-citation in articles of these two different areas of knowledge, considering their role in the construction of the argumentative dimension of scientific discourse. The corpus consists of 40 (forty) scientific articles published in Brazilian Portuguese, 20 (twenty) from each sub-area, 10 (ten) from each of the chosen platforms (Web of Science and SciELO), with a high citation rate according to bibliometric indicators, from 2015 to 2019. The hypothesis is that there is a prevalence of other fields of knowledge in relation to the Humanities, according to a dynamic of power and authority relations in academic-scientific literate practices. The process of self-citation also corroborates the counting of citations in bibliometric indicators, with promotion of visibility in discursive communities. Therefore, this work is interested in investigating the process of production of scientific texts in relations of power and authority, such as those derived from bibliometric indicators that promote the identification and recognition of certain authors, according to quantification, with consequences for visibility in database platforms.

Keywords: academic literacies; self-citation; scientific article; Soft Sciences; Hard Sciences; scientific discourse.

Lista de figuras

Figura 1 – Cálculo de fator de impacto (FI)	23
Figura 2– Modos de referência ao discurso do outro	38
Figura 3 – Busca no Portal da plataforma CAPES	43
Figura 4 – Seleção dos filtros e ordenação dos artigos	44
Figura 5– Acesso à plataforma SciELO através da WoS	44
Figura 6 – Organização do conjunto do material	45
Figura 7 - Busca por artigos da subárea de Matemática	46
Figura 8 – Resultado final da busca de artigos na subárea Matemática, plataforma WoS	46
Figura 9 – Interface da tela inicial do software MAXQDA	54
Figura 10 - Classificação dos códigos em cores no software MAXQDA	55
Figura 11 – Print da tela do software MAXQDA, com indicação de “busca lexical”	55
Figura 12 – Exemplo de empréstimo, retirado de artigo da subárea de Linguística	56
Figura 13 – Exemplo de evocação, retirado de um artigo da subárea de Linguística	56
Figura 14 – Exemplos de reformulação e citação direta, retirados de artigo da subárea de Linguística	57
Figura 15– Estrutura IMRD a partir de Feak e Swales (2009)	59
Figura 16 – Funcionamento geral das autocitações nas subáreas de Linguística e Matemática Aplicada	66
Figura 17 – Funcionamento das autocitações na subárea de Linguística	67
Figura 18 – Funcionamento das autocitações na subárea de Matemática	68
Figura 19 – Autocitação por seção em artigos da subárea de Linguística, em ambas as bases	69
Figura 20 – Autocitação por seção em artigos da subárea de Matemática Aplicada, em ambas as bases	70
Figura 21 – Tipos de citação dos artigos na subárea de Linguística	74

Figura 22 – Tipos de citação dos artigos na subárea de Matemática Aplicada	75
Figura 23 - Data de publicação dos artigos por ano	77
Figura 24 - Demonstração de autocitação por base de dados	77

Lista de tabelas

Tabela 1 – Dados gerais dos artigos na subárea de Linguística	63
Tabela 2– Dados gerais dos artigos na subárea de Matemática Aplicada	64
Tabela 3 – Modos de referênciação por seção de artigos na subárea de Linguística	71
Tabela 4 - Modos de referênciação por seção de artigos na subárea de Matemática Aplicada	71

Lista de quadros

Quadro 1 – Critérios que permitem diferenciar os modos de referência ao discurso do outro	38
Quadro 2 – Artigos da subárea de Matemática Aplicada coletados junto à plataforma Web of Science	47
Quadro 3 – Artigos da subárea de Matemática Aplicada coletados junto à plataforma SciELO CI	49
Quadro 4 – Artigos da subárea de Linguística coletados juntos à plataforma Web of Science	51
Quadro 5 – Artigos da subárea de Linguística coletados junto à plataforma SciELO CI	52
Quadro 6 – Estrutura composicional dos artigos na subárea de Linguística	59
Quadro 7– Estrutura composicional dos artigos na subárea de Matemática Aplicada	61
Quadro 8 – Exemplos dos tipos de citação nos artigos de Linguística	73

Lista de abreviaturas e siglas

BIREME - Biblioteca Regional de Medicina

CAPES - Coordenação de Aperfeiçoamento de Pessoal de Nível Superior

CAFE - Comunidade Acadêmica Federada

CNPq - Conselho Nacional de Desenvolvimento Científico e Tecnológico

FI - Fator de Impacto

ISI - Institute of Scientific Information

JCR - Journal Citation Reports

LWoS - Linguística/ Web of Science

LSci - Linguística/ SciELO

MCTI - Ministério da Ciência, Tecnologia e Inovações

MSCI - Matemática/SciELO

MWoS - Matemática/Web of Science

NEL - Novos Estudos do Letramento

NLS - New Literacy Studies

QDA - Qualitative Data Analysis

SCI - Science Citation Index

SciELO - Scientific Eletronic Library Online

UNESP - Universidade Estadual Paulista Júlio de Mesquita Filho

WoS - Web of Science

Sumário

<i>Introdução</i>	13
<i>Capítulo 1 – Fundamentação teórica</i>	17
1.1. A perspectiva dos estudos de letramentos	17
1.2. Produção, organização e distribuição do conhecimento científico	21
1.3. O fenômeno da autocitação em culturas disciplinares distintas	28
<i>Capítulo 2 – Material e Métodos</i>	40
2.1. A escolha das subáreas de Linguística e Matemática	40
2.2. A escolha das plataformas de bases de dados	41
2.3. O conjunto do material e a constituição do <i>corpus</i>	45
2.4. Procedimentos metodológicos de organização e critérios para análise dos dados	53
<i>Capítulo 3 – Análise dos dados</i>	58
3.1. Da macroestrutura do artigo científico nas duas subáreas do conhecimento	58
3.2. Autocitação nas subáreas de Linguística e Matemática Aplicada	63
3.3. Modos de referenciação nas subáreas de Linguística e Matemática Aplicada	66
3.2.1. Autocitação por seção do artigo científico nas duas subáreas do conhecimento	68
3.2.2. Modos de referenciação por seção de artigo nas duas subáreas do conhecimento	70
3.3. Categorias de autocitação em Linguística e Matemática Aplicada	72
3.4. A visibilidade nas plataformas de bases de dados	76
<i>Considerações Finais</i>	79
<i>Referências</i>	83

Introdução

Esta Dissertação de Mestrado está associada a projetos de pesquisa mais amplos, em âmbito nacional e internacional, a saber:

- (i) “Escrita acadêmica/escrita científica: das formas de presença do autor, do outro, das áreas de conhecimento e seus domínios disciplinares”, coordenado pela Profa. Dra. Juliana Alves Assis (PUC Minas), com financiamento do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq/ Chamada Universal), no período de 2019 a 2021, e com a participação de pesquisadores do Brasil (FURB, IFMG, PUC Minas, UEMG, UFOP, Unesp e USP) e do exterior (Université de Lille, França);
- (ii) “Literacies in Different Fields of Knowledge”, coordenado pela Profa. Dra. Fabiana Komesu (Unesp), com financiamento CAPES-PrInt-Unesp, com início em 2019 (em andamento) e com a participação de pesquisadores do Brasil (FURB, IFMG, PUC Minas, UEMG, UFOP, Unesp e USP) e do exterior (Dartmouth College, Estados Unidos; The Ohio State University, Estados Unidos; The Open University, Reino Unido; Universidade de Lisboa, Portugal; Université de Lille, França; Université Libre de Bruxelles, Bélgica; University of Rochester, Estados Unidos).

Os projetos mais amplos e este trabalho em nível de Dissertação buscam investigar práticas letradas acadêmico-científicas em diferentes grandes áreas de conhecimento, levando em consideração, dentre outros aspectos, a distinção entre culturas disciplinares; o “atravessamento” de demandas do contexto sócio-histórico da produção do conhecimento, em nível global; bem como a circulação desse conhecimento, orientada por tecnologias algorítmicas e de plataformas de dados utilizados na avaliação para concessão de recursos financeiros (distribuídos, entretanto, de maneira não equânime, como discutido adiante) para docentes e pesquisadores, laboratórios e instituições.

Considerando, pois, que tanto a atividade do pesquisador quanto a área ou a subárea em que atua são impactadas por aspectos que envolvem desde a produção até a divulgação do saber científico, partimos da hipótese de que, por um lado, há uma prevalência das demais áreas de conhecimento em relação às Ciências Humanas, segundo uma dinâmica de relações de poder e autoridade em práticas letradas acadêmico-científicas (LEA; STREET, 2014;

SALGADO; CLARES, 2017) e, por outro, o processo de autocitação corrobora a contagem de citações em indicadores bibliométricos, com promoção de visibilidade em comunidades discursivas.

Assim, com base em pressupostos teórico-metodológicos dos letramentos acadêmicos, o objetivo principal desta Dissertação é investigar padrões de autocitação em artigos científicos de duas grandes áreas de conhecimento, Ciências Humanas (subárea de Linguística) e Ciências Exatas (subárea de Matemática), tencionando contribuir com uma reflexão sobre o modo de funcionamento do discurso científico em culturas disciplinares distintas.¹ Este objetivo geral desdobra-se em outros, específicos: (i) estudar padrões de autocitação em artigos científicos dessas duas diferentes grandes áreas de conhecimento, considerando o modo como esses artigos se autorreferenciam no processo de autocitação; (ii) estudar as funções da autocitação nos artigos dessas duas diferentes grandes áreas de conhecimento, considerando as suas atuações na construção da dimensão argumentativa do discurso científico. O conjunto do material é formado de 40 (quarenta) artigos científicos publicados em língua portuguesa, 20 (vinte) de cada subárea (Linguística e Matemática), 10 (dez) de cada uma das plataformas escolhidas (*Web of Science* e *SciELO*), com alto índice de citação segundo indicadores bibliométricos, no período de 2015 a 2019, como será explicado adiante na seção sobre Material e Métodos.

Estudos sobre autocitação em produções científicas têm se tornado objeto de pesquisadores em diferentes áreas do conhecimento, a exemplo da cientometria, da bibliometria e dos estudos linguísticos. Uma maneira de entender o conceito de autocitação é pensar, com Ioannidis (2015, p. 7), que se trata de citar um trabalho anterior em trabalhos acadêmicos subsequentes. É ainda esse autor que considera que a autocitação pode ser concebida como um componente do discurso acadêmico-científico, operando como uma variável moduladora importante no campo da ciência (IOANNIDIS, 2015, p.10). Da perspectiva assumida neste trabalho, trata-se de compreender o modo como a prática de autocitação está engendrada na produção do discurso científico e como pode indicar aspectos

¹ A nomenclatura das grandes áreas do conhecimento e das subáreas, utilizada nesta pesquisa, foi orientada pela tabela de classificação das Áreas do Conhecimento da Coordenação de Pessoal de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), fundação subordinada ao Ministério da Educação (MEC), responsável pela consolidação da pós-graduação *stricto sensu* em todos os Estados do Brasil, e do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), fundação subordinada ao Ministério da Ciência, Tecnologia e Inovações (MCTI), responsável pelo incentivo à pesquisa no Brasil. Assim, entende-se “grande área” como a concentração em uma área do conhecimento em virtude da afinidade de objetos e métodos amplamente reconhecidos e utilizados, que refletem, por sua vez, contextos sociopolíticos específicos, segundo a CAPES. Por “subárea” entende-se uma segmentação da área de conhecimento. Assim, na grande área de Ciências Humanas, selecionamos a subárea de Linguística; em Ciências Exatas, a subárea de Matemática.

importantes da construção do conhecimento e dos padrões que caracterizam relações disciplinares (HYLAND, 2003).

Segundo Salgado e Clares (2017, p. 31), uma das questões fundamentais que regem a comunicação científica no Brasil é a dominação das Ciências Exatas e Naturais sobre as Ciências Humanas. Essa prevalência baseia-se, de acordo com as autoras, numa concepção de ciência universalista. No caso do estudo conduzido pelas autoras, referente à investigação de normas reguladoras de periódicos científicos, a escrita científica em humanidades é considerada desafiadora, uma vez que essas normas são dadas como gerais para qualquer área de conhecimento e são frequentemente reproduzidas com base em modelos internacionais que não necessariamente representam as particularidades das pesquisas locais. Haveria, pois, uma espécie de “matematização” na estrutura do artigo científico – incluído o produzido em Linguística –, que pode ser encontrada em seções como “Material e Métodos”, dentre outras, características de uma pesquisa como a laboratorial. É, pois, dessa afirmação das autoras que elaboramos a hipótese segunda a qual haveria uma dominação das Ciências Exatas e Naturais sobre as Ciências Humanas, segundo uma concepção de ciência universalista, que desconsidera distinções entre domínios disciplinares.

Esta Dissertação de Mestrado está organizada da seguinte maneira: no Capítulo 1, *Fundamentação Teórica*, apresenta-se o referencial teórico sobre os Novos Estudos do Letramento e sobre letramentos acadêmicos. São discutidos em subseções, conceitos fundamentais à Dissertação, como o fenômeno da autocitação nos estudos de letramentos; modos de produção, circulação e distribuição do conhecimento científico; conceitos fundamentais à Dissertação, como o fenômeno da autocitação; fundamentos da bibliometria em estudos que visam à criação de indicadores sobre a produtividade de autores e periódicos científicos.

No Capítulo 2, *Material e Métodos*, são apresentadas as justificativas para a escolha das subáreas de Linguística e Matemática (Matemática Aplicada, como discutido adiante); das plataformas de bases de dados; do conjunto do material e a constituição do *corpus*; dos procedimentos metodológicos de organização e critérios para análise dos dados.

No Capítulo 3, *Análise de dados*, faz-se a exposição e análise dos resultados referentes aos critérios de análise previamente estabelecidos, considerando as variáveis: macroestrutura do artigo científico nas duas diferentes subáreas; autocitação nas duas diferentes subáreas, segundo modos de referência e seções do artigo em que a autocitação figura; categorias de autocitação nas duas diferentes subáreas; visibilidade nas plataformas de bases de dados.

Por fim, em *Considerações finais*, expomos uma síntese dos principais resultados obtidos nesta pesquisa.

Capítulo 1 – Fundamentação teórica

1.1. A perspectiva dos estudos de letramentos

Este trabalho ancora-se em pressupostos dos estudos de letramentos, mais particularmente, dos letramentos acadêmicos. Inicialmente, traremos à discussão autores como Street (1984), Kleiman (1995) e Soares (2003), dentre outros, procurando fazer uma retomada dessas contribuições para a investigação de práticas sociais de leitura e escrita em âmbito acadêmico-científico. Vamos nos apoiar ainda nas reflexões de Lea e Street (1998, 2014) e Lillis e Curry (2011), que nortearão a discussão sobre a produção textual escrita em contextos acadêmicos-científicos e de internacionalização da pesquisa científica. Considerando os objetivos deste trabalho, trataremos, em seguida, da produção, organização e distribuição do conhecimento científico, com base na problematização dos chamados índices bibliométricos. Discutiremos a posição de autores que estudam o funcionamento do fenômeno da autocitação do ponto de vista da linguagem, como Hyland (2003, 2010), Hyland e Jiang (2018) e Boch e Grossmann (2002, 2015).

Os Novos Estudos do Letramento (NEL) – em inglês, *New Literacy Studies* – levam em conta a natureza social e heterogênea de práticas de leitura e escrita. A partir da contribuição de pesquisadores filiados a uma perspectiva sociocultural (STREET, 1984, 2001), esses estudos se opõem a uma visão segundo a qual os usos da escrita seriam decorrentes de habilidades individuais. Os NEL, portanto, se contrapõem a uma visão que considera “efeitos universais do letramento” (KLEIMAN, 1995, p.16), que seriam independentes dos sujeitos e das condições sócio-históricas de produção de sentidos, propondo, em seu lugar, compreender a diversidade de práticas sociais letradas, bem como suas especificidades condicionadas a relações de poder e autoridade, a valores e crenças de cada sociedade.

As contribuições de Street (1984), na década de 1980, que nortearam os estudos brasileiros sobre letramentos, problematizam dois modelos, o autônomo e o ideológico. Como é sabido, o primeiro concebe práticas de leitura e de escrita na escola como independentes de seu contexto de uso e do processo de sua produção. Esse modelo concebe leitura e escrita como neutras, resultantes de processos cognitivos e individuais do aluno, desconsiderando, portanto, a dimensão de poder existente nesse processo. No âmbito desse

modelo, “o” letramento por excelência seria o escolar. O modelo ideológico, por sua vez, refere-se a práticas sociais de uso da linguagem. Esse modelo, defendido por Street (1984), reconhece os diversos tipos de letramento (não restritos ao letramento escolar, razão pela qual o autor trata de letramentos, no plural) e considera que práticas sociais de leitura e escrita estão fortemente relacionadas a um contexto específico de produção, não sendo, portanto, neutras. A esse respeito, Kleiman (1995, p. 38) afirma que, para Street (1984), o modelo ideológico compreende que todas as práticas de letramentos são aspectos de uma cultura e de estruturas de poder de uma sociedade.

No Brasil, no campo da Linguística Aplicada, Kleiman (1995) define letramentos como um conjunto de práticas sociais que compreende a leitura e a escrita enquanto sistema simbólico e tecnológico, em contextos específicos para objetivos específicos. Para enfatizar o efeito potencializador ou conferidor de poder do letramento, a autora argumenta que: “O domínio de outros usos e funções da escrita significa, efetivamente, o acesso a outros mundos, públicos e institucionais, como o da mídia, da burocracia, tecnologia, e através deles, a possibilidade de acesso ao poder.” (KLEIMAN, 1995, p. 8).

No que se refere aos contextos específicos, Kleiman (1995) avalia que a escola é a mais importante das agências de letramento, ocupando-se como instituição, entretanto, de um tipo de prática, a da alfabetização, mediante aquisição (tomada como de competência individual) de códigos necessários para a promoção da própria escola. Observa-se, pois, com Kleiman (1995, p.19) que o uso da escrita valorizado pela escola se apoia num modelo autônomo de letramento (STREET, 1984), que pressupõe que há apenas uma maneira de o letramento ser desenvolvido: na própria escola. Esse modelo, como mencionado, contrapõe-se ao modelo ideológico que concebe a existência de outras práticas de letramento, socioculturalmente determinadas e igualmente dependentes de relações institucionais.

A partir desses estudos dos letramentos, estudiosos voltaram o olhar para outros contextos, como o do ensino superior. A abordagem direcionada a práticas de leitura e escrita acadêmico-científicas foi desenvolvida, dentre outros, por Street (2010) e Lea e Street (2014). A esse respeito, Street (2010, p. 545) avalia que o modelo de letramentos acadêmicos reconhece a escrita na universidade como prática social em contexto institucional e disciplinar determinados, considerando relevantes fatores como o poder e autoridade no processo de produção textual.

Lea e Street (2014) discutem um modelo de letramentos acadêmicos em dois contextos distintos, no de estudantes do Programa de Desenvolvimento do Letramento Acadêmico e no de estudantes da Faculdade de Direito da Universidade Aberta de Londres.

Esse estudo buscou destacar a variedade de práticas institucionais, o enfrentamento dos estudantes para que essas práticas fizessem sentido e as identidades envolvidas no aprendizado dentro e fora dos contextos acadêmicos. Os autores buscavam, assim, criticar o modelo dominante do *déficit* e a mera classificação dos textos acadêmicos como “bons” ou “ruins”. Com base em resultados de pesquisa, os autores descreveram três modelos no processo de produção textual em nível superior (mas não somente), cuja atuação não é mutuamente exclusiva, mas sobreposta. Trata-se do modelo das habilidades de estudos, modelo de socialização acadêmica e modelo dos letramentos acadêmicos.

Segundo Lea e Street (2014, p. 481), o modelo de habilidade de estudos tem como foco habilidades cognitivas, concebendo a escrita como produto individual, voltado ao domínio das regras gramaticais e sintáticas, com pouca atenção ao contexto. O segundo modelo, o da socialização acadêmica, dizem os autores, pressupõe que os alunos, uma vez apresentados aos gêneros disciplinares específicos, e tendo, portanto, se apropriado de regras básicas do discurso acadêmico, teriam condições de reproduzi-los de forma eficaz. O modelo dos letramentos acadêmicos, por sua vez, relaciona-se com a construção dos sentidos, de identidade, poder e autoridade, cuja prioridade é destacar o caráter institucional do que é tido como conhecimento nos diferentes contextos acadêmicos. Para Lea e Street (2014), esse último modelo assemelha-se em grande medida ao de socialização acadêmica, salvo pelo fato de que o modelo de letramentos acadêmicos compreende a aquisição e apropriação dos discursos acadêmicos como sendo mais “complexos, dinâmicos, com nuances diferenciadas, envolvendo epistemológicas quanto processos sociais, incluindo as relações de poder entre as pessoas e instituições, e as identidades sociais.” (LEA; STREET, 2014, p. 546).

Embora cada modelo seja considerado pelos autores como complementar ao outro, o modelo dos letramentos acadêmicos é o que melhor estabelece relações entre a produção escrita e práticas institucionais, relações de poder e identidades envolvidas no processo de construção de sentido. É esse modelo o responsável por viabilizar projetos de aprendizagem disciplinares específicos, como também é responsável por orientar educadores para desenvolvimento de currículos pedagógicos presentes em instituições de ensino básico e superior. É levando em conta, pois, a dimensão institucional de práticas letradas acadêmico-científicas que propomos a investigação, num primeiro plano, da autocitação e de seu funcionamento e, num segundo, dos artigos científicos e dos modos de circulação do conhecimento em práticas disciplinares distintas.

Da perspectiva dos estudos de letramentos e dos letramentos acadêmicos, é importante mencionar a questão das relações de poder na ciência, entendidas, neste trabalho,

como um dos pilares que norteiam a produção da pesquisa, segundo um processo de internacionalização da publicação acadêmica. Embora o conjunto do material desta pesquisa seja formado exclusivamente de publicações em língua portuguesa (como explicado adiante), a discussão é relevante, pois contribui para a reflexão sobre o (baixo) impacto e a (ausência de) visibilidade das publicações coletadas em bases de dados, em distintas comunidades disciplinares, até mesmo em seu país de origem. A internacionalização da pesquisa científica, como padrão em lugares onde o inglês não é o idioma oficial, reforça a ideia de que letramentos, em contextos situados, são instituídos de acordo com relações de poder, no caso da produção científica, em âmbito mundial. Lillis e Curry (2011, p. 186) problematizam o *status* do inglês como língua franca, isto é, como um idioma que um grupo de falantes multilíngues elegeria intencionalmente para que todos conseguissem se comunicar uns com os outros, desempenhando, assim, um papel central nessa globalização sistemática, como “linguagem da ciência global”.

O inglês, no contexto institucional mundial, configura-se como o idioma padrão da pesquisa e da divulgação acadêmico-científica, uma vez que a utilização desse idioma seria uma forma de “valorizar” a ciência local, segundo um prestígio conferido fora do próprio contexto anglófono. Para Lillis e Curry (2011, p. 15, tradução nossa),² “esse *status* global de inglês, junto ao crescimento documentado de publicações em inglês mediano, significa que estudiosos de todo o mundo estão sob considerável pressão para publicar em inglês”, uma vez que essa língua é assumida na produção de conhecimento científico em níveis nacional e internacional.

As autoras observam que a política de localização na produção de textos acadêmicos se depara com pelo menos três dimensões que precisam ser levadas em consideração: (i) a *geográfica* (em contexto local imediato, como o de departamento, instituição, bem como em nível de Estado ou região), (ii) a da *geolocalização linguística* (idiomas usados ou não usados na produção escrita para publicação acadêmica, com *status* diferencial) e (iii) a da *geopolítica* (notadamente políticas que influenciam sistemas de investigação e avaliação em níveis local, nacional e supranacional) (LILLIS; CURRY, 2011, p. 6).

² No original: “This global status of English alongside the documented growth in English-medium publications means that scholars from around the world are under considerable pressure to publish in English”. (LILLIS; CURRY, 2011, p. 15).

Lillis e Curry (2011, p. 15, tradução nossa)³ esclarecem ainda que “onde quer que a escrita acadêmica para publicação ocorra e para quem quer que ela seja destinada, essa escrita é global na medida em que emerge de um mercado global (Gibbs 1995a), em que os textos são literalmente de valor diferente, segundo uma ‘economia de signos’ global (Blommaert 2005).” No caso do conjunto do material selecionado para este trabalho, observamos, por exemplo, que o periódico *Bulletin of the Brazilian Mathematical Society* tem título em língua inglesa. Trata-se de uma revista científica criada em 1970 e editada pela Sociedade Brasileira de Matemática. É distribuída pela editora multinacional Springer Verlag, com apoio do CNPq/MCTI e CAPES/MEC. Tem fator de impacto (FI) 1.246, como discutido a seguir. Segundo o texto de apresentação da revista, “a fim de facilitar sua circulação em todo o mundo, os manuscritos devem ser submetidos em inglês ou francês.”⁴ É, pois, buscando essa visibilidade internacional que o título do periódico também aparece em língua inglesa. Na subseção seguinte, discutiremos a produção, a organização e a distribuição do conhecimento científico no formato de artigos de periódicos de alto impacto.

1.2. Produção, organização e distribuição do conhecimento científico

A atividade científica tem sido avaliada com base em aspectos quantitativos, segundo políticas de produção de alto impacto num cenário de mercado editorial científico mundial.⁵ Segundo Clares (2017, p.36), as normas reguladoras da produção científica que são adotadas no tratamento editorial de textos acadêmicos condicionam não só ajustes técnicos, estruturais, linguísticos, mas também as circunstâncias nas quais esses ajustes devem ocorrer, de acordo com prazos, exigência (ou não) de coautoria e de modos de citação.

Nesse contexto, periódicos científicos configuram-se como dispositivos comunicacionais que apresentam o conhecimento acumulado, com indicativo de prestígio e reconhecimento de estudos já realizados. Conforme aponta Gruszynski, Golin e Castelo (2008), o periódico garante

³ No original: “Wherever academic writing for publication takes place and whoever it is aimed at, such writing is global to the extent that it occurs within a global market (Gibbs 1995a), where texts are quite literally accorded different value, and within a global ‘economy of signs’ (Blommaert 2005)”. (LILLIS; CURRY, 2011, p. 15).

⁴ Disponível em: <https://sbm.org.br/bulletin-of-the-brazilian-mathematical-society/>. Acesso em: 09 mar. 2023.

⁵ Komesu e Assis (2022) também discutem esse cenário de produtivismo científico, apontando, com base em estudos, a tentativa de adoção de parâmetros qualitativos de avaliação.

a memória da ciência, aponta seu grau de evolução, estabelece a propriedade intelectual, legitima novos campos de estudos e disciplinas, constitui-se em fonte para o início de novas pesquisas, dando visibilidade e prestígio aos pesquisadores entre um público altamente especializado, os seus pares. (GRUSZYNSKI; GOLIN; CASTELO, 2008, p. 4)

É sabido, entretanto, que a produção científica nem sempre foi disseminada e medida como o é atualmente. Segundo Amante (2011), no século XVII, por exemplo, os cientistas comunicavam ideias e resultados experimentais por meio de cartas. Foi publicado na França, em 1665, o primeiro número do que seria uma revista científica, *Journal des Sçavans* e, dois meses mais tarde, na Inglaterra, *Philosophical Transactions*, provavelmente, publicações decorrentes da necessidade de os investigadores apresentarem seus trabalhos fora de círculos profissionais restritos. Ainda de acordo com Amante (2011, p. 5), tratou-se das primeiras revistas científicas da Europa, as quais instituíram o início do sistema de comunicação científica baseado em periódicos, como hoje reconhecido.

Em decorrência do crescimento científico permanente, principalmente, por meio de revistas, surge a necessidade de codificar e organizar o conhecimento de maneira a ser disponibilizado para consulta. Desde então, os métodos de análise da produção científica e tecnológica cresceram exponencialmente.

Marziale e Mendes (2002) explicam que, na década de 1960, Eugene Garfield, diretor do *Institute of Scientific Information* (ISI) e fundador da base de dados *Science Citation Index*, foi o responsável por introduzir o *Fator de Impacto*, doravante FI, na avaliação dos periódicos científicos. O FI é considerado um instrumento bibliométrico que objetiva avaliar a produção científica dos autores, a qualidade das publicações, e por fim, classificar os períodos inseridos no *Journal Citation Reports* (JCR) (MARZIALE; MENDES, 2002, p. 466). Esse instrumento que, até nos dias de hoje, mede a frequência com que um artigo é citado, como forma de classificar e avaliar as revistas indexadas nas bases de dados, funciona da seguinte forma:

O valor do fator de impacto é obtido dividindo-se o número total de citações dos artigos, acumulados nos últimos dois anos, pelo total acumulado de artigos publicados pela revista no referido período. Somente os periódicos indexados no ISI são considerados para o cálculo do fator de impacto internacional. (MARZIALE; MENDES, 2002, p. 466)

A Figura 1 a seguir, retirada de Thomaz, Assad e Moreira (2011, p. 91), busca demonstrar essa dinâmica, com a observação de que “o fator de impacto é eficiente em

avaliar a qualidade de um periódico, não sendo, porém, útil na análise da qualidade científica de um artigo isolado, de um pesquisador ou de uma instituição”:

Figura 1 – Cálculo de fator de impacto (FI)

Tabela 1 - Cálculo do fator de impacto do periódico *Circulation* durante 2009

Ano da publicação	Artigos publicados	Número de citações em 2009
2007	670	11.420
2008	607	7.500
2007 + 2008	1.277	18.920

Fator de Impacto: $18.920 \div 1.277 = 14,816$. Fonte: Journal Citation Reports

Fonte: Thomaz, Assad e Moreira (2011, p. 91)

A análise bibliométrica tem sido realizada desde o século XIX e, desde então, as informações bibliográficas, agrupadas em banco de dados, foram foco de estudos que resgataram ou originaram novas designações, como é o caso das ciências da biblioteconomia, da cientometria, e infometria, por exemplo (BOUSTANY, 1997, p.156 *apud* SANTOS; KOBASHI, 2009, p.157).

Cada uma dessas designações mencionadas – sintomas de crescimento do campo científico – gera controvérsias, pois, mesmo entre os especialistas das áreas não há consenso sobre as terminologias. Segundo Santos e Kobashi (2009), a bibliometria, por exemplo, tem como objeto de estudo livros ou revistas científicas, cujas análises se vinculam à gestão de bibliotecas e bases de dados; a cientometria, por sua vez, preocupa-se com a dinâmica da ciência, como atividade social, tendo como objeto de análise a produção, a circulação e o consumo da produção científica; a infometria, por fim, abarca as duas primeiras, tendo desenvolvido métodos e ferramentas para mensurar e analisar os aspectos cognitivos da ciência (SANTOS; KOBASHI, 2009, p.156).

A bibliometria é uma disciplina que emerge na Biblioteconomia e na Ciência da Informação, cujo enfoque é o estudo dos aspectos quantitativos da produção, disseminação e

socialização da informação registrada (MACIAS-CHAPULA, 1998 *apud* MELO RIBEIRO, 2017, p. 2). A criação do termo “bibliometria” é atribuída a Pritchard, que a caracterizou como um “conjunto de métodos e técnicas quantitativos para a gestão de bibliotecas e instituições envolvidas com o tratamento de informação” (MELO RIBEIRO, 2017, p. 157). Segundo Garcia (2008 *apud* AMANTE, 2011), a bibliometria evolui ao longo do tempo e recebeu um forte impulso nos anos 1970, quando os dados ainda eram coletados manualmente. Nos anos 1980, a bibliometria tornou-se uma disciplina científica com vários subdomínios e estruturas de comunicação científica, tendo como expoente a revista *Scientometrics*. Foi a partir da década de 1990 que a disciplina bibliometria passou a ser concebida como técnica de análise padrão para diversas ciências (PATRA; BHATTACHARYA; VERMA, 2006 *apud* MELO RIBEIRO, 2017, p. 3).

Garcia (2007 *apud* AMANTE, 2011, p. 3, tradução nossa) argumenta que o desenvolvimento dos estudos bibliométricos se deve a:

(i) razões técnicas, fundamentalmente resultantes da informação das bases de dados bibliográficos em todas as áreas, do fácil acesso a essas bases (com o desenvolvimento das telecomunicações) e da crescente padronização que permite o desenvolvimento de novos indicadores; (ii) necessidade de avaliar a atividade científica, devido aos crescentes cortes de financiamento, das restrições econômicas e da exigência de avaliar a eficácia das políticas científicas enquanto resultados de investigação.⁶

Amante (2011) considerava, passados 12 anos, que a pesquisa bibliométrica se dirigia a três grupos-alvo: (i) bibliometria para especialistas, na investigação metodológica (metodologia); (ii) bibliometria para disciplinas científicas (informação científica); (iii) bibliometria para gestão da política de ciência (política científica). Para o que nos interessa nesta pesquisa, trata-se de três aspectos relevantes para a investigação do funcionamento de culturas disciplinares num contexto de forte produtividade científica.

A cientometria é um modo de medir a produção científica. Emerge da preocupação de avaliar diferentes campos científicos, seguindo a aplicação de técnicas numéricas analíticas. Silva e Bianchi (2001) observaram, já há 22 anos, o seguinte:

⁶ Do original: “(i) Razones técnicas, fundamentalmente debidas a la informatización de las bases de datos bibliográficos en todas las áreas, el fácil acceso a las mismas (desarrollo de las telecomunicaciones) y la creciente estandarización que permiten el desarrollo de los nuevos indicadores. (ii) Necesidad de evaluar la actividad científica, debido a los crecientes costes de la investigación y a las restricciones económicas y a la necesidad de evaluar la eficacia de las políticas científicas en cuanto a resultados de la investigación.” (GARCIA, 2007 *apud* AMANTE, 2011).

A partir da análise cuidadosa destes números, pode-se acompanhar a evolução ou o declínio de campos da ciência e também se podem identificar áreas emergentes que necessitam de maiores suportes financeiros ou de recursos humanos para melhor progredirem. (SILVA; BIANCHI, 2001, p. 6)

O uso de citação é a principal métrica quantitativa, pois seria um índice transparente e inequívoco do prestígio acadêmico-científico de seus autores. Amante (2011) menciona que o movimento em favor da adoção do número de citações teve início com a publicação do *Science Citation Index*, em 1961:

Segundo Glanzel, Debackere e Schubert (2006) a discussão sobre a utilização de indicadores de citação ganhou uma dimensão nova quando os indicadores bibliométricos, para além de serem usados para avaliar o desempenho da investigação institucional e nacional, passaram a integrar a fórmula de financiamento da investigação científica. (AMANTE, 2011, p. 4)

Dadas essas condições, a contagem de citações é consolidada como importante “indicador de valor científico”, pois refletiria a “qualidade da investigação” (AMANTE, 2011, p. 13). O número de citações recebidas por um artigo é entendido, na comunidade científica, como indicador da influência ou do impacto que seu conteúdo produziu num determinado grupo, para além das especificidades das diferentes culturas disciplinares. Se determinado texto é citado, é porque tem relevância – teórica, metodológica, analítica –, mesmo que tenha sido mobilizado para discussão por confronto de ideias.

Em conformidade com Silva e Bianchi (2001, p. 6), os indicadores bibliométricos utilizados por pesquisadores são responsáveis por mensurar o número de pessoas que recebem titulações acadêmicas; o número de artigos científicos publicados, o número de cientistas que publicam artigos científicos; a quantidade de referências citadas nos artigos científicos, de citações recebidas por artigo científico, de auxílios à pesquisa recebidos pelos cientistas e a quantidade de recursos destinados às atividades de pesquisa fomentadas pelas agências.

Silva e Bianchi (2001, p. 7) mencionam dois conjuntos de indicadores bibliométricos. O primeiro deles é constituído por indicadores quantitativos da atividade científica, aferida pelo número de publicações numa área de pesquisa e num determinado país de origem. De acordo com os autores, esses indicadores permitem comparar e analisar a quantidade de publicação científica em relação à de outros locais. O relatório *Clarivate Analytics*, produzido pelo grupo *Web of Science*, a pedido da CAPES, e publicado em 2019, apontava que o Brasil havia produzido 50 mil artigos na *Web of Science* e ocupava o 13º lugar no *ranking* de

produção científica mundial no período de 2013 a 2018, tendo colaborado com pesquisadores de mais de 205 países, no mesmo período. Trata-se, nos termos vistos em Silva e Bianchi (2001), de indicadores quantitativos da atividade científica.

O segundo conjunto mencionado por Silva e Bianchi (2001) é formado de indicadores de impacto. O FI, como discutido, visa avaliar a qualidade da publicação, com base, principalmente, no número de citações que o periódico recebe. Trata-se, bem entendido, da expressividade de um periódico numa comunidade, não se confundindo, portanto, com a expressividade de um pesquisador individual ou ainda de uma instituição. Atualmente, esse instrumento tornou-se também um guia político-científico, isto é, governos e agências de fomento o concebem como norteador para a alocação de recursos financeiros a pesquisadores e instituições (MARZIALE; MENDES, 2002, p. 466). O chamado *índice h* é outro instrumento bibliométrico, que objetiva quantificar a produtividade e o impacto científico baseado em artigos mais citados de cada pesquisador.

Esses medidores bibliométricos e, por conseguinte, a questão da citação, são decisivos para a aferição da qualidade e para a visibilidade de um periódico científico. Sobre a política de citação como maneira estratégica de visibilidade, pode-se dizer que

Obviamente, órgãos de maior divulgação e visibilidade são mais lidos, daí originam mais referências, criando competição de autores para neles verem publicados seus trabalhos e, portanto, ganharem maiores oportunidades de serem citados. De fato, para um autor, vencer numa competição dessas já seria motivo de orgulho pelo reconhecimento da qualidade de seu artigo. A maior visibilidade garante-lhe, além disso, *o possível prestígio de citações e de "impacto"*. A tendência é o fechamento de um círculo vicioso sobre a qualidade (e, pois, demanda) desses periódicos. (BICAS; ROTHER; BRAGA, 2002, p.2, grifos no original)

Segundo Melo Ribeiro (2017), tanto para os modelos bibliométricos, quanto para os cientométricos, os artigos científicos são objetos empíricos privilegiados nos estudos métricos. Ao longo dos anos, a investigação dos indicadores de citação passou a ser uma das mais importantes medidas de impacto da literatura científica.

De acordo com Silva e Bianchi (2001, p. 7), a publicação como principal meio comunicacional nesse contexto pode ser justificada de duas perspectivas: do ponto de vista cognitivo ou epistemológico, o artigo científico é eficaz, pois é possível avaliar, por meio dele, a qualidade da pesquisa e, conseqüentemente, acompanhar sua contribuição para o avanço científico numa comunidade; do ponto de vista social, a publicação é valorizada como

representativa de novas descobertas científicas, sendo considerada uma etapa importante do processo de investigação e da produção do conhecimento.

Na era do Sistema Eletrônico de Editoração de Revista (SEER), a comunicação por meio de periódicos científicos eletrônicos altera a distribuição e aumenta a visibilidade das produções. Salgado e Clares (2017, p. 30) afirmam que os artigos são uma função algorítmica recuperável na *web* e, para serem encontrados nas bases de dados, é preciso haver cliques e *downloads*. No início dos anos 2000, Lawrence (2001 *apud* GRUSZYNSKI; GOLIN; CASTELO, 2008, p.5) já apontava que a circulação eletrônica dos artigos, à época, aumentava cerca de 336% as citações *online* em relação à fonte impressa e endossava a busca por visibilidade e prestígio em um campo competitivo. De acordo com o relatório *Clarivate Analytics* (2019, p. 30), “publicações citam outras publicações. Esses links de citação crescem em redes e é provável que seus números estejam relacionados à importância ou ao impacto da publicação.”⁷ A análise de citações e a análise de conteúdo tornaram-se, portanto, procedimento comum à metodologia bibliométrica.

Numa conjuntura em que as métricas se valem da citação como medida para aferir o fator de impacto, a necessidade de os pesquisadores terem seus trabalhos citados por outros, ao longo de suas carreiras, fica mais evidente. Para Hyland (2003, p. 252), a autocitação emerge como uma consequência retórica desse cenário acadêmico marcadamente competitivo. As altas taxas de autocitação, entretanto, são contestadas em alguns campos de conhecimento e entendidas como forma duvidosa de autopromoção. Em vista dessas condições de produção do discurso acadêmico-científico, o fenômeno da autocitação aparece como maneira estratégica de relatar outros trabalhos de um mesmo autor, com o intuito de enfatizar suas contribuições e dar credibilidade ao seu trabalho numa comunidade científica (HYLAND, 2003, p. 257).

Assim, a discussão sobre autocitação tem sido objeto de diversos campos da ciência. Amante (2011, p. 7) afirma que ainda está em curso um debate sobre a interpretação do papel da autocitação no processo de comunicação científica. A autora aponta que há quem condene a utilização da autocitação por considerá-la uma forma artificial de inflacionar taxas de citação em bases de dados; os especialistas em bibliometria, entretanto, avaliam o uso da autocitação como um elemento constituinte das práticas da comunicação científica. Numa conjuntura em que há, por um lado, exigência de originalidade na produção científica e, por outro, dispositivos de averiguação dessa originalidade, por meio de ferramentas digitais de

⁷ No original: “Publications cite other publications. These citation links grow into networks, and their numbers are likely to be related to the significance or impact of the publication.” (CLARIVATE ANALYTICS, 2019).

detecção de plágio, a prática da autocitação aparece também como uma prática de preservação do histórico de trabalho do pesquisador. Para o que interesse a este trabalho, o fenômeno da autocitação é investigado no âmbito de culturas disciplinares distintas, como prática social acadêmico-científica que “responde” a determinadas condições sócio-históricas da produção e distribuição do conhecimento científico, como discutido na subseção seguinte.

1.3. O fenômeno da autocitação em culturas disciplinares distintas

A autocitação, comumente entendida como o ato de citar trabalhos anteriores sobre os quais o trabalho atual se fundamenta, tem sido discutida como fator crucial na produção e comunicação científica, em estudos bibliométricos e infométricos, no campo da Biblioteconomia, da Ciência da Informação e dos Estudos Linguísticos. Em bibliometria, a discussão sobre o fenômeno da autocitação é recorrente, pois, como apontado, esse campo é responsável por organizar e analisar quantitativamente a atividade científica por meio da aplicação de métodos estatísticos e matemáticos, bem como da avaliação sistemática da pesquisa, com o objetivo de respaldar decisões estratégicas sobre políticas de financiamento da pesquisa científica (AMANTE, 2011, p. 3).

Amante (2011, p. 3), em seu trabalho sobre o impacto da autocitação na investigação científica, busca abordar o fenômeno de duas perspectivas, teórica e aplicada. A autora apresenta uma síntese das origens, da evolução e dos objetivos da Bibliometria, considerando que, para os profissionais dessa área, trata-se de um assunto que fomenta discussões polarizadas. Como comentado, Amante (2011, p. 4) menciona que, por um lado, uma parte desses especialistas julga a autocitação como forma artificial de aumentar os índices de citação; por outro, como elemento constitutivo do processo de comunicação científica e, portanto, adequado ao contexto de sua divulgação.

Ao apresentar crítica em relação à autocitação, a autora baseia-se numa categorização de citações, concebida segundo uma sua natureza e função no texto científico. Segundo Moravcsik (s/d *apud* AMANTE, 2011, p. 5), essa classificação tem como objetivo avaliar e distinguir as citações necessárias das dispensáveis em materiais classificados por critério de qualidade. Os tipos de citação, segundo Moravcsik, citado por Amante (2011, p. 5), podem ser classificados em conceitual, orgânico, evolutivo e confirmativo/negativo:

- (a) citação conceitual, numa perspectiva teórica: trata de um conceito ou de uma teoria quando empregada direta ou indiretamente. Em termos metodológicos, esse

tipo de citação ocorre quando o conceito se apresenta como elemento que dá suporte à afirmação do autor;

- (b) citação orgânica essencial e orgânica superficial: a primeira se refere a citações que correspondem a teorias sobre as quais o tema do artigo se fundamenta; a segunda (superficial ou não-essencial) funciona, por sua vez, no caso de referências a procedimentos e métodos alternativos, considerada dispensável para o tratamento da temática do texto;
- (c) citação evolutiva: baseia-se em referências que contribuem para o desenvolvimento lógico de uma ideia ou do tema do artigo; pode abarcar também citações que referenciam trabalhos que tratam de procedimentos análogos, mesmo que esses não colaborem para o desenvolvimento lógico do tema;
- (d) citação confirmativa/negativa: tem como característica apoiar um argumento em duas circunstâncias, a saber, quando o autor cita outro autor com a finalidade de defendê-lo, ou quando o cita com a finalidade de se opor a uma ideia apresentada pelo autor citado.

Ainda sobre essa classificação, Zulueta Garcia (2007 *apud* AMANTE, 2011, p. 6) argumenta também que existem motivações distintas para a autocitação, que podem ser agrupadas da seguinte maneira:

- (a) razões sérias ou funcionais, que consistem em dar crédito a outro trabalho, identificar metodologias, introduzir o tema e situar outros antecedentes, apoiar afirmações, autenticar dados, identificar artigos-fontes nos quais uma ideia ou conceito foi difundido, disputar prioridade de uma descoberta, dentre outros;
- (b) razões estratégicas ou menos sérias (motivações sociais), que se baseiam em reconhecer os autores citados, promover trabalhos anteriores que não foram difundidos ou citados, referenciar trabalhos para obter benefícios, mencionar trabalhos selecionadas de acordo com as preferências editoriais dos periódicos, intercambiar referências com colegas para aumentar mutuamente o número de citações, dentre outros;
- (c) motivações políticas e ideológicas, julgadas como demonstração aparente da influência de um autor sobre outros. Segundo Zulueta Garcia (2007 *apud* AMANTE, 2011, p. 6), esse é o motivo pelo qual a análise de citações constitui-se em um instrumento para medir a rede de influência das ideias científicas, uma vez que “as citações e as ideias num texto não aparecerem isoladas, mas combinadas com outras citações e ideias”, sendo “possível aceder a

um sistema complexo de símbolos que o autor partilha com outros e que está em permanente mutação.”.

Com base nessas classificações, Amante (2011) conclui que a autocitação se enquadra no tipo “conceitual”, originado da “variação de hábitos de citação e de publicação de acordo com as diferentes disciplinas, da sobrerepresentação de artigos de caráter científico de acordo com as diferentes disciplinas, motivo pelo qual não são indicadas as comparações entre disciplinas” (AMANTE, 2011, p. 7).

O estudo de White (2001), por sua vez, no campo da Biblioteconomia e da Ciência da Informação, propõe investigar a *identidade científica* dos autores da área por meio das obras por eles citadas em trabalhos subsequentes. Para o pesquisador, a *imagem de citação do autor* é determinada pelo conjunto de autores citados em artigos. A autocitação, de acordo com White (2001, p. 89), afeta significativamente essa imagem, uma vez que o próprio nome do autor fará parte da lista de referências do próprio trabalho e, conseqüentemente, terá impacto em índices bibliométricos. White (2001, p. 89) defende que, embora a autocitação seja uma atividade muitas vezes confundida com certo tipo de “egoísmo” – ver, por exemplo, o que diz Zulueta Garcia (2007 *apud* AMANTE, 2011) a propósito de “razões estratégicas ou menos sérias” –, é também uma prática comum, pois integra a identidade do autor, o “núcleo do núcleo” de sua imagem diante da (sua) comunidade científica.

Nesse cenário, emergem duas formas de avaliar a intensidade da autocitação na produção científica. Lawani (1982 *apud* WHITE, 2001, p. 89), que defende a natureza heterogênea da autocitação, empresta os conceitos da taxonomia para classificar a autocitação em duas principais categorias, sincrônicas e diacrônicas, com o objetivo de avaliar a intensidade das ocorrências em trabalhos científicos. A autocitação sincrônica avalia a quantidade de autocitação de um autor em relação ao total de citações em sua obra; a autocitação diacrônica avalia a intensidade das autocitações de um autor em relação ao total de citações em suas obras.

Essa classificação, sobretudo, no que respeita à autocitação sincrônica, é constantemente mobilizada no campo científico, uma vez que as autocitações são consideradas segundo seu “potencial efeito de distorção que afeta sobre o valor das citações como medida de impacto” (COSTAS; VAN LEEUWEN; BORDONS, 2010, p. 518, tradução nossa).⁸ Para White (2001, p. 88), uma citação é considerada sincrônica quando alguém cita

⁸ No original: “potential distorting effect on the value of citations as a measure of impact.” (COSTAS; VAN LEEUWEN; BORDONS, 2010, p. 518).

um autor mais de uma vez em uma única obra, enquanto a diacrônica ocorre quando alguém cita o mesmo autor em trabalhos publicados em épocas diferentes. Aplicando esses conceitos para a prática de autocitação, Lawani (1982 *apud* WHITE, 2001, p. 89) argumenta que uma alta taxa de autocitação sincrônica não implica “egoísmo”, ao passo que uma taxa elevada de autocitação diacrônica, sim.

Da perspectiva dos letramentos acadêmicos, valemo-nos dos estudos de Hyland (2003) sobre autocitação, na assunção de uma posição distinta da de autores que criticam o excesso de autocitação em trabalhos científicos e a consideram como forma duvidosa de autopromoção. Para o autor, o fenômeno da autocitação pode ser concebido de dois pontos de vista: (i) um em que a autocitação é uma forma neutra de referenciar outros trabalhos que culminam num “tipo de egoísmo acadêmico”; (ii) um em que a autocitação pode ser concebida com uma maneira estratégica de relatar outros trabalhos de um mesmo autor, com o intuito de enfatizar suas contribuições e dar credibilidade ao trabalho na (sua) comunidade científica (HYLAND, 2003, p. 253).

Hyland (2003, p. 251) propõe pensar a autocitação de um ponto de vista amplo, distinguindo-a da automenção autoral. A autocitação, segundo o autor, é uma forma de referência a trabalhos anteriores, enquanto, a automenção é caracterizada pelo uso de pronomes de primeira pessoa (*eu, meu, minha, nós, nosso, nossa*) que referenciam, portanto, trabalhos conduzidos pelo autor do texto e termos que se referem ao autor. Tal como a autocitação, a automenção é, segundo Hyland, uma consequência relacionada a mudanças na avaliação e na aferição do fazer científico, constituindo uma “ameaça” às “convenções estabelecidas em relatórios impessoais” (HYLAND, 2003, p. 252).⁹ Dito de outro modo, com o autor, considerando uma convenção do fazer científico que pressupõe que a escrita científica seja “puramente empírica e objetiva”, não haveria espaço nesse fazer para a presença de agenciamento humano, isto é, haveria resistência ao uso da automenção, caracterizado pelo emprego de pronomes pessoais (HYLAND, 2003, p. 252).

Estudiosos citados por Hyland (2003, p. 252) consideram que evitar a automenção promoveria maior credibilidade ao autor do texto científico, com ênfase na objetividade, o que justificaria, em manuais de redação acadêmico-científica, a não recomendação do uso de pronomes pessoais. Para Hyland (2003, p. 251), tanto a automenção – “uso estratégico de referência pessoal” (contribuição ou reconhecimento individual) –, quanto a autocitação – forma retórica de referenciar trabalhos anteriores e informar que há produção científica

⁹ No original: “Another, closely related, consequence of the changes in science has been the threat posed to the established conventions of impersonal reporting.” (HYLAND, 2003, p. 252).

(reconhecimento institucional) – são consequências de um cenário competitivo, em que o uso desses mecanismos estratégicos constitui um modo de “sobrevivência”.

Esse estudo de Hyland (2003) se voltou a uma investigação das escolhas linguísticas feitas por autores de artigos científicos, as quais podem demonstrar aspectos importantes da construção do conhecimento e dos padrões que caracterizam relações disciplinares. O autor analisou 240 artigos de periódicos e 800 resumos de artigos científicos de oito diferentes áreas do conhecimento (Sociologia, *Marketing*, Filosofia, Linguística Aplicada, Biologia, Engenharia Mecânica, Engenharia Eletrônica e Física), com o objetivo de investigar como a automenção é empregada, no âmbito da autocitação, e de que maneira esses usos indicariam estratégias promocionais e práticas epistemológicas dessas disciplinas.

Os resultados apresentados por Hyland (2003, p. 255) permitiram observar que a autocitação foi amplamente empregada no conjunto de artigos analisados. O autor organizou os resultados de acordo com os padrões de autocitação em relação às disciplinas, mostrando o funcionamento das citações de acordo com cada área do conhecimento. Segundo os dados, quase 70% de todos os casos de autorreferência ocorreram em artigos de Ciências Humanas e Sociais, em comparação com apenas 17% em Engenharia. Outro dado relevante são as diferenças no uso de automenção e autocitação: 68% das automenções aparecem em artigos de Humanas (com alta incidência em artigos da área de *Marketing* e de Filosofia) e 60% de todas as autocitações nas chamadas *Hard Sciences* (HYLAND, 2003, p. 254). O estudo revelou, portanto, que em se tratando de autocitação e automenção há um desequilíbrio entre as Ciências Humanas e as Ciências Exatas. Esse resultado mostra, na avaliação do autor, diferentes maneiras de as disciplinas conduzirem pesquisas e persuadirem os leitores a aceitar resultados (HYLAND, 2003, p. 258).

Em Ciências Humanas, a maior ocorrência de empregos de pronomes pessoais está relacionada, na avaliação do autor, ao desejo de se apresentar como um pesquisador *informado*, uma estratégia de identificação com o leitor. Ainda sobre a automenção:

A automenção explícita aqui não enfatiza a credibilidade pessoal por meio de competências processuais, mas destaca o papel único do escritor na construção de uma interpretação plausível para um fenômeno, estabelecendo uma autoridade pessoal baseada na confiança e no comando de um argumento. (HYLAND, 2003, p. 258, tradução nossa)¹⁰

¹⁰ No original: “Explicit self-mention here does not stress personal credibility through procedural competencies but underlines the writer’s unique role in constructing a plausible interpretation for a phenomenon, establishing a personal authority based on confidence and command of an argument.” (HYLAND, 2003, p. 258).

Em estudo publicado em 2018, Hyland e Jiang se voltaram à investigação da autocitação, na observação da variação de sua quantidade em campos disciplinares distintos. Para os autores, essa variação provavelmente é determinada pela natureza dos trabalhos e de suas especificidades, como na subárea das Engenharias, por exemplo, em que os objetos de pesquisa necessitam ser avaliados em laboratórios (HYLAND; JIANG, 2018, p. 370). De acordo com os autores, a autocitação nesses contextos é inevitável, pois é através dela que há continuidade da pesquisa, requisito nessas áreas; por outro lado, as humanidades apresentam menores índices de autocitação, pois a ocorrência de trabalho cumulativo de pesquisa é menos expressiva (HYLAND; JIANG, 2018, p. 370).

Diferentemente de Hyland (2003), que considera a distinção entre os fenômenos da autocitação e da automenção, nesta Dissertação voltamos o olhar para a dinâmica da autocitação em dois domínios disciplinares distintos, as subáreas de Linguística e de Matemática. Essa escolha, que implica a exclusão da investigação de pronomes pessoais, é motivada pelo fato de considerarmos que, ainda que a presença de pronomes pessoais possa indiciar uma estratégia de identificação com o leitor, sua ausência pode igualmente indiciar um processo semelhante. Em outras palavras, a ausência de uso de pronomes pessoais é condizente com uma cultura disciplinar do domínio científico, em que a chamada “imparcialidade” – portanto, a tentativa de exclusão do agenciamento humano – deve ser construída. A exemplo de Hyland (2003), assumimos que a autocitação é a consequência retórica do crescimento da competitividade acadêmica, e que investigar padrões ou outras possíveis formas de como os autores se referem a si próprios, pode permitir compreender o importante papel da autocitação no processo de produção científica e no funcionamento das disciplinas, possibilitando, assim, averiguar a hipótese de que há prevalência de outras ciências em relação às Ciências Humanas, uma vez que indicadores bibliométricos corroborariam a contagem de (auto)citações.

Como dito, para o desenvolvimento deste trabalho, fundamentamo-nos na perspectiva sociocultural dos letramentos, mais particularmente, dos letramentos acadêmicos, no âmbito da investigação de um contexto institucional, disciplinar e mundializado, que considera relevante o “atravessamento” de aspectos como os de poder e autoridade na constituição da produção textual acadêmico-científico (LEA; STREET, 1998; 2014; LILLIS; CURRY, 2011). Com base nessa perspectiva, discutiremos a autocitação e seu funcionamento em artigos de periódico, buscando investigar práticas acadêmico-científicas em duas diferentes culturas disciplinares.

A propósito da distinção entre culturas disciplinares, Hyland (2006, p. 86), apoiando-se em Swales (1990), observa que diferentes disciplinas moldam diferentemente argumentos, na construção do conhecimento. Para Hyland (2006, p.18), “de fato, gênero e comunidade fornecem juntos uma estrutura descritiva e explicativa de como os significados são construídos socialmente, considerando as forças externas ao indivíduo que ajudam a orientar os propósitos e a moldar a escrita.” (HYLAND, 2006, p.18).¹¹

O conceito de *comunidade* mobilizado por Hyland (2006, p. 18) parece ser peça-chave para compreender a diferença entre disciplinas, uma vez que se volta à produção escrita e falada e à comunicação, auxiliando não apenas a interpretar e a compreender melhor o uso da linguagem, como também a explicar a variação em diferentes grupos disciplinares. Hyland (2006, p. 19) considera que, embora os discursos constituam o conhecimento e as atividades de determinada comunidade, os membros de um corpo social raramente se constituem numa massa homogênea, pois, ao adotar uma voz associada a determinado campo de estudo, os autores precisam se alinhar àquelas práticas de construção do conhecimento. As comunidades são, portanto, heterogêneas, formadas de sujeitos com experiências, origens, conhecimentos diversos, que podem decidir até que ponto compartilham de seus métodos, convenções e valores (HYLAND, 2006, p. 19).

O autor propõe o conceito de *comunidade de discurso*, baseada em um conjunto de textos e contextos mais amplos e socialmente informados. A comunidade de discurso permite compreender como o significado é produzido na interação e como contribui para a identificação de escolhas retóricas dos autores, na relação com o público. Para Hyland (2006, p. 21), a comunidade de discurso, quando associada à noção de disciplina, permite ao pesquisador avaliar convenções e práticas que influenciam determinada comunidade científica – caso deste trabalho em que são avaliadas duas distintas áreas de conhecimento.

Ainda no que se refere à disciplina, Hyland (2006, p. 21), com base em Becher (1989), avalia como cada uma se configura segundo um conjunto de convenções (de normas, categorizações, corpos de conhecimento) e modos de investigação que compreendem, portanto, uma cultura “reconhecível”. Segundo Hyland (2006, p. 22), no âmbito de cada disciplina, os sujeitos desenvolvem competências em discursos especializados, escrevendo como um membro da disciplina, “textualizando” o trabalho de modo que os demais membros dessa comunidade possam identificar determinado fazer científico. Para o autor:

¹¹ No original: “In fact, *genre* and *community* together provide a descriptive and explanatory framework of how meanings are socially constructed by considering the forces outside the individual which help guide purposes and shape writing.” (HYLAND, 2006, p. 18).

Assim, o discurso ajuda a criar uma visão disciplinar do mundo. Influenciada por práticas sociais e formas de pensamento numa disciplina, a escrita acadêmica envolve conjuntos de escolhas retóricas empregadas para estimular apoio, expressar cortesia, resolver dificuldades e evitar discordâncias de formas que correspondam mais de perto a pressupostos, métodos e conjuntos específicos de conhecimentos. (HYLAND, 2006, p. 21, tradução nossa)¹²

Dentre os fatores que contribuem para a diferença entre disciplinas está, segundo Hyland (2006, p. 25), o estilo da escrita. Pesquisadores citados por Hyland consideram que os textos das *Hard Sciences* são mais legíveis do que os das Ciências Sociais e os do campo das Artes. Outro fator pelo autor considerado é a citação, pois, por meio dela, determinado autor reconhece pesquisas anteriores, demonstra fidelidade a uma comunidade, cria uma “lacuna retórica” para sua pesquisa e estabelece um *ethos* de autor confiável (HYLAND, 2006, p. 25). O estudo de Hyland (2006, p. 26) mostra ainda que a frequência e o uso de citações diferem de acordo com o contexto, influenciados pelos modos como determinadas disciplinas concebem o mundo e realizam (as suas) pesquisas; além disso, argumenta que a citação auxilia a definir um contexto específico de conhecimento e contribui para um senso de progressão linear.

Neste trabalho, o objetivo principal é o de investigar padrões de autocitação em artigos científicos das Ciências Humanas (subárea de Linguística) e das Ciências Exatas (subárea de Matemática), tencionando contribuir com uma reflexão sobre o modo de funcionamento do discurso científico em culturas disciplinares distintas. No que respeita ao campo disciplinar da Matemática, McGrath e Kuteeva (2012) apresentam reflexões interessantes sobre essa cultura disciplinar. As autoras argumentam que, embora os estudos de Hyland (2006) ofereçam uma visão ampla de *corpus*, uma investigação mais aprofundada sobre práticas disciplinares de comunidades de discurso é necessária, sobretudo, no que se refere à Matemática Pura, considerando a escassez de trabalhos. Essa “negligência” seria

¹² No original: “Discourse thus helps to create a disciplinary view of the world. Influenced by a discipline’s social practices and ways of thinking, academic writing involves sets of rhetorical choices employed to galvanise support, express collegiality, resolve difficulties, and avoid disagreement in ways which most closely correspond to the community’s assumptions, methods, and bodies of knowledge.” (HYLAND, 2006, p. 21).

resultante, na avaliação das autoras, da percepção de que o discurso matemático é construído por meio de um “código padronizado” (HYLAND, 2005, p. 189 *apud* McGRATH; KUTEEVA, 2012, p. 165) e de disciplinas marcadamente teóricas (SWALES; FEAK, 2004 *apud* McGRATH; KUTEEVA, 2012, p. 162).

No âmbito das Ciências Exatas, McGrath e Kuteeva (2012, p. 162) consideram que o conhecimento em Matemática é gerado de maneira cumulativa e há critérios de raciocínio lógico, reconhecidos pela comunidade, para a averiguação desse conhecimento. Elas destacam ainda uma singularidade epistemológica nessa área na medida em que novos resultados são consubstanciados por provas, enquanto nas Ciências Naturais o resultado da pesquisa é “experimentalmente falsificável”. Segundo as autoras, a Matemática Pura não é facilmente classificada como empírica; contrariamente às Ciências Naturais, o resultado não depende da interpretação de dados, sendo essencialmente limitada à questão do verdadeiro ou falso (McGRATH; KUTEEVA, 2012, p. 162).

De acordo com as autoras, o objetivo da matemática é o de alcançar a simplicidade e a generalização, reduzindo a explicação que fundamenta os resultados. A questão apresentada por elas, diante das características dessa área, é: “como essa visão do conhecimento e das diferentes práticas de pesquisa associadas a uma disciplina baseada em provas impactam as práticas de escrita da comunidade de discurso da matemática pura?”¹³ (McGRATH; KUTEEVA, 2012, p. 162). As autoras consideram que a pesquisa sobre a produção escrita em Matemática é limitada. Mencionam os estudos de Swales *et al.* (1998), cujo enfoque eram os “imperativos” em Matemática. Referem-se ainda a O’Halloran (2005), que explorou a linguagem e o simbolismo no discurso matemático e, por fim, a Shaw (2006) que aponta como o “código matemático está embutido no texto em gêneros educacionais em três disciplinas (matemática pura, engenharia e física).” (McGRATH; KUTEEVA, 2012, p.162).

Interessa-nos nesse estudo, refletir sobre o funcionamento disciplinar da subárea da Matemática, em virtude do cumprimento de uns dos objetivos específicos: o de estudar padrões de autocitação de subáreas distintas. O conjunto de artigos selecionado é referente ao campo da chamada Matemática Aplicada. Na investigação de artigos de outras áreas de conhecimento, é preciso que o pesquisador em estudos linguísticos consiga “ler” os textos, para além de uma sua codificação. Superar esse obstáculo foi um dos desafios deste trabalho, que pretende ser uma contribuição para o estudo do campo.

¹³ Do original: “How does this view of knowledge and the different research practices associated with a roof-based discipline impact the writing practices of the pure mathematics discourse community?” (McGRATH; KUTEEVA, 2012, p. 162).

1.4. O estudo da autocitação como modo de referência

A investigação da autocitação e de seu funcionamento em artigos de periódico de duas distintas áreas de conhecimento contará, de uma perspectiva dos estudos de letramentos, com o aporte teórico dos estudos de modos de referência fundamentados em Boch e Grossmann (2002) e na releitura feita por Daunay e Delcambre (2017) e Daunay (2020).

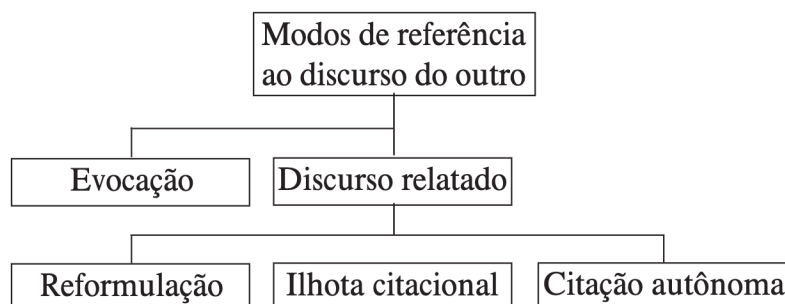
Boch e Grossmann (2002) apresentam em seu estudo critérios que permitem diferenciar os modos de referenciar o discurso do outro. Fundamentam-se, para tanto, em um *corpus* constituído de artigos científicos produzidos por especialistas (pesquisadores em Linguística) e noutro de relatórios de estágio produzidos por universitários. Para os autores, além da citação, há, na produção da escrita retórica, outras maneiras de referenciar o discurso de outrem, segundo *evocação* ou *discurso relatado* (com reformulação, ilhota citacional ou citação autônoma).

A *evocação*, segundo Boch e Grossmann (2002), é uma forma de convocar o dizer do outro, assinalada pela presença de nome próprio do autor, seguido do ano de publicação do artigo. É marcada pela ausência de marcas introdutórias (*segundo X, de acordo com X*) e de desenvolvimento temático.

“No discurso reportado, o escritor indica que resume, reformula ou cita o discurso do outro.” (BOCH; GROSSMANN, 2002, p. 100). Assim, pode ser subdividido em reformulação, ilhota citacional ou citação autônoma. Para Boch e Grossmann (2002), a *reformulação* possibilita ao escritor integrar o discurso do outro em seu próprio, servindo para marcar uma apropriação implícita do ponto de vista presente no discurso do outro. Consequentemente, permite que o escritor do texto não se “perca” no fio da análise. Diferentemente da citação, a reformulação não apresenta marcas introdutórias como as aspas, por exemplo. Por outro lado, é marcada pela presença de indicativos introdutórios do discurso reportado e pela ausência de marcas escriturais, como as já citadas aspas, ou ainda as verbais (*eu cito X, para retomar as palavras de X*). A ilhota citacional, por sua vez, “permite a integração quanto a colocação em evidência do segmento citado, pela marca escritural, graças ao itálico e às aspas.” (BOCH; GROSSMANN, 2002, p. 100). Na avaliação de Boch e Grossmann (2002), a citação é o modo de referência menos empregado por pesquisadores experientes. Considerada por criar um “espaço autônomo no plano enunciativo” (BOCH; GROSSMANN, 2002, p. 100), caracteriza-se pela presença de marcas escriturais, como aspas, itálico ou bloco tipográfico, que permitem identificar um segmento do texto extraído de um texto fonte. A Figura 2 e o

Quadro 1 foram extraídos do artigo de Boch e Grossmann (2002) e mostram esquematicamente a proposta dos autores:

Figura 2– Modos de referência ao discurso do outro



Esquema 1: Tipologia dos modos de referência ao discurso do outro

Fonte: Boch; Grossmann (2002, p. 100)

Quadro 1 – Critérios que permitem diferenciar os modos de referência ao discurso do outro

Quadro 1
Critérios que permitem diferenciar os modos de referência ao discurso do outro

Evocação	Reformulação	Citação
- Ausência de marcas introdutórias de discurso reportado (tais como: <i>segundo X</i>, <i>como afirma X</i>, ou equivalentes). - Ausência de desenvolvimento temático do dizer do outro. - Presença de um nome próprio de autor, frequentemente com data à qual o autor do artigo se refere, sem precisar o teor do texto.	- Presença de marcas introdutórias do discurso reportado (<i>segundo X ...</i>, <i>de acordo com X ...</i>, <i>para X ...</i>, <i>como X afirma ...</i>, <i>como X pretende ...</i>, etc.). - Ausência de marcas escriturais tais como aspas (ou verbais, como <i>eu cito X</i>, <i>para retomar as palavras de X</i>). - O discurso do outro é integrado no discurso de quem escreve e não tem autonomia enunciativa.	- Marcas, geralmente escriturais, como aspas, itálico ou bloco tipográfico, permitem identificar um segmento do texto como extraído de uma fonte externa; essas marcas podem, às vezes, ser substituídas por comentários metalingüísticos: <i>eu cito X ..., para retomar as palavras de X ...</i>); - Autonomia enunciativa do segmento citado (salvo no caso da “ilhota citacional”).

Fonte: Boch; Grossman (2002, p. 101)

Recorremos também à tipologia adaptada por Daunay e Delcambre (2017 *apud* DAUNAY, 2020), cuja proposta retoma aquela de Boch e Grossmann (2002). Daunay e

Delcambre (2017) optam por dividir o conceito de “ilhota citacional”, de Boch e Grossmann (2002), em duas formas:

- (1) *empréstimo*: constitui-se por uma citação curta (uma palavra ou um sintagma);
- (2) *hibridação*: assemelha-se ao empréstimo, porém, localizada em uma reformulação.

Dessa perspectiva, em termos metodológicos, organizamos o conjunto do material formado de artigos de periódicos das subáreas de Linguística e de Matemática em termos de *modalidades enunciativas* (DAUNAY, 2020; DAUNAY; DELCAMBRE, 2017). Essas nomenclaturas foram desenvolvidas a partir dos casos de menção ao discurso do outrem e para os casos de inserção do discurso próprio, tal como a autocitação se constitui. Neste trabalho, no emprestamos essas tipologias como um modo de classificar o posicionamento retórico discursivo no corpo do texto.

Capítulo 2 – Material e Métodos

Esta seção está organizada em quatro subseções. Na primeira delas, é apresentada a justificativa para a escolha das subáreas de Linguística e Matemática. Em seguida, buscamos justificar a escolha das plataformas de bases de dados utilizadas. Os critérios para a composição do conjunto do material e para a constituição do *corpus* são tratados em seguida. Por fim, são apresentados os procedimentos metodológicos da organização e os critérios para análise dos dados.

2.1. A escolha das subáreas de Linguística e Matemática

A justificativa das grandes áreas e subáreas selecionadas para este trabalho está fundamentada, em parte, no relatório *Clarivate Analytics*, produzido pelo Grupo *Web of Science* (WoS) a pedido da CAPES e publicado no ano de 2019. Esse relatório aponta que, no período de 2013 a 2018, a grande área de Ciências Exatas e da Terra (na qual a subárea de Matemática se encontra) foi a mais produtiva em publicação de artigos indexados na WoS, com maior número de colaboração em nível internacional. A grande área da Linguística, Letras e Artes (na qual a subárea de Linguística se encontra), por sua vez, é considerada, pelo relatório, uma das áreas de pesquisa mais frágeis em termos de colaboração internacional. Esse fato seria explicado pela língua de publicação dos periódicos da área (língua portuguesa), com número limitado de leitores/pares interessados.

Em *Avaliação da Pesquisa em Humanidades*, relatório publicado pela CAPES em 2019, pontua-se que uma das características dessa grande área é a publicação no idioma local, em decorrência das “diferentes culturas nacionais de intelectuais nas áreas de humanas”. Esse fato pode justificar a dificuldade da área de se firmar como “ciência” na competição por apoio financeiro. Em relação aos recursos financeiros destinados, especificamente, à área de Linguística, Letras e Artes, o *Diagnóstico das Ciências Humanas, Sociais, Sociais Aplicadas, Linguística, Letras e Artes* (CHSSALLA, 2019),¹⁴ mostra que, no período de 2006 a 2016, a área obteve o menor índice de financiamento de bolsas para a pesquisa em relação a outras áreas das Humanidades. Diante desse cenário, em que a subárea de Linguística – na qual este

¹⁴ Disponível em:

https://www.anpepp.org.br/boletimartigo/view?ID_BOLETIM=12&ID_BOLETIM_ARTIGO=106. Acesso em: 09 mar. 2023.

trabalho se filia, por ser realizado em Programa de Pós-graduação em Estudos Linguísticos – apresenta-se como mais atingida por falta de reconhecimento institucional em relação à subárea de Matemática, por exemplo, considera-se relevante investigar como o processo de autocitação atua e corrobora a contagem de citações em indicadores bibliométricos, numa conjuntura em que se busca visibilidade na produção científica.

Além desse critério institucional fundamentado em relatórios institucionais, há outra justificativa para a escolha das subáreas de Linguística e Matemática. Esta Dissertação de Mestrado está inserida em projetos de pesquisa mais amplos, já mencionados, que buscam investigar o processo de escrita acadêmico-científica. Na divisão do trabalho realizado em grupo, as subáreas de Linguística e Matemática nos foram atribuídas, diante de outras subáreas investigadas (Ciências Biológicas, Ciências Contábeis, Direito, Educação, Física, Química, Saúde, dentre outras). Também nos foi atribuída a investigação de artigos de periódicos em língua portuguesa, em meio à possibilidade de investigação de artigos científicos em outras línguas (espanhol, francês, inglês).

Considerando-se que um dos critérios de fator de impacto é o tempo de cinco anos (*5-year journal impact factor*),¹⁵ tomou-se a decisão de que o período de coleta de conjunto do material seria de cinco anos, no intervalo de 2015 a 2019. O conjunto do material foi coletado assim no ano de 2020, ano do início do Curso de Mestrado. Os artigos foram selecionados de acordo com o número de citações, segundo indicadores bibliométricos das plataformas, como explicitado adiante.

2.2. A escolha das plataformas de bases de dados

Dada a importância da *Scientific Electronic Libray Online* (doravante *SciELO*) e *Web of Science* (doravante *WoS*) para a indexação e publicação na internet de uma grande coleção de periódicos científicos – considerando-se o objetivo de aumentar visibilidade, acessibilidade, uso e impacto, como discutido na primeira seção desta Dissertação –, a escolha dessas plataformas é justificada por seu caráter multidisciplinar e por apresentarem informações acerca do fator de impacto (FI) de periódicos, artigos, com seleção desses artigos por ordem de maior número de citações.

Sobre as plataformas, de acordo com relatório *Clarivate Analytics* (2019), as colaborações internacionais do indexador *SciELO* têm crescido de forma constante nos

¹⁵

Disponível

em:

<http://help.prod-incites.com/inCites2Live/indicatorsGroup/aboutHandbook/usingCitationIndicatorsWisely/fiveYearJIF.html>. Acesso em: 09 mar. 2023.

últimos anos e chegou a 10% de todas as publicações em 2018. No que se refere a WoS, destacamos que se trata de um serviço *online* de indexação de citações científicas baseado em assinatura, originalmente produzida pelo *Institute for Scientific Information* (ISI) e atualmente segurado pela *Clarivate Analytics* (antiga Instituição *Thomson Reuters*). Segundo informações que constam no *site* da *Clarivate*,¹⁶ a WoS é um provedor de informações e tecnologia destinado à comunidade científica mundial, com mais de 12 milhões de artigos em cinco mil periódicos de acesso aberto. Essa plataforma fornece aos pesquisadores “dados, análises e *insights*”, por meio de ferramentas e informações bibliográficas publicadas nos periódicos indexados em diversas áreas do conhecimento. As informações que constam no *site* da WoS, na aba *Web of Science Core Collection*,¹⁷ mostram que o número de periódicos indexados nessa Coleção chega a 21.100, em mais de 250 áreas do conhecimento. Além disso, a plataforma fornece, por meio da ferramenta *Citation Reports*, relatórios de citações, que incluem uma exibição gráfica de publicações e citações organizadas por período, com indicação do índice h.

Já a *SciELO*,¹⁸ plataforma cooperativa de periódicos científicos na internet, é resultado de um programa da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) e do Centro Latino-Americano e Caribenho de Informação em Ciências da Saúde, também conhecido por seu título original, Biblioteca Regional de Medicina (BIREME).¹⁹ A partir de 2002, a *SciELO* passou a contar com o apoio do CNPq.

O *SciELO Citation Index* (*SciELO CI*), integrado à plataforma WoS, por sua vez, é fruto da parceria do Programa *SciELO/FAPESP* com a Thomson Reuters, proprietária da WoS. Em termos funcionais, o *SciELO CI* compartilha dos mesmos recursos, filtros e navegabilidade da interface da WoS. No entanto, segundo informações apresentadas no *site* da *Clarivate Analytics*,²⁰ apenas na funcionalidade integrada à base do WoS é possível rastrear o impacto das citações em artigos indexados e ter acesso aos “mais de 1,5 bilhão de conexões de citações nas ciências, ciências sociais, artes e humanidades.”.

¹⁶ Disponível em: <https://clarivate.com/webofsciencegroup/about-us/>. Acesso em: 09 mar. 2023.

¹⁷ Disponível em: <https://clarivate.com/webofsciencegroup/solutions/web-of-science-core-collection/>. Acesso em: 09 mar. 2023.

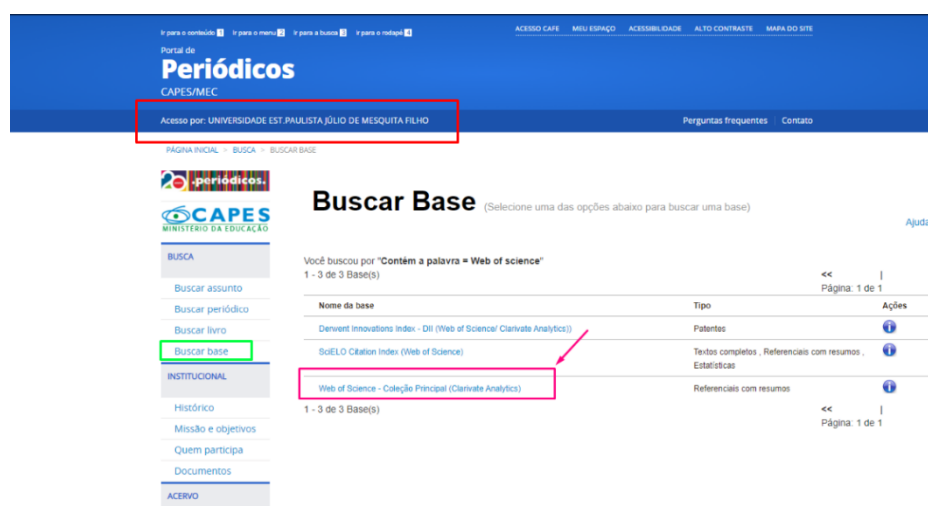
¹⁸ Disponível em: <https://www.SciELO.org/pt>. Acesso em: 09 mar. 2023.

¹⁹ Disponível em: <https://www.paho.org/pt/bireme>. Acesso em: 09 mar. 2023.

²⁰ Disponível em: <https://clarivate.com/webofsciencegroup/solutions/webofscience-SciELO/>. Acesso em: 09 mar. 2023.

Tanto a plataforma *WoS* quanto a *SciELO CI* foram acessadas através do Portal de Periódicos da CAPES,²¹ por meio da Comunidade Acadêmica Federada (CAFE).²² Esse acesso remoto se faz necessário, pois permite que o usuário busque conteúdos exclusivos de assinantes, caso da universidade à qual estamos filiados (Unesp). Por meio da concessão permitida pela CAPES, o usuário com *login* do *e-mail* Unesp tem, portanto, acesso a conteúdos que não são gratuitos. Após realizar o *login* na plataforma, o acesso à *WoS* pode ser realizado, num primeiro momento, pela categoria “base” e, em seguida, por “título”. A Figura 2 a seguir mostra o passo a passo de acesso à *WoS*: em destaque em vermelho, o *login* na CAFE; em verde e rosa, a busca por base:

Figura SEQ Figura 1* ARABIC 3 – Busca no Portal da plataforma CAPES



Fonte:

elaborado pela autora.

Dentre as palavras-chave pesquisadas, foram escolhidas: *Linguística* e *Matemática*, respectivamente. Seleccionamos os filtros:

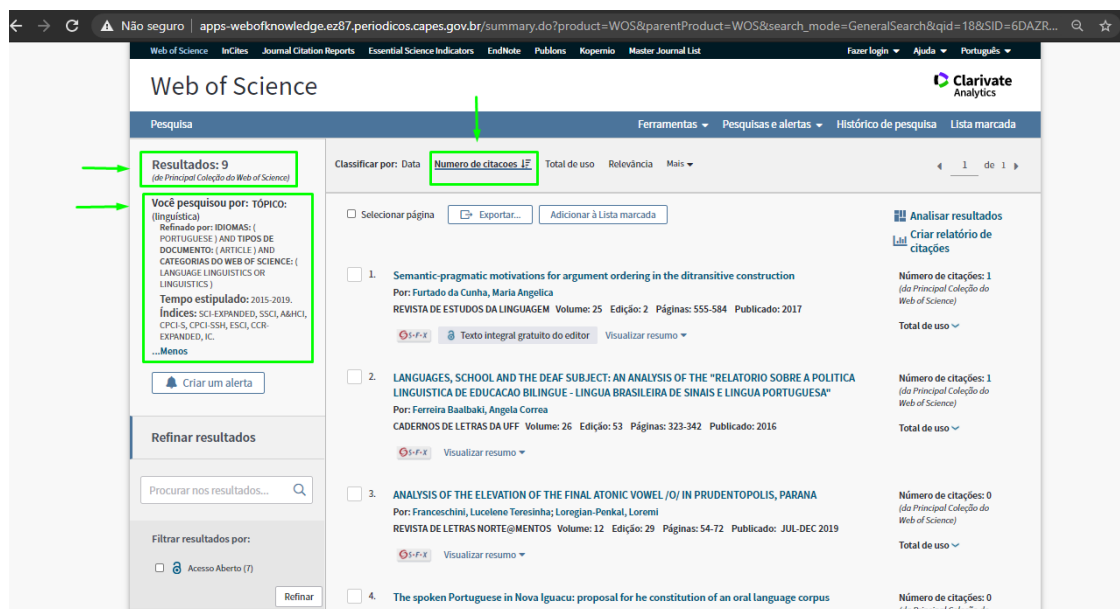
- (i) idioma: português;
- (ii) tipo de documento: artigo;
- (iii) categorias: Linguística; Matemática;
- (iv) tempo estipulado: 2015-2019.

²¹ Disponível em: <http://www.periodicos.capes.gov.br/>. Acesso em: 09 mar. 2023.

²² Disponível em: http://www-periodicos-capes-gov-br.ez1.periodicos.capes.gov.br/?option=com_plugin&ym=3&pds_handle=&calling_system=primo&institute=CAPES&targetUrl=https://www-periodicos-capes-gov-br.ez1.periodicos.capes.gov.br/&Itemid=155&pagina=CAFe. Acesso em: 09 mar. 2023.

Os artigos foram ordenados do mais citado para o menos citado. A seleção dos filtros pode ser conferida na Figura 4:

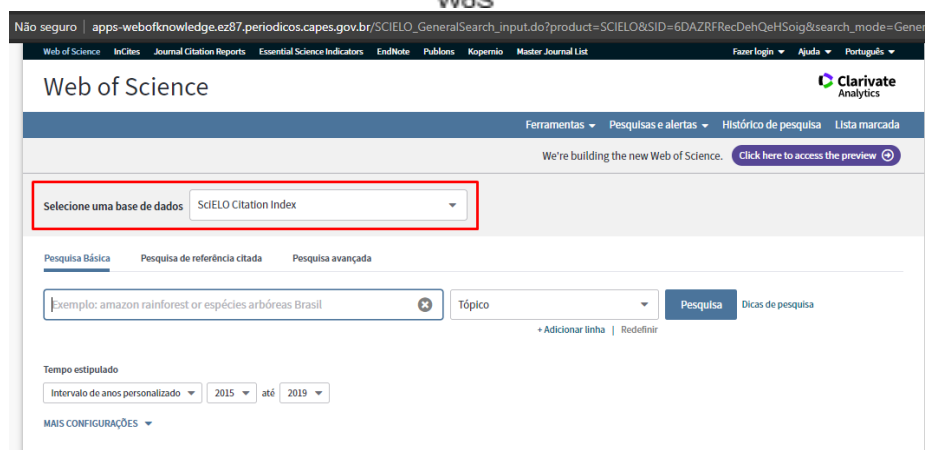
Figura SEQ Figura 1* ARABIC 4 – Seleção dos filtros e ordenação dos artigos



Fonte: elaborado pela autora.

Para acessar o *SciELO Citation Index*, por meio da WoS, é necessário seguir os passos iniciais descritos. Foi selecionado “*SciELO Citation Index*” no campo de busca, conforme ilustra o destaque em vermelho na Figura 5, a seguir:

Figura SEQ Figura 1* ARABIC 5– Acesso à plataforma SciELO através da WoS



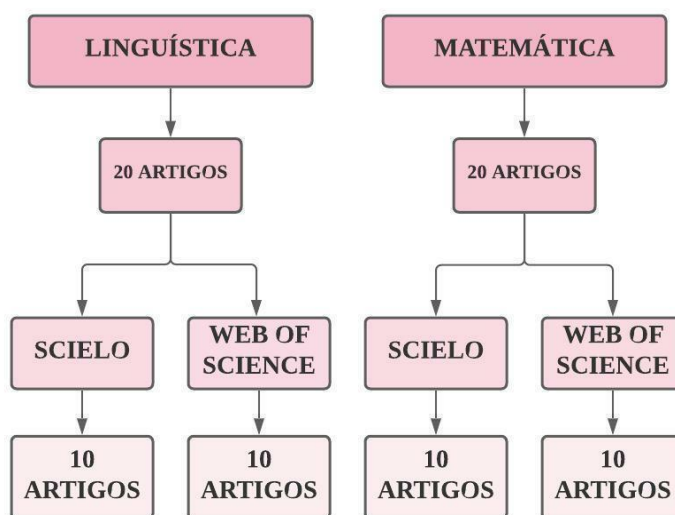
Fonte: elaborado pela autora.

Foram utilizados os mesmos filtros e critérios para a coleta dos artigos científicos das duas subáreas em estudo, Linguística e Matemática (idioma: português; tipo de documento: artigo; tempo estipulado: 2015-2019).

2.3. O conjunto do material e a constituição do *corpus*

O conjunto do material foi formado de um total de 40 (quarenta) artigos científicos, 20 (vinte) de cada subárea (Linguística e Matemática), 10 (dez) de cada uma das plataformas escolhidas (*SciELO* e *WoS*). A organização dos artigos científicos, em função das subáreas de conhecimento e das plataformas de base de dados, deu-se da seguinte maneira:

Figura SEQ Figura * ARABIC 6 – Organização do conjunto do material



Fonte: elaborado pela autora.

A partir desse conjunto mais amplo de textos, foram escolhidos 27 (vinte e sete) artigos científicos das subáreas de Linguística e Matemática, publicados entre janeiro de 2015 e dezembro de 2019, nas já referidas bases. No processo da coleta, constatou-se uma inconsistência. Parte dos artigos apresentados pelas plataformas advinha de subáreas distintas daquelas indicadas no critério de seleção. Isso se deu, provavelmente, pelo fato de que a seleção automática levou em conta estudos interdisciplinares ou ainda palavras-chave que remetem aos campos, sem necessariamente a eles pertencerem. Vamos explicar de maneira circunstanciada essa inconsistência, com base no conjunto de material coletado junto à WoS para a subárea de Matemática.

A busca por artigos científicos da subárea da Matemática, na WoS, indicou, num primeiro momento, 11 (onze) resultados, conforme destacado na Figura 7 a seguir. Os artigos listados já estavam organizados por maior número de citações. Ao analisar cada um dos artigos, constatou-se que três (03) deles não pertenciam à subárea indicada, mas a outras, caso da Informática, por exemplo.

Figura SEQ Figura 1* ARABIC 7 - Busca por artigos da subárea de Matemática

Fonte: elaborado pela autora.

Os resultados obtidos mostram uma prevalência de trabalhos com “aplicações interdisciplinares da Matemática” (oito dos 11 artigos localizados). O resultado final obtido nessa busca é ilustrado na Figura 8, seguinte:

Figura 8 – Resultado final da busca de artigos na subárea Matemática, plataforma WoS

Essa inconsistência resultou na constituição de um *corpus* com quantidade menor de artigos do que inicialmente previsto, tanto para a subárea de Matemática, quanto para a de Linguística. Observa-se que os artigos coletados na subárea de Matemática e que foram apresentados pelas bases são do campo de Matemática Aplicada e não de Matemática Pura. Essa prevalência pode ser explicada pela seleção do idioma (língua portuguesa) nas bases, o que levou ao encontro de artigos científicos em Matemática Aplicada, mas não em Matemática Pura, que tradicionalmente publica em línguas estrangeiras (língua inglesa, principalmente). A questão que se assenta sobre esse achado da pesquisa é: como essa organização pode impactar nos indicadores bibliométricos da área da Matemática, sobretudo, em relação à Linguística?

A seguir, apresentamos a relação dos artigos coletados em função das subáreas e das plataformas de bases de dados, organizada em quadros para a melhor visualização. O Quadro 2 mostra informações gerais dos 10 (dez) artigos de Matemática Aplicada coletados junto à plataforma *Web of Science*. Os artigos foram codificados da seguinte maneira: MWoS-n indica, respectivamente, artigo da subárea de Matemática (Aplicada) – “M” – coletado junto à plataforma – “WoS” –, e o número do artigo – “n” – no conjunto do material. Esse número foi distribuído aleatoriamente.

Quadro 2 – Artigos da subárea de Matemática Aplicada coletados junto à plataforma Web of Science

Código	Título	²³ Subárea	Número de páginas	Número de autores	Ano de publicação	Formação acadêmica dos autores	Qualis do periódico (2013-2016)	Quantidade de citações na base (até 2019)
MWoS-1	Dyscalculia: A Brazilian Articles Mapping	Matemática Aplicada	21	4	2017	não constam	-	2
MWoS-2	Dynamic Mathematics to understand the exponential function Teaching	Matemática Aplicada	10	3	2016	não constam	C	1
MWoS-3	Combinatory Analysis by Mobile Technology: an experience report	Aplicações Interdisciplinares da Matemática	12	2	2019	não constam	C	0
MWoS-4	Application of the IRT Nominal Response Model to	Aplicações Interdiscipli	6	4	2019	não constam	C	0

²³ Tanto as subáreas que dizem respeito aos artigos de Matemática quanto as de Linguística, aqui apresentadas em seus respectivos quadros, foram nomeadas e classificadas pela lógica de ambas plataformas de dados, *Scielo Ci e Web of Science*.

	Large-Scale Educational Assessment	nares da Matemática						
MWoS-5	Contributions of Technoracy and Ethnomathematics to the Promotion of Financial Education for Deaf Students	Aplicações Interdisciplinares da Matemática	14	2	2019	não constam		
MWoS-6	Problem Solving and Linear Systems: a case study in high school involving a simplified scheduling method	Matemática Aplicada	8	1	2017	não constam	C	0
MWoS-7	Redefinition of notable products using concrete material	²⁴ Aplicações Interdisciplinares da Matemática	13	2	2016	não constam	-	0
MWoS-8	Training of Initial Grades Teachers and development of Mathematics Education: Mapping of some Brazilian productions	Aplicações Interdisciplinares da Matemática	14	2	2016	não constam	-	0
MWoS-9	Descriptions of Free and Freeware Software in the Mathematics Teaching	Aplicações Interdisciplinares da Matemática	16	3	2016	não constam	-	0
MWoS-10	Mathematic literacy and training of the early years teacher: The mapping of scientific production	Matemática Aplicada	11	2	2015	não constam	-	0

Fonte: elaborado pela autora.

Na constituição do *corpus*, foram excluídos os artigos MWoS-1, MWoS-2 e MWoS-10, por não terem relação com a subárea de Matemática (Aplicada ou Pura).

No que respeita ao conjunto de artigos da subárea de Matemática coletados na plataforma *SciELO CI* (identificada como Sci na codificação do artigo, em observação ao título e ao tema de cada um, notamos que quatro (04) deles, nomeados como MSci-3,

²⁴ Para esta análise, mantivemos tanto o artigo MWoS-8 quanto seu subsequente MWoS-9, os quais foram classificados pela base de dados como pertencentes ao campo “Aplicações Interdisciplinares da Matemática”, pois consideramos que ambos estão ainda sob o domínio da “Educação em Matemática” e, consequentemente, da subárea da Matemática Aplicada.

MSci-4, MSci-5 e MSci-9, não pertenciam ao campo da Matemática, mas, sim, de disciplinas como Enfermagem e Tecnologia, os quais foram, portanto, excluídos do *corpus*. Somente seis (06) desses artigos foram selecionados: MSci-1, MSci-2, MSci-6, MSci-7, MSci-8 e MSci-10. Abaixo, listamos todos os artigos coletados em Quadro 3, seguinte:

Quadro 3 – Artigos da subárea de Matemática Aplicada coletados junto à plataforma SciELO CI

Código	Título	Subárea	Número de páginas	Número de autores	Ano de publicação	Formação acadêmica dos autores	Qualis do periódico (2013-2016)	Quantidade de citações na base (até 2019)
MSci-1	A análise de conteúdo como uma metodologia	Matemática Aplicada	22	2	2017	não constam	B5	9
MSci-2	O Estudo de Aula como Processo de Desenvolvimento Profissional de Professores de Matemática	Matemática Aplicada	23	4	2016	não constam	A1	8
MSci-3	Lee Shulman: Contribuições para a investigação da formação docente em enfermagem e saúde	Enfermagem	-	6	2017	não constam	-	6
MSci-4	Liderança do diretor, clima escolar e desempenho dos alunos: qual a relação?	Educação	20	2	2016	não constam		6
MSci-5	Mineração em bases de dados do INEP: uma análise exploratória para nortear melhorias no sistema educacional brasileiro	Educação	24	2	2016	não constam	B5	6
MSci-6	Necessidades de aprendizagem de alunos da Educação Profissional de Nível Técnico em Enfermagem	Enfermagem	5	4	2015	não constam	-	5
MSci-7	A articulação entre a avaliação somativa e a formativa, na prática pedagógica: uma impossibilidade ou um desafio?	Matemática Aplicada	32	1	2016	não constam	-	4
MSci-8	Seres humanos - com - internet ou internet - com - seres humanos: uma troca de papéis?	Matemática Aplicada	25	1	2016	não constam	A2	4
MSci-9	O impacto da criação da Funpresp sobre os benefícios previdenciários	Administração Pública	26	2	2015	não constam		4

dos servidores públicos
federais

MSci-10	Insubordinação Criativa: um convite à reinvenção do educador matemático	Matemática Aplicada	16	2	2015	não constam	A1	4
---------	---	------------------------	----	---	------	----------------	----	---

Fonte: elaborado pela autora.

O *corpus* de Matemática formou-se, portanto, dos artigos coletados nas plataformas *WoS*, os quais somaram-se sete (7) e *Scielo CI*, seis (6). Totalizando, assim, treze (13) artigos.

Já a constituição do *corpus* da subárea da Linguística, como já mencionado, também foi realizada por meio da coleta em duas plataformas, *WoS* e *SciELO CI*. Na plataforma *WoS*, foram coletados nove (09) artigos no período investigado. Dois artigos desse conjunto foram excluídos do *corpus*. Essa exclusão justifica-se, pois o artigo LWoS- 5 trata de um dossiê comemorativo dos 30 anos de um Grupo de Trabalho em Teoria da Gramática e da Associação Nacional de Pesquisa e Pós-graduação em Linguística (ANPOLL) e, por isso, não foi concebido como artigo científico. O artigo LWoS- 8, por sua vez, foi publicado em língua espanhola, fora, portanto, do critério do idioma língua portuguesa. O *corpus* em Linguística, junto à *WoS*, resultou em 7 (sete) artigos, como indicado no Quadro 4 a seguir:

Quadro 4 – Artigos da subárea de Linguística coletados juntos à plataforma Web of Science

Código	Título	Área	Número de páginas	Número de autores	Ano de publicação	Formação acadêmica dos autores	Qualis do periódico (2013-2016)	Quantidade de citações na base (até 2019)
LWoS-1	Motivações semântico-pragmáticas para a ordenação dos argumentos na construção ditransitiva/ Semantic-pragmatic motivations for argument ordering in the ditransitive construction	Linguística	30	3	2015	doutor/doutor/ doutora	B1	1
LWoS-2	Línguas, Escola e sujeito surdo: Análise do “Relatório sobre a política Linguística de Educação Bilingue - Língua Brasileira de Sinais e Língua Portuguesa”	Linguística Aplicada	21	1	2016	doutora	B1	1
LWoS-3	Análise de elevação da vogal átona final /o/ em Prudentópolis	Linguística	19	2	2019	doutora/douto r	-	0
LWoS-4	The spoken Portuguese in Nova Iguaçu: proposal for he constitution of an oral language corpus	Linguística	25	1	2019	doutor	B1	0
LWoS-5	30 Anos de teoria da gramática na ANPOLL: Uma apresentação	Dossiê Histórico		1	2018	doutora		0
LWoS-6	Galego e Português Brasileiro história, variação e mudança	Linguística	29	-	2017	-		0
LWoS-7	A cientificidade da linguística inatingível	Linguística	39	2	2017	doutora/mestr e		0
LWoS-8	ASinEs: Prolegómenos de un atlas de la variación sintáctica del español	Linguística		2		doutora/douto r		0
LWoS-9	Palavras para fazer ouvir interrogações	Linguística	19	1	2015	doutor	B1	0

Fonte: Elaborado pela autora

Já os artigos coletados junto à plataforma de dados *SciELO CI* totalizaram nove (09) textos na subárea de Linguística, conforme Quadro 5 a seguir:

Quadro 5 – Artigos da subárea de Linguística coletados junto à plataforma SciELO CI

Código	Título	Área	Número de páginas	Número de autores	Ano de publicação	Formação acadêmica dos autores	Qualis do periódico (2013-2016)	Quantidade de citação na base (até 2019) ²⁵
LSci-1	Learning to read and its implications on memory and cognition	Linguagem e Linguística	30	2	2016	doutor	A1	3
LSci-2	Knowledge without borders: Policies for Post-modern migrations	Linguística	21	1	2016	doutor	A1	4
LSci-3	Narrative analysis and practices of understanding social life	Linguística Aplicada		2	2015	pós-doutorado/pós-doutorado	A1	2
LSci-4	Language practices in classroom reality: contributions of ethnographic research perspective in Applied Linguistics	Linguística Aplicada	19	1	2015		A1	2
LSci-5	From differences and hierarchies in the TV sketch Adelaide to situated and critical analysis in Applied Linguistics	Linguagem e Linguística	25		2015	mestre	A1	1
LSci-6	Language teaching as a stage of language politics: mobilization of gender repertoires	Linguística Aplicada		1	2017		A1	0
LSci-7	Language as space of power: classroom research from a critical perspective	Linguística Aplicada	29	1	2015		A1	0
LSci-8	Genres in language practices at school: curriculum and teacher education	Linguística Aplicada	39	4		2015	A1	4

²⁵ A plataforma *SciELO CI* não apresenta a contagem de citações que um artigo recebe no âmbito da própria base, como realizado na *Web of Science*. Esses números, portanto, foram buscados de maneira autônoma, artigo por artigo, na base *Google Scholar*, em janeiro 2020.

	Globalization and Multilingualism in Brazil Linguistic Competence and the Science Without Borders Program	Linguística	19	2	doutoranda	A1	-
--	---	-------------	----	---	------------	----	---

Fonte: elaborado pela autora.

O *corpus* da pesquisa é assim constituído de 29 (vinte e nove) artigos científicos. Desses, treze (13) são da subárea de Matemática Aplicada e 16 (dezesesseis), da Linguística. A seguir, descrevemos os procedimentos de organização para posterior análise dos dados.

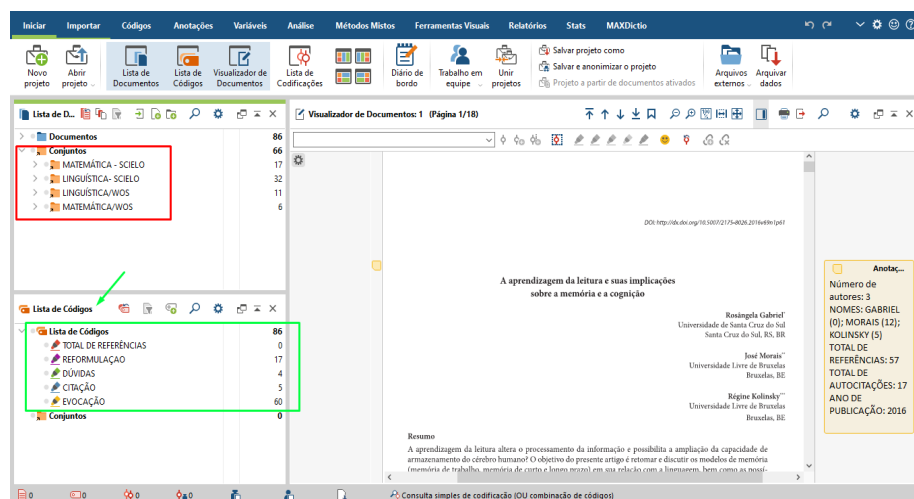
2.4. Procedimentos metodológicos de organização e critérios para análise dos dados

O *software* MAXQDA *Analytics Pro* 2020 (VERBI SOFTWARE, 2019)²⁶ é desenvolvido pela empresa *VERBI Software* na Alemanha e está disponível, mediante pagamento, como aplicação para os dois principais sistemas operacionais no mercado, *Windows* e *IOS*. Segundo informações que constam do *site*, trata-se de uma ferramenta para análise de dados qualitativos e métodos mistos em pesquisas acadêmicas, científicas e comerciais, cujas funcionalidades permitem a análise de documentos, artigos, entrevistas transcritas ou arquivos de texto em geral. Observamos, entretanto, que os chamados *softwares* de apoio à análise de conteúdo, formalmente denominados de programas *Qualitative Data Analysis* (QDA), não realizam as análises de forma automatizada (GARCIA; ESTEVÃO, 2016), sendo necessário o empenho e o trabalho do pesquisador.

Neste trabalho, o *software* foi utilizado no processo de identificação das citações, na contagem das ocorrências, na representação gráfica e visual dos resultados. Seu uso permitiu criar, organizar códigos e subcódigos no processo de análise, com a finalidade de classificar os dados. Na Figura 9 seguinte, observamos a *interface* da área de trabalho do programa, na aba “Iniciar”, e como se dá a disposição dos dados no interior dela:

²⁶ Disponível em: <https://www.maxqda.com/brasil/software-analise-qualitativa>. Acesso em: 09 mar. 2023.

Figura SEQ Figura * ARABIC 9 – Interface da tela inicial do software MAXQDA



Fonte: elaborado pela autora.

Para realizar a primeira etapa de identificação das ocorrências de citações, agrupou-se todo o material em um único projeto no MAXQDA. Cada área do conhecimento e sua respectiva base de dados foi organizada em pastas individuais, formando um conjunto de documentos. Com todo material organizado, o passo seguinte constituiu-se em realizar as buscas das autocitações pelos sobrenomes dos autores de cada artigo, utilizando o recurso “Pesquisa Lexical Avançada”, na aba “Análise”. Foi habilitada a função “ativar”, que consiste em selecionar o conjunto de artigos que se quer analisar. No campo de busca foi preenchido o sobrenome do autor e foram selecionados os campos: “documentos ativados”; “dentro de todo documento”; “dentro da página de PDF” e “dentro de todo o documento”, dando início à fase de codificação das ocorrências de citações no corpo dos artigos.

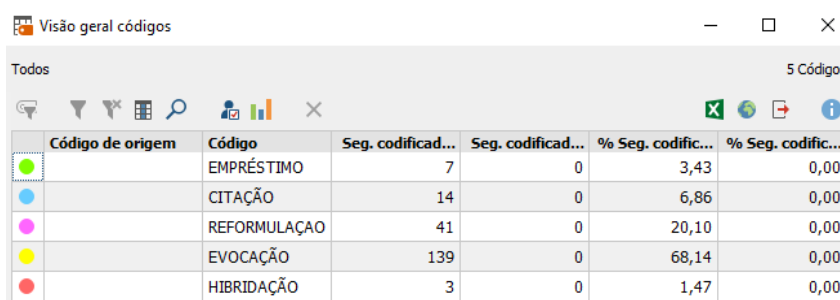
Nessa fase, os desafios se impuseram. Alguns artigos são indexados de forma digitalizada, em formato de imagem (JPG). Assim, o *software* não reconhece a busca lexical como ocorre nos casos dos arquivos em formato PDF. Nesses casos, realizamos as buscas de modo manual, fazendo uma leitura atenta de cada seção dos textos em busca das autocitações.

No procedimento seguinte, o de classificação das categorias de análises, valemo-nos, num primeiro momento, dos critérios assumidos por Boch e Grossmann (2002) na investigação do modo de referência ao discurso de outrem, a saber, evocação, reformulação e citação. Recorremos também à tipologia adaptada por Daunay e Delcambre

(2017 *apud* DAUNAY, 2020), cuja proposta retoma Boch e Grossmann (2002), e consideramos, assim, ainda: empréstimo e hibridação.

Para cada um desses modos de referência, criou-se um código, identificado por uma cor, de acordo com as ferramentas disponíveis no *software MAXQDA*, a fim de proceder à contagem e à disposição visual no texto. Estabeleceu-se: amarelo para evocação, rosa para reformulação, azul para citação direta, verde para empréstimo e vermelho para hibridação, como observado em seguida:

Figura SEQ Figura 1* ARABIC 10 - Classificação dos códigos em cores no software MAXQDA

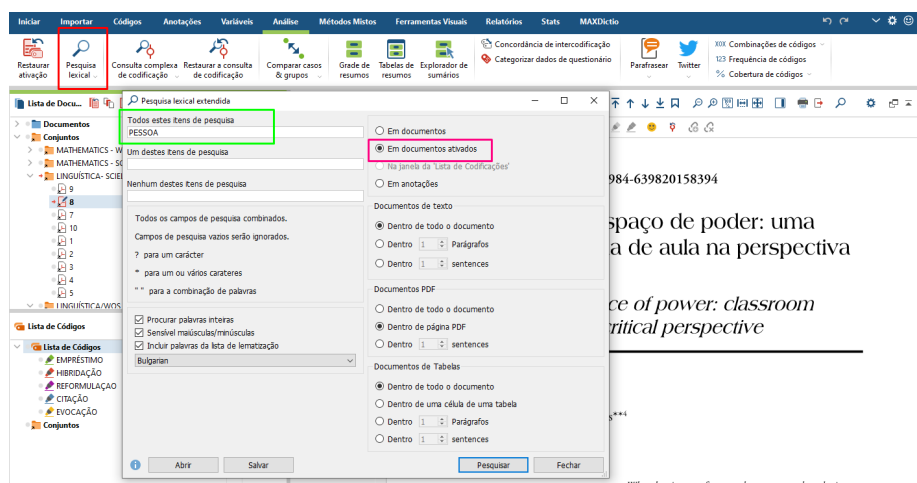


Código de origem	Código	Seg. codificad...	Seg. codificad...	% Seg. codific...	% Seg. codific...
	EMPRÉSTIMO	7	0	3,43	0,00
	CITAÇÃO	14	0	6,86	0,00
	REFORMULAÇÃO	41	0	20,10	0,00
	EVOCAÇÃO	139	0	68,14	0,00
	HIBRIDAÇÃO	3	0	1,47	0,00

Fonte: elaborado pela autora.

Buscamos os sobrenomes dos autores e coautores dos artigos por meio da ferramenta “busca lexical” do *software*. Essa pesquisa nos permitiu identificar as ocorrências de autocitações no corpo do texto, conforme os destaques no *print*, a seguir:

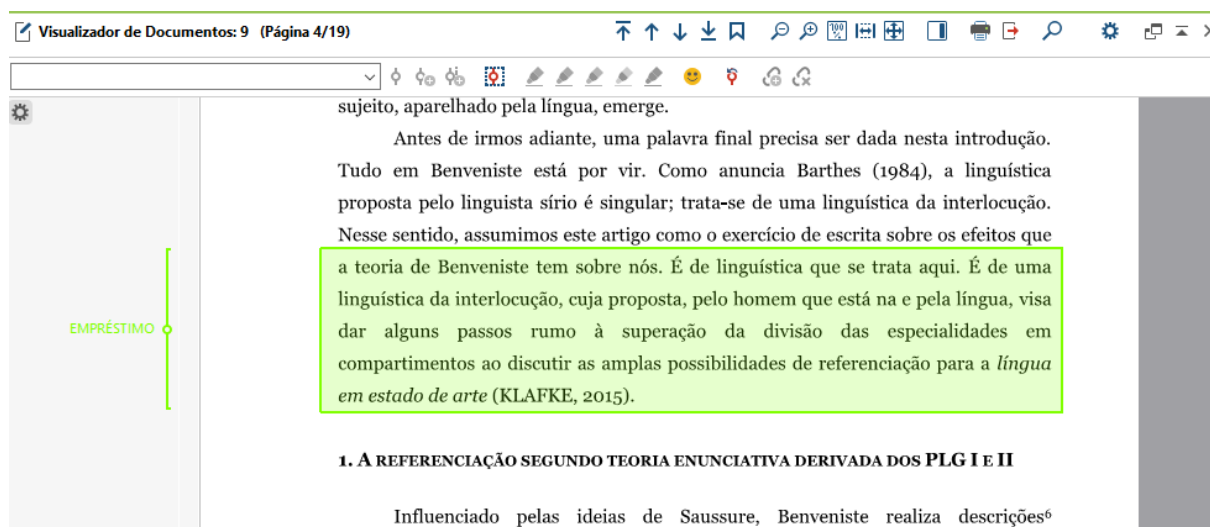
Figura SEQ Figura 1* ARABIC 11 – Print da tela do software MAXQDA, com indicação de “busca lexical”



Fonte: elaborado pela autora.

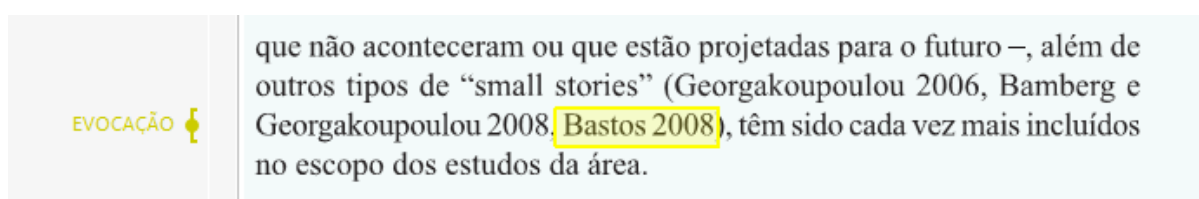
Ao encontrar a citação em que o autor referencia a si próprio, o primeiro passo foi o de identificar em qual classificação se enquadra a referenciação, marcando-a com cor, de acordo com o que foi estabelecido. Os *prints* de telas a seguir indicam casos de *empréstimo*, destacado na cor verde, *evocação* em amarelo, *reformulação* e *citação* em rosa e azul, respectivamente:

Figura 12 – Exemplo de empréstimo, retirado de artigo da subárea de Linguística



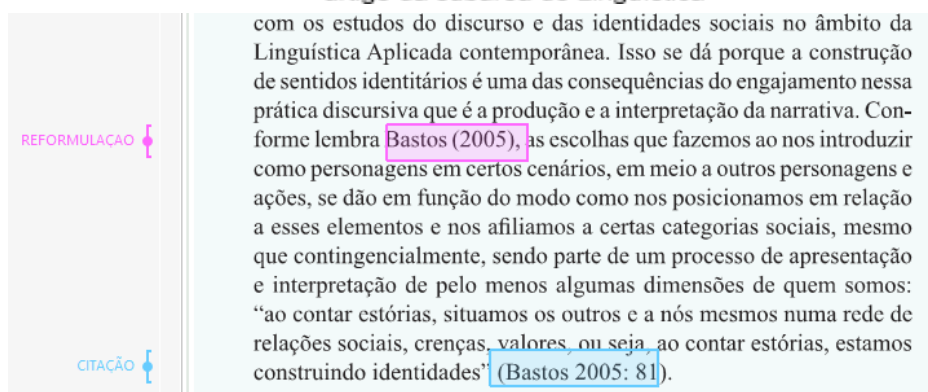
Fonte: elaborado pela autora.

Figura 13 – Exemplo de evocação, retirado de um artigo da subárea de Linguística



Fonte: elaborado pela autora.

Figura SEQ Figura 14 – Exemplos de reformulação e citação direta, retirados de artigo da subárea de Linguística



Fonte: elaborado pela autora.

Para efeito da análise, escolhemos discutir, primeiramente, os modos de referenciação ou de autorreferenciação dos autores de suas produções científicas. Em seguida, buscamos analisar a localização dessas autocitações dentro do corpo do texto, com base no modelo de organização retórica (SWALES, 1990), bem como seu funcionamento, considerando suas atuações na construção da dimensão argumentativa do discurso científico, buscando padrões disciplinares.

Procuramos assim, responder aos objetivos específicos da dissertação, a saber: (i) estudar padrões de autocitação nos artigos científicos dessas duas subáreas de conhecimento, considerando-se o modo como esses artigos se autorreferenciam no processo de autocitação; (ii) estudar as funções da autocitação nos artigos dessas duas subáreas de conhecimento, considerando as suas atuações na construção da dimensão argumentativa do discurso científico. A seguir, no Capítulo 3, mostramos como se deram as análises dos artigos científicos.

Capítulo 3 – Análise dos dados

Neste capítulo, temos por objetivo descrever e analisar o padrão de autocitação em artigos científicos das subáreas de Linguística e Matemática Aplicada. O capítulo divide-se em quatro subseções. Na primeira seção, buscamos apontar para a macroestrutura composicional dos artigos que compõem o *corpus*. Em seguida, discutimos padrões de autocitação, levando em consideração os modos de referência nessas duas subáreas, o modo como a autocitação emerge em cada seção do artigo científico e o modo de referência por seção do artigo científico. Avaliamos em seguida as categorias de autocitação em Linguística e em Matemática Aplicada. Por fim, discutimos a questão da visibilidade nas plataformas das bases de dados investigadas.

3.1. Da macroestrutura do artigo científico nas duas subáreas do conhecimento

Com base no modelo de organização retórica proposto por Swales (1990), que consiste na identificação da estrutura do artigo científico como sendo composta de *Introdução, Métodos, Resultados e Discussão* (IMRD), buscamos avaliar a macroestrutura dos artigos que compõem o *corpus*. Sabe-se que o padrão de organização IMRD tem sido considerado não somente um fator norteador do processo de descoberta científica como também um fator de conhecimento empírico verificável (TESSUTO, 2015). Esse formato, (re)conhecido no meio acadêmico, vem sendo desenvolvido e aprimorado desde o século XIX e tornou-se quase universal nos periódicos científicos (ARAGÃO, 2011, p.159). Jesus (2020) busca sintetizar as principais contribuições dos estudos de Feak e Swales (2009), Motta-Roth e Hendges (2010) e de Marcuzzo (2006), a propósito da macroestrutura da organização retórica, como podemos observar a seguir:

Figura 15– Estrutura IMRD a partir de Feak e Swales (2009)

Seções do Artigo	Propósitos comunicativos
Introdução	Apresentar de forma breve e precisa o conteúdo proposicional de seu texto: objetivos, questões de pesquisa, justificativa; Contextualizar o problema de pesquisa na área de conhecimento pertinente, conforme a literatura prévia (antecedentes); Informar a organização retórica do trabalho.
Metodologia	Descrever os procedimentos de coleta de dados de pesquisa; Descrever os instrumentos e procedimentos de pesquisa adotados Relatar como a pesquisa foi desenvolvida e que dados ela obteve;
Revisão da Literatura	Fazer referência à literatura prévia: utilizar, reconhecer e dar crédito à criação intelectual de outras(os) autoras(es); Demonstrar que a pesquisa se situa numa área de conhecimento, fundamenta-se em e estende pesquisas prévias da área.
Análise e Discussão dos Resultados	Apresentar os resultados encontrados/achados de pesquisa. Comparar os resultados com a literatura prévia Sumarizar

Fontes: Feak; Swales (2009), Motta-Roth; Hendges (2010) e Marcuzzo (2006).

Fonte: (JESUS, 2020)

Embora essa estrutura seja assumida por estudiosos da área como modelo universal para esse tipo de produção textual escrita, pode sofrer variações a depender das normas e convenções que regem diferentes comunidades discursivas, segundo determinados contextos de produção, como argumentado por Aragão (2011, p. 157). Com base nessas considerações, analisamos a estrutura composicional dos artigos em busca das possíveis variações do modelo universal IMRD, voltando-se às especificidades das normas editoriais que regem a publicação de periódico de cada área do conhecimento. Para tanto, organizamos, por área, os quadros a seguir, nos quais observamos comparativamente algumas variações.

Interessa-nos, no Quadro 6 a seguir, discutir o modo como cada autor nomeou as seções dos seus respectivos artigos, levando-se em conta as especificidades disciplinares da área de Linguística e a estrutura composicional de cada um deles.

Quadro 6 – Estrutura composicional dos artigos na subárea de Linguística

ARTIGOS EM LINGÜÍSTICA	INTRODUÇÃO	MÉTODOS	RESULTADOS	DISCUSSÃO	CONCLUSÃO
LSC-1	Introdução	_* ²⁷	-	Discussão	Conclusão
LSC-2	Introdução	Sujeito	Um caso ilustrativo	Discussão e perspectiva	-
LSC-3	Introdução	Narrativa e metodologia da pesquisa qualitativa	Os caminhos analíticos recorrentes	-	Considerações Finais

²⁷ Seção não identificada no corpo do texto.

LSC-4	Introdução	Sobre o distanciamento das pesquisas linguísticas da pedagogia de línguas	-	Etnografia Linguística	Considerações Finais
LSC-5	-	Contextualização da Pesquisa	Análise do quadro "Adelaide"	Avanços	Conclusão: Uma posição compartilhada
LSC-7	Campos em disputa na base nacional comum curricular	-	-	Ensino de inglês para além da colonialidade	Considerações finais
LSC-8	Introdução	O estudo	Significados fluidos	As palavras impõem	Reflexões Finais
LSC-9	Introdução	Aspectos teóricos: gênero e ensino	Linguagem como conteúdo de ensino	Encaminhamentos para o planejamento didático	Considerações finais
LSC-10	Introdução	A globalização	Ensinar e aprender línguas estrangeiras	A pesquisa	Conclusão
LW0S-1	Introdução	Alinhamento Teórico	-	Correlação forma-função na construção distransitiva	Considerações Finais
LW0S-2	Introdução	A memória de uma língua interditada	-	O sujeito surdo e as línguas entrelaçadas na escola- uma análise do relatório	Conclusões
LW0S-4	Introdução	Materiais e Métodos	Descrição do roteiro de entrevista e procedimentos de recolha de dados	A integração do Projeto Nova Iguaçu sob o viés da Sociolinguística ao Curso de Letras da Universidade Federal Rural do Rio de Janeiro – Campus Instituto Multidisciplinar (UFRRJ-IM)	Considerações Finais
LW0S-5	Introdução	-	-	Estrutura do dossiê comemorativo dos 30 anos do do Grupo de Trabalho Teoria da Gramática	-
LW0S-6	Introdução	O possível caminho da construcionalização	-	Comparação entre as construções	Concluindo que não há conclusão
LW0S-7	Aurouz e a rede de retrospecto	-	-	A questão das rupturas: esmaecimento e psicanálise	Conclusão
LW0S-9	Introdução	-	-	A referência segundo teoria enunciativa derivada dos PLG e II	Distribuindo algumas sementes

Fonte: elaborado pela autora.

Nos artigos científicos investigados, preserva-se “Introdução”. Todas as seções correspondentes a “Métodos”, “Resultados” e “Discussão” apresentam alterações terminológicas, isto é, propõem títulos referentes ao tema abordado pelo artigo, como é o caso das seções equivalentes à “Discussão”, que apresentam títulos como “Discussão e perspectiva”, “Etnografia Linguística”, “Avanços”, “Ensino de inglês para além da colonialidade”, “As palavras impõem”, “Encaminhamentos para o planejamento didático”, “A pesquisa”, “Correlação forma-função na construção distransitiva”, “A questão das rupturas: esmaecimento e psicanálise” e “A referenciação segundo teoria enunciativa derivada dos PLG e II”.

Diferentemente da “Introdução”, em que a nomenclatura da seção aparece fixada, invariável e recorrente em todos os artigos, a “Conclusão” aponta para algumas diferenças: “Considerações finais”, “Reflexões Finais”, “Concluindo que não há conclusão” aparecem como títulos dessa seção de encerramento da reflexão, nem sempre sendo mantida de maneira fixa.

O Quadro 7 a seguir apresenta a estrutura composicional dos artigos em Matemática Aplicada.

Quadro 7– Estrutura composicional dos artigos na subárea de Matemática Aplicada

ARTIGOS EM MATEMÁTICA	INTRODUÇÃO	MÉTODOS	RESULTADOS	DISCUSSÃO	CONCLUSÃO
MSCI-1	Introdução	-	Resultados	Conclusão e Análise	-
MSCI- 2	Introdução	Metodologia de investigação	Trabalho em tarefas matemáticas	Discussão e conclusão	Discussão e conclusão
MSCI- 6	Introdução	Bases usadas e preparação dos dados	Resultados e avaliação	-	Conclusões
MSCI- 7	Introdução	-	A avaliação somativa	Pontos fortes e constrangimento	A concluir
MSCI- 8	Introdução	Aspetos metodológicos e pedagógicos	-	Discussão e Resultados	-
MSCI-10	Introdução	-	-	-	Considerações finais
MWoS-3	Introdução	A Metodologia da Pesquisa	Análise e Resultados	-	Considerações Finais

MWoS-4	Introdução	Metodologia	-	-	Considerações Finais
MWoS-5	Introdução	O método de escalonamento simplificado	-	-	Alguns Resultados do estudo de caso
MWoS-6	Introdução	Metodologia	-	-	Conclusões
MWoS-7	Introdução	Metodologia e de Elaboração das Atividades	Aplicação das Atividades	-	Considerações Finais
MWoS-9	Introdução	Metodologia	Programações Computacionais Livres e Gratuitos no Ensino de Matemática	-	Considerações Finais

Fonte: elaborado pela autora.

Com base nesse Quadro 7, é possível observar regularidade na composição do modelo matriz IMDR, em relação aos de Linguística (Quadro 6). Assim como em Linguística, todos os artigos preservam a seção “Introdução”, sem variação. No que se refere à seção de “Métodos”, no entanto, constatamos a existência de: “Metodologia de investigação”, “Bases usadas e preparação dos dados”, “Aspectos metodológicos e pedagógicos”, “Material e Método”, “O método de escalonamento simplificado” e “Metodologia”.

Outro aspecto a ser considerado nos artigos de Matemática Aplicada é que há recorrente fusão das seções de “Resultados e Avaliação”, “Resultados e Discussão” e “Discussão e Conclusão”. Além disso, essas seções apresentam precisões terminológicas, como visto em MSCI-7 e MWoS-3, respectivamente. As seções finais dos artigos são intituladas “Conclusões”, “Discussão e conclusão”, “A concluir” ou ainda “Considerações finais”.

De acordo com os dados dos Quadros 6 e 7, observamos variantes do modelo IMRD no *corpus*: o modelo ICD (Introdução-Desenvolvimento-Conclusão) e IMRD (Introdução-Métodos-Resultados-Discussão). Conforme apontam os estudos de Aragão (2011), há três maneiras possíveis de agrupar e classificar as variações do modelo sociorretórico discursivo, as quais utilizaremos em nossas análises. O primeiro grupo compreende variantes que, embora sejam formadas das mesmas seções que a “matriz” (IMRD), apresentam diferenças terminológicas, como é o caso da maioria dos artigos em Linguística (87,5%) em relação aos da Matemática (62,5%).

O segundo grupo refere-se àquelas variantes que apresentam as quatro seções da matriz e uma ou mais seções diferentes. Nesse segundo grupo, podemos observar que há, nos

artigos em Linguística, uma tendência de as seções se subdividirem ou ainda acrescentarem mais seções do que previsto no modelo IMRD. Um indicativo desse movimento na subárea da Linguística pode ser comprovado pelo maior número de páginas dos artigos analisados dessa subárea, em comparação aos da outra. Os artigos em Linguística apresentam-se mais extensos do que os da Matemática. Essa diferença é uma característica disciplinar que aponta para diferentes modos de argumentação.

No terceiro grupo apresentado por Aragão (2011) estão as variantes, que além de apresentarem diferenças terminológicas, apresentam também uma seção adicional. Os artigos em Linguística compõem, em sua grande maioria, esse último grupo. Nele podem posar observadas variações terminológicas das seções, a exemplo do artigo LSC-8, em que as seções são nomeadas de acordo com o tema de cada tópico (*Introdução – O estudo Significados fluidos – As palavras impõem – Reflexões Finais*). Embora seja possível apontar para variações terminológicas nas seções dos artigos, observamos, como mencionado, que a terminologia das seções iniciais não sofreu variação nas duas subáreas, mantendo a nomenclatura “Introdução”.

Há outro ponto a se destacar com base no estudo de Aragão (2011):

No IMRD, admite-se fusão nas seções de resultados e de discussão, compondo a seção “Resultados e Discussão”. Ainda, nas variantes que incluem uma seção para conclusões, admitem-se três configurações que contêm fusões: 1) “Resultados”, “Discussão” e “Conclusões” em uma única seção; 2) “Resultados” e “Discussão” em uma só seção e “Conclusões” à parte; 3) a seção “Resultados” separada e as seções “Discussão” e “Conclusões” juntas. (ARAGÃO, 2011, p. 157)

Essa mesma tendência de variação, apontada pelo autor, foi observada nos artigos analisados, como é o caso do artigo de Matemática Aplicada (MSCI-2), em que houve uma fusão das seções “Discussão” e “Conclusão” para uma única seção “Discussão e Conclusão”.

Consideramos, a partir do exposto, que ao analisarmos essas variações das seções nessas duas diferentes culturas disciplinares, é possível observar um comportamento e/ou tendência interdisciplinar. Conforme os apontamentos de Jesus (2020), a análise sociorretórica prevê que tanto a organização macro como o microtextual sejam levadas em consideração na produção e análise dos gêneros acadêmicos. E é com base nesse modelo de estrutura de artigo que passamos a analisar as diferenças retóricas das seções como o intuito de buscar padrões interdisciplinares.

3.2. Autocitação nas subáreas de Linguística e Matemática Aplicada

Nesta subseção, avaliamos as autocitações encontradas nos artigos científicos das duas subáreas de conhecimento e nas duas plataformas de bases de dados.

Essa análise se justifica, pois é por meio dela que podemos discutir padrões de autocitação de acordo com a cultura disciplinar. A contagem da porcentagem de autocitação se deu a partir do número absoluto de citações gerais do artigo, isto é, a média foi calculada em relação ao número total de referências por artigo, de acordo com cada área, como vamos explorar nas Tabelas 1 e 2 a seguir.

Tabela 1 – Dados gerais dos artigos na subárea de Linguística

Artigos	Nº de ocorrências de autocitações no corpo do texto	Nº de autocitações nas referências	Nº de referências	% de autocitação em relação às referências
LWoS-1	8	6	56	10,7
LWoS-2	1	1	24	0
LWoS-3	0	0	17	4,1
LWoS-4	1	1	16	6,2
LWoS-6	2	1	30	3,3
LWoS-7	*	1	36	2,8
LWoS-9	7	2	25	8
LSci-1	17	7	33	21,2
LSci-2	0	0	31	0
LSci-3	10	6	47	12,7
LSci-4	0	0	45	0
LSci-5	4	5	46	10,9
LSci-7	15	3	31	9,6
LSci-8	2	2	27	7,4
LSci-9	0	19	44	43,1
TOTAL	67	54	508	10,60%

Fonte: elaborado pela autora.

A autocitação foi amplamente empregada pelos autores dos artigos científicos da subárea de Linguística, considerando-se os artigos encontrados em ambas as plataformas. A Tabela 1 mostra que 6,25% deles apresentaram, pelo menos, um caso de autocitação. Desses, o caso com maior índice de autocitação no corpo do texto apresenta 17 ocorrências (LSci-1). A Tabela 1 também mostra um número expressivo de citações nesta subárea, considerando os

dados da coluna “Nº (número) de referências”, em relação à subárea da Matemática, como explorado a seguir na Tabela 1.

Tabela 2– Dados gerais dos artigos na subárea de Matemática Aplicada

Artigos	Nº de ocorrências de autocitações no corpo do texto	Nº de autocitações nas referências	Nº de referências	% de autocitação em relação às referências
MWoS-2	0	0	17	0
MWoS-4	5	4	11	36,3
MWoS-6	0	0	4	0
MWoS-7	0	0	14	0
MWoS- 8	0	0	18	0
MWoS-9	0	0	11	0
MSci-1	3	1	13	7,69
MSci-2	5	3	9	33,33
MSci-6	8	3	42	7,14
MSci-7	8	3	48	6,25
MSci-8	17	8	29	27,59
MSci-10	7	4	23	17,3
TOTAL	53	26	196	13, 27%

Fonte: elaborado pela autora

Na subárea da Matemática Aplicada, a autocitação foi empregada em cerca de 77,8% dos artigos, ainda que com um baixo índice de ocorrências por artigo. Os dados da Tabela 2 apontam que apenas dois dos artigos do *corpus* não apresentam quaisquer ocorrências de autocitações (MWoS-2; MWoS-6). O artigo MSci-8 apresenta 17 ocorrências de autocitação no corpo do texto, o que representa 27,59% das citações em todo artigo.

Com base nos dados apresentados até o momento, e embora o *corpus* da Matemática Aplicada seja menor em quantidade, se comparado ao de Linguística, é possível constatar uma variação nos números que correspondem às ocorrências de autocitações localizadas no corpo dos artigos de ambas as subáreas. Há um índice maior de autocitação na subárea de Linguística (67) do que em Matemática Aplicada (53). Esse resultado corrobora o que é apontado nos estudos de Hyland e Jiang (2018), para quem há um considerável aumento de autocitações em Linguística, se comparado com outras áreas das engenharias e biológicas, por exemplo. Contrapõem-se, portanto, aos resultados de seu estudo anterior (HYLAND, 2003), cujos dados apontaram as *Hard Sciences* como a que mais investe em autocitação do que as *Soft Sciences*.

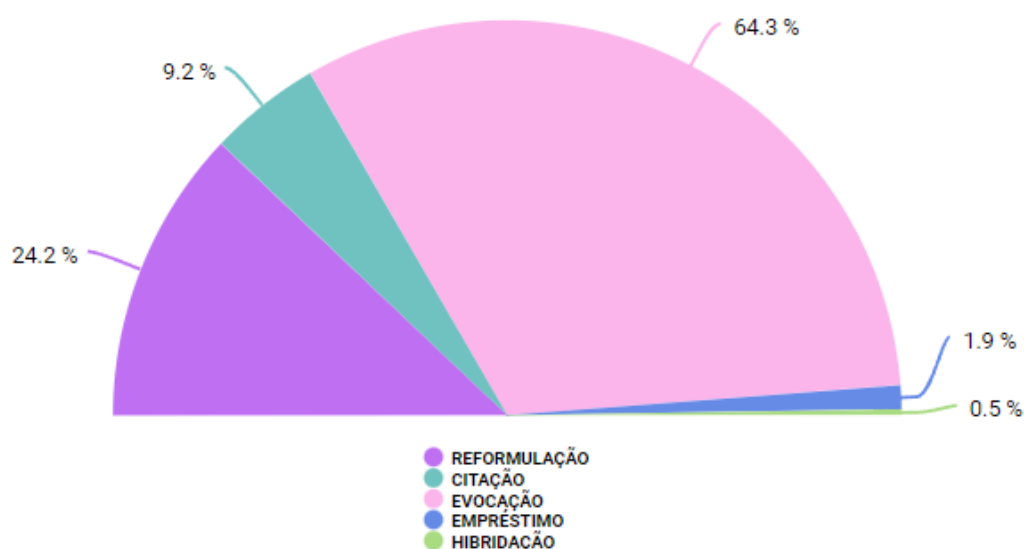
Há também, como mencionado, uma maior tendência de os autores da subárea de Linguística citarem trabalhos de outros, em comparação aos da Matemática Aplicada, por exemplo. Observando os Quadros 2, 3, 4 e 5, é possível afirmar que: (i) o artigo MSci-7, que apresentou o maior número de referência, chegou a 48 referências em 32 páginas; o artigo menos extenso (MWoS-4), com seis páginas, apresenta quatro (04) referências. Em Linguística, o artigo LWoS-1 apresenta o maior número de referências, com indicação de 56 trabalhos citados em 30 páginas. O artigo LWoS-3 é o que apresenta o menor número de referências (17), também em 30 páginas.

Observamos, portanto, que em Linguística os autores tendem a mobilizar mais autores para a produção científica do saber, na argumentação retórica, comparativamente a autores pertencentes à área da Matemática Aplicada.

3.3. Modos de referenciação nas subáreas de Linguística e Matemática Aplicada

Ao analisarmos os modos de referenciação ao discurso do outro, conforme postulados por Boch e Grossmann (2002), Daunay e Delcambre (2017) e Daunay (2020), em ambas as subáreas, concluímos que o maior número de ocorrências de autocitação corresponde à evocação, com 64,3%; seguida de reformulação, com 24,2%; de citação, com 9,2%; de empréstimo, com 1,9%; por fim, de hibridação, com 0,5% das ocorrências, conforme indicado na Figura 16 a seguir.

Figura SEQ Figura 1* ARABIC 16 – Funcionamento geral das autocitações nas subáreas de Linguística e Matemática Aplicada

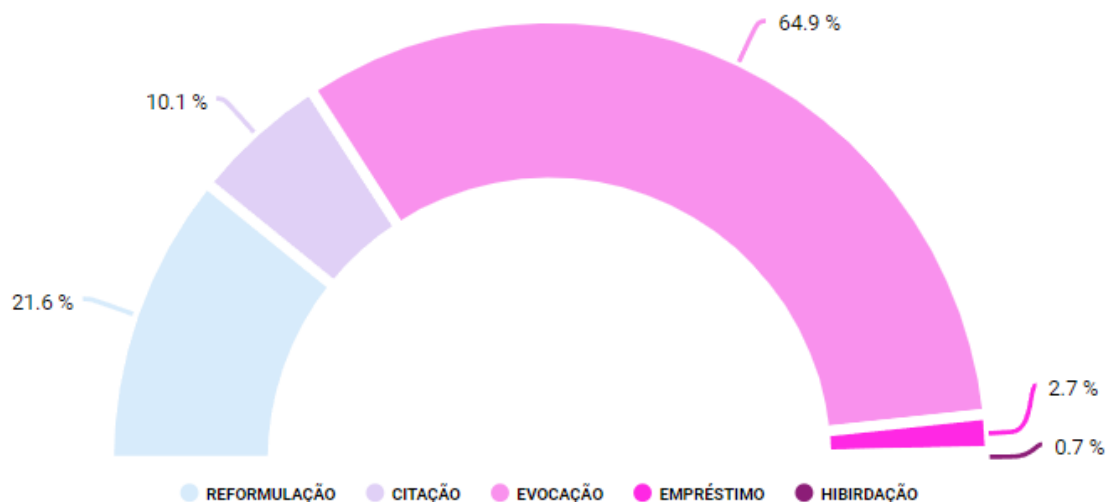


O fato de a evocação despontar como o modo de referenciação privilegiado em ambas as subáreas pode ser interpretado a partir de Boch e Grossmann (2002), cujo estudo mostrou que a evocação é o modo mais recorrente, sobretudo, no que se refere a textos de especialistas (caso dos artigos científicos coletados), indicando que os autores têm uma longa trajetória no campo disciplinar. A autocitação como reformulação, como discurso reportado, aparece, no quadro geral, como o segundo modo de referenciação mais recorrente, com 24,2 % das ocorrências.

A baixa recorrência de autocitação por citação (2,9%), contudo, pode ser justificada com base nas hipóteses de Boch e Grossmann (2002, p. 102), para quem a citação é uma “forma canônica de referência ao discurso do outro”, neste caso, a do próprio autor, e que “fornece modelos de estilo científico e tem, por isso, um papel de aculturação à escrita teórica, permitindo, também, “controlar os conceitos do campo de referência”.

No que se refere às modalidades enunciativas de empréstimo e hibridação, o levantamento mostra baixa ocorrência, de 1,5% no total, para ambas as subáreas. Especificamente, na subárea da Linguística, os dados a seguir mostram que, de modo similar à Matemática Aplicada, a evocação aparece em maior número de vezes nos artigos, com 64,9%, seguida da reformulação com 21,6%, citação com 10,1%, empréstimo com 2,7% e hibridação, com 0,7%. Esses dados permitem observar a preferência dos autores da subárea de Linguística em evocar seus dizeres anteriores no discurso atual, segundo uma integração não evidenciada do dizer anterior em seu próprio discurso, utilizando marcas introdutórias.

Figura SEQ Figura * ARABIC 17 – Funcionamento das autocitações na subárea de Linguística

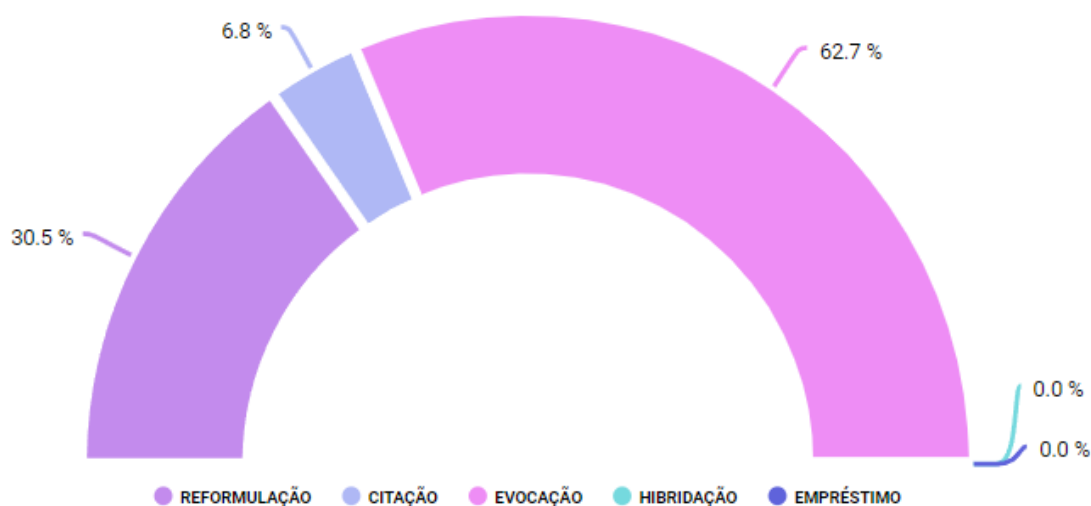


Fonte: elaborado pela autora.

Em Matemática Aplicada, assim como em Linguística, a Figura 18 a seguir demonstra que a taxa de autocitação por evocação aparece como predominante nos artigos, com 75,8%, seguida da reformulação 18,9% e com apenas 5,3% de autocitação por citação direta. A preferência pela autocitação por evocação, nesta área, pode ser justificada também pelo fato de essa ser a forma mais econômica de referenciar-se, por “permite colocar em segundo plano os conhecimentos compartilhados” no fio do texto, sem, no entanto, “parar” para desenvolvê-los, redizê-los ou reformulá-los (BOCH; GROSSMANN, 2002, p.100).

Figura SEQ Figura * ARABIC 18 – Funcionamento das autocitações na subárea de Matemática

MODOS DE REFERENCIAÇÃO DE MATEMÁTICA



Fonte: elaborado pela autora.

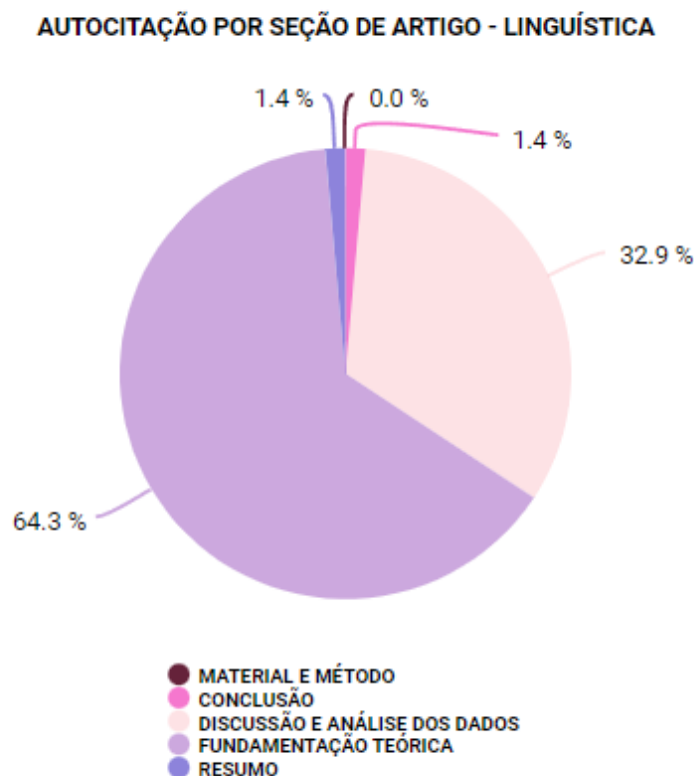
A ausência das modalidades enunciativas de hibridação e empréstimo, na subárea de Matemática Aplicada, é algo a ser destacado, pois pode apontar para um comportamento disciplinar, uma vez que, como mencionado, as modalidades enunciativas são formas retóricas da subjetividade do autor, podendo revelar de modo mais ou menos visível seu posicionamento em relação ao discurso do outro, nesse caso, ao próprio (tomado como alheio). Essa conjuntura, em que não há registros dessas modalidades, pode significar, segundo Daunay (2020, p. 366), que se trata de discursos do tipo pesquisa, nos quais os autores mantêm uma relação mais impessoal com o texto e que não evidenciam, portanto, sua “própria voz” inscrita em trabalhos anteriores.

3.2.1. Autocitação por seção do artigo científico nas duas subáreas do conhecimento

Num segundo momento, analisamos em quais seções dos artigos científicos concentram-se o maior número de autocitações em ambas as subáreas, pois, conforme indica Hyland (2011 *apud* FISCHER; FERREIRA; SILVA, 2020), a alta taxa de autocitações em determinadas seções pode relacionar-se às exigências do campo disciplinar. Em Linguística, observou-se que há uma tendência de os autores se autocitarem na “Fundamentação Teórica”

e “Discussão e Análise de dados”, principalmente; em seguida, em “Introdução”, mas não em “Material e método”. Evidentemente, o conjunto do material é restrito e precisaria ser ampliado para uma discussão mais aprofundada. Vejamos a Figura 19 a seguir:

Figura SEQ Figura * ARABIC 19 – Autocitação por seção em artigos da subárea de Linguística, em ambas as bases



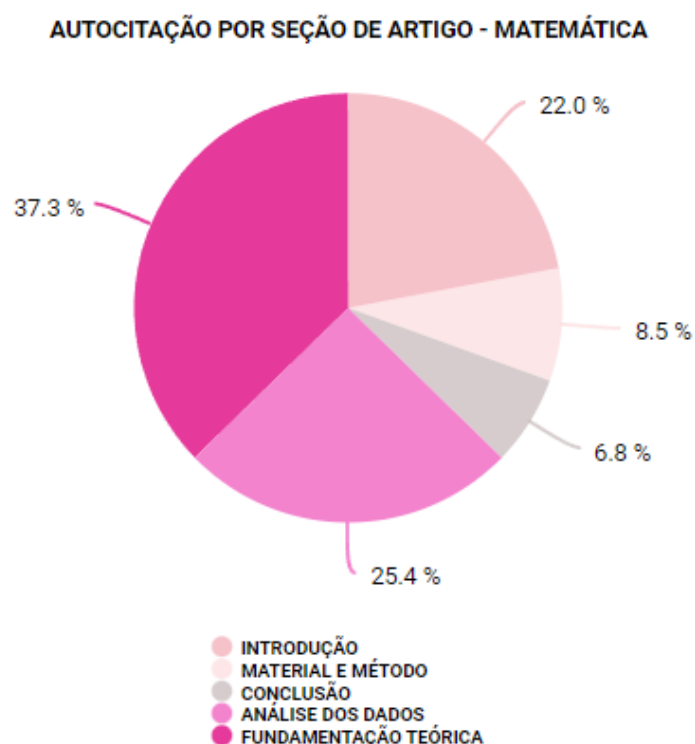
Fonte: Elaborada pela autora

Holmes (1997 *apud* Hyland, 2006, p. 332) oportunamente observa que “as ciências sociais, ao contrário das ciências naturais, tendem à complexidade e à elaboração no início do artigo e não no final”. Holmes, na citação de Hyland (2006), atribui essa variação à tendência de a pesquisa nas ciências naturais refletir tendências de maior gasto e burocratização medida por dados quantitativos, por autoria colaborativa, suporte financeiro externo e padrões de discurso uniformes.

Na subárea da Matemática Aplicada, por sua vez, observamos que as autocitações aparecem predominantemente em “Fundamentação Teórica”, seguida da seção de “Análise dos dados”. Áreas como essas, segundo Hyland (2003), tendem à autocitação, pois essa dinâmica contribui para o sentido de progressão linear, muitas vezes tido como característica

das ciências exatas. Esse raciocínio pode, portanto, justificar a predominância de autocitações na seção de “Fundamentação Teórica”, por exemplo. Além disso, observamos um resultado diverso do da Linguística, no que diz respeito às ocorrências, embora poucas, de autocitação na seção de “Conclusão”, com 6,8%, conforme indica a Figura 20 a seguir:

Figura SEQ Figura 1* ARABIC 20 – Autocitação por seção em artigos da subárea de Matemática



Fonte: elaborado pela autora.

3.2.2. Modos de referenciação por seção de artigo nas duas subáreas do conhecimento

Também analisamos os modos de referenciação por seção de artigo, buscando observar pontos de aproximação ou de distanciamento entre as culturas disciplinares. Em Linguística, conforme indica a Tabela 3 a seguir, nota-se que há uma tendência de os autores se autocitarem na “Introdução” e “Fundamentação Teórica”, principalmente, por meio da evocação e reformulação, especificamente. A evocação também aparece predominantemente nas seções “Material e método”, “Resultados”, “Análise de dados”. A autocitação por reformulação, por seu turno, aparece em segundo lugar na “Introdução”, “Fundamentação Teórica”, “Análise de dados” e predominantemente em “Resumo”. A autocitação por citação

aparece em terceiro lugar nas seções “Resultados” e “Discussão” e “Considerações Finais”. As modalidades enunciativas (hibridação e empréstimo) aparecem somente em “Fundamentação Teórica” nessa subárea.

Tabela 3 – Modos de referência por seção de artigos na subárea de Linguística

SEÇÃO DO ARTIGO	REFORMULAÇÃO	CITAÇÃO	EVOCÇÃO	EMPRÉSTIMO	HIBRIDAÇÃO
RESUMO	0	0	1	0	0
INTRODUÇÃO	2	1	23	0	0
FUNDAMENTAÇÃO TEÓRICA	14	3	45	3	1
MATERIAL E MÉTODO	0	1	2	0	0
RESULTADOS E DISCUSSÃO	8	4	12	1	0
ANÁLISE DE DADOS	5	1	12	0	0
CONSIDERAÇÕES FINAIS	3	5	1	0	0
REFERÊNCIAS	0	0	0	0	0
TOTAL	32	15	96	4	1

Fonte: elaborado pela autora.

Conforme apontam os dados da Tabela 4 a seguir, os autores da subárea da Matemática Aplicada tendem a citar, por meio da evocação e reformulação, seus trabalhos anteriores principalmente nas seções iniciais, “Introdução” e “Fundamentação Teórica”.

Tabela 4 - Modos de referência por seção de artigos na subárea de Matemática Aplicada

SEÇÃO DO ARTIGO	REFORMULAÇÃO	CITAÇÃO	EVOCÇÃO	EMPRÉSTIMO	HIBRIDAÇÃO
RESUMO	0	0	0	0	0
INTRODUÇÃO	2	0	12	0	0
FUNDAMENTAÇÃO TEÓRICA	11	1	14	0	0
MATERIAL E MÉTODO	0	1	2	0	0
RESULTADOS E DISCUSSÃO	0	0	3	0	0
ANÁLISE DE DADOS	4	1	6	0	0
CONSIDERAÇÕES FINAIS	0	1	0	0	0
REFERÊNCIAS	1	0	0	0	0
TOTAL					

Fonte: elaborado pela autora.

McGrath e Kuteeva (2012), em seu estudo sobre a cultura disciplinar da Matemática Pura, apontam que embora a divisão estrutural dos artigos de Matemática seja retoricamente simples, composta de seções como “Métodos”, “Resultados” e “Introduções e Discussões”, a escrita pode revelar os esforços do autor de persuadir o público com base em suas afirmações. Esse estudo aponta ainda que, de modo similar aos artigos do *corpus* de Matemática Aplicada, os delas também não aderem necessariamente à estrutura prototípica do IMRD, pois constroem-se da seguinte forma: Resumo, Introdução, Corpo Principal e Discurso de encerramento.

Embora os artigos analisados pertençam ao domínio da Matemática Aplicada e não ao da Pura, observamos um padrão mais próximo ao da organização retórica da área pura. A justificativa pode estar relacionada ao fato de a inscrição do autor do artigo, no campo aplicado, ainda estar subordinada à área principal (no caso, a Matemática), atendendo a demandas institucionais de sua *comunidade discursiva* de origem, isto é, da área das exatas, da Matemática Pura.

3.3. Categorias de autocitação em Linguística e Matemática Aplicada

Ainda sobre a análise dos modos de referência, classificamos as autocitações de acordo com as mencionadas nomenclaturas: conceitual, evolutiva, confirmativa/negativa e justaposta (MORAVICSIK sem data *apud* AMANTE, 2011). Para essa análise, fizemos um levantamento qualitativo, baseado na identificação de cada autocitação e classificação de acordo com as características apresentadas na subseção 1.3, “O fenômeno da autocitação em culturas disciplinares distintas”. Na investigação dessas duas subáreas, observamos que algumas autocitações que não se encaixam em nenhuma das categorias ou, ainda, se adequam a mais de uma (por exemplo: conceitual-confirmativa ou conceitual-negativa), como no caso a seguir:

Cunha Vieira et. al. (2016) propõem “que as construções clivadas sirvam para focalizar elementos frasais e reforçá-los na memória de trabalho (memória focal) do ouvinte/leitor”, o que poderia levar um leitor desavisado a concluir que o Princípio da Não Sinonímia estaria sendo violado. (LWoS - 6, P. 16)

No Quadro 8, veremos, a seguir, os três padrões de autocitação mais recorrentes, por meio dos excertos retirados dos artigos:

Quadro SEQ Quadro 1* ARABIC 8 – Exemplos dos tipos de citação nos artigos de Linguística e Matemática

	CONCEITUAL	CONFIRMATIVA	EVOLUTIVA
LINGUÍSTICA	A análise da construção ditransitiva segue os pressupostos teórico metodológicos da Linguística Funcional Centrada no Uso (FURTADO DA CUNHA et al., 2013)[...]. - (LWOS -1, P. 3)	[...] como demonstrado por Morais <i>et al.</i> (1979) por outro lado, isso não significa que não sejam sensíveis à codificação fonológica, como por exemplo à rima [...] (LSC -1, P. 14)	Por fim, apresentaremos um breve recorte desse nosso trabalho, retirado de uma pesquisa de mestrado, cujo contexto foi uma sala de aula de língua inglesa de uma escola pública de Goiânia (HOELZLE, 2016).” - (LSC-7, P.1)
MATEMÁTICA	Um conceito central desse construto é a noção de moldagem recíproca (Borba, 1993, 1999), segundo a qual, os feedbacks dados por uma determinada mídia influenciam no raciocínio de quem interage com elas, em outras palavras, a mídia molda o ser humano [...] - (M8 -SCI, P. 7)	Conforme apresentado em detalhes em Fonseca (2014), a tabela TS_QUEST_PROFESSOR teve seus dados transformados, de modo que fosse possível identificar as classes às quais os professores pertenciam, possibilitando que, na mineração de dados, tarefas apresentadas na próxima seção, fizesse-se uso desses dados para identificação de padrões relevantes. (M4-SCI, P. 12)	Mais detalhes sobre o processo de seleção de dados podem ser encontrados em Fonseca (2014) - (M4- SCI, P. 10)

Fonte: elaborado pela autora.

As autocitações classificadas como conceituais foram majoritariamente encontradas na seção “Fundamentação Teórica”. No excerto a seguir, o autor do artigo na subárea de Matemática Aplicada qualifica o trabalho autocitado como “central”:

Excerto 01

Um conceito central desse construto é a noção de moldagem recíproca (Borba, 1993, 1999), segundo a qual, os feedbacks dados por uma determinada mídia influenciam no raciocínio de quem interage com elas, em outras palavras, a mídia molda o ser humano [...] - (M8 -SCI, P. 7)

Em Linguística, a autocitação conceitual indica ao leitor que os pressupostos teóricos que fundamentam a pesquisa baseiam-se em conceitos mobilizados em trabalhos anteriores. Em geral, as autocitações confirmativas, aquelas em que os autores apoiam ou atestam que a obra citada está correta, aparecem predominantemente compostas pelos tipos conceitual-afirmativo, ou seja, os que indicam defesa de um conceito autocitado e são comumente estruturadas por meio de reformulações (quase nunca por meio de evocações). O excerto seguinte é representativo desse padrão:

Excerto 02

[...] como demonstrado por Morais *et al.* (1979) por outro lado, isso não significa que não sejam sensíveis à codificação fonológica, como por exemplo à rima [...] (LSC -1, P. 14)

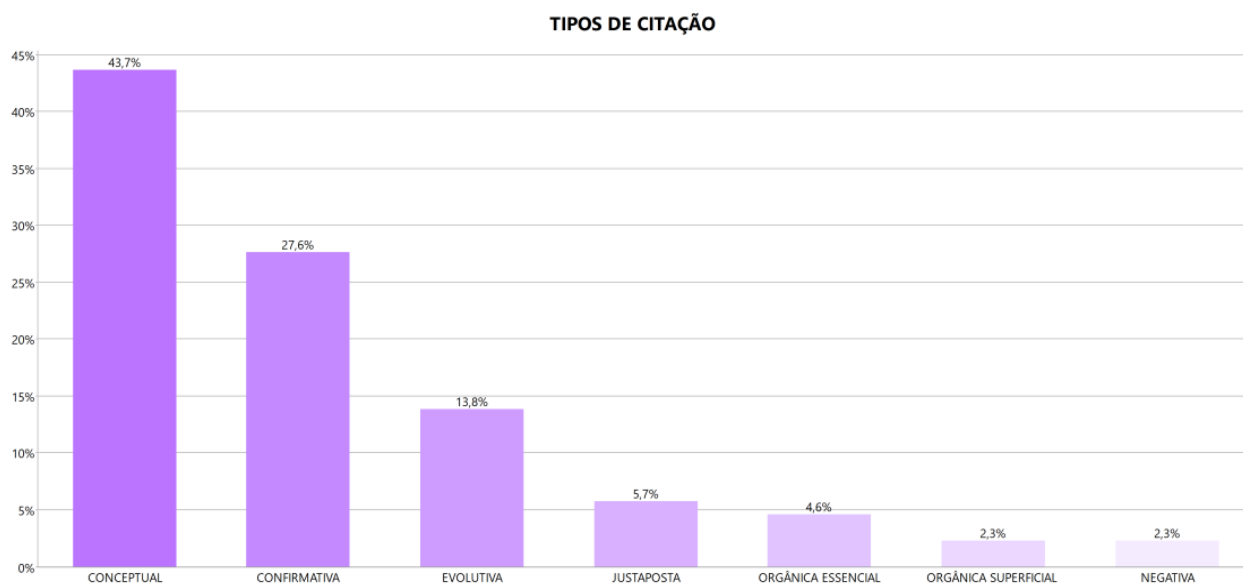
Já as autocitações evolutivas, que se baseiam na menção a trabalhos análogos, são, em grande parte, referenciadas por meio de evocações, conforme o excerto seguindo extraído de artigo da subárea de Linguística:

Excerto 03

Por fim, apresentaremos um breve recorte desse nosso trabalho, retirado de uma pesquisa de mestrado, cujo contexto foi uma sala de aula de língua inglesa de uma escola pública de Goiânia (HOELZLE, 2016)." - (LSC-7, P.1)

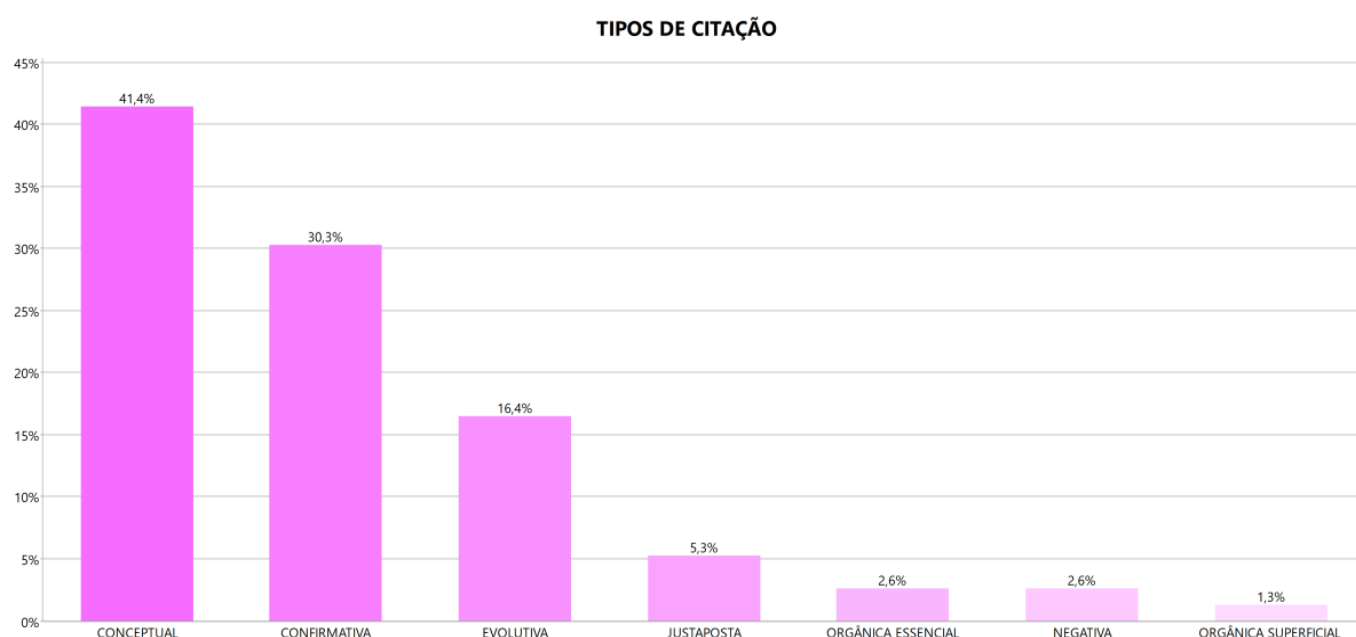
No que diz respeito aos tipos de autocitação mais recorrentes nos artigos das subáreas de Linguística e de Matemática Aplicada, vejamos as Figuras 21 e 22:

Figura SEQ Figura * ARABIC 21 – Tipos de citação dos artigos na subárea de Linguística



Fonte: elaborado pela autora.

Figura SEQ Figura * ARABIC 22 – Tipos de citação dos artigos na subárea de Matemática Aplicada



Fonte: elaborado pela autora.

De acordo com as Figuras 21 e 22, nas subáreas de Linguística e de Matemática Aplicada, 43,7% e 41,4% das autocitações, respectivamente, podem ser tomadas como conceituais. Isto é, quase metade das autocitações em ambas as subáreas tratam de um conceito ou uma teoria empregada de forma direta ou indireta (reformulada ou evocada).

Quando entrecruzamos esse dado de que quase metade das autocitações encontradas nessas subáreas tem como objetivo mobilizar ou expor algum conceito já tratado em trabalho anterior, ao dado de que a maioria das autocitações aparece nas primeiras seções dos artigos (“Introdução” e “Fundamentação Teórica”), podemos dizer que os autores, de ambos os campos disciplinares, buscam citar seus trabalhos anteriores a fim de apoiar-se em conceitos já difundidos ou trabalhados por eles em outro momento, com o claro objetivo de dar suporte às suas afirmações atuais.

Swales (2004 *apud* MACEDO; PAGANO, 2011, p. 276) aponta que nas seções iniciais (Introdução, Fundamentação Teórica) as obras de outros têm importante papel, enquanto a pesquisa que as cita assume papel secundário. No caso das autocitações, em que o autor evoca conceitos difundidos por ele próprio em trabalhos anteriores, o foco principal passa a ser a voz e o trabalho próprios, tomados como um terceiro no funcionamento do discurso científico. Esse movimento retórico pode evidenciar também o processo de produção

do conhecimento científico contínuo, produzido dentro das respectivas comunidades disciplinares.

A segunda maior ocorrência de citação identificada em ambas as subáreas é a confirmativa. Esse padrão de autocitação opõe-se à negativa, à medida que aquela tem como característica apoiar um argumento quando o autor cita outro autor, com a finalidade de defendê-lo. Na autocitação, a confirmativa ocorre quando o autor referencia algum outro trabalho anterior com o intuito de confirmá-lo no trabalho atual.

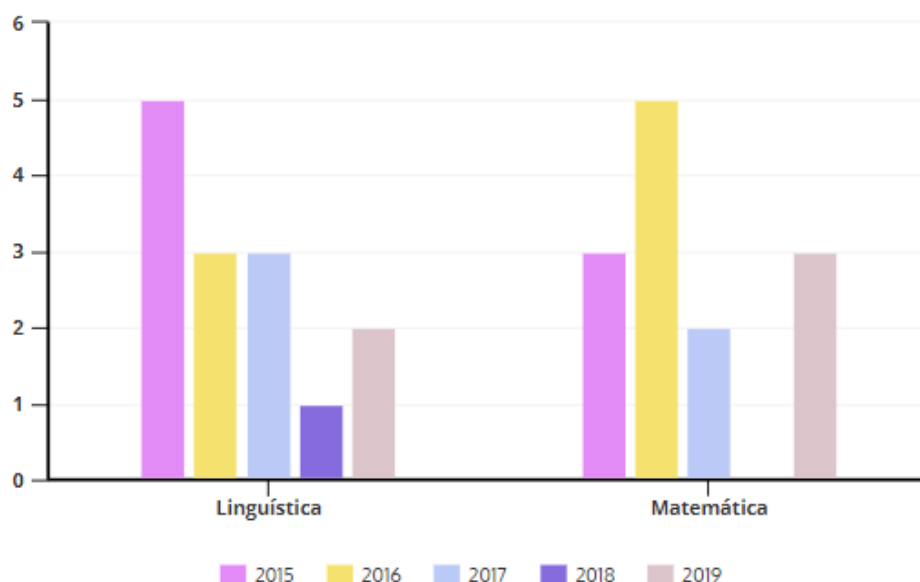
Nas Figuras 21 e 22, é possível visualizar que a ocorrência de autocitação do tipo negativa, ou seja, aquela em que o autor cita seus próprios trabalhos com o intuito de opor-se às suas ideias ou refutá-las, não chega a 3% em ambas as subáreas. O baixo índice de citações do tipo negativa ou justaposta, segundo Moravcsik e Murugesan (1975 *apud* MACEDO; PAGANO, 2011, p. 284), pode indicar que a “área de estudo em que esses artigos estão inseridos pode ser caracterizada em relação à atividade científica, mas talvez não quanto ao progresso científico, uma vez que é a citação justaposta que traz o conhecimento novo para a área de estudo em questão”. Além do mais, os autores observam que a ocorrência ou não desses tipos de citação pode estar relacionada a questões “culturais e à maneira como a comunidade discursiva que produz esses artigos se relaciona com a autoridade/obra citada e à maneira como organiza a produção e veiculação de conhecimento.” (MACEDO; PAGANO, 2011, p. 284).

A autocitação evolutiva aparece, no quadro geral, como sendo a terceira mais empregada pelos autores das duas subáreas. Esse tipo de citação tem como objetivo qualificar aquelas autocitações que contribuem para o desenvolvimento lógico do tema ou da ideia do autor. Tanto em Linguística, como em Matemática Aplicada, o número de vezes em que a autocitação evolutiva ocorre é similar, 13,8% e 16,4%, respectivamente.

3.4. A visibilidade nas plataformas de bases de dados

A Figura 23 a seguir mostra o quantitativo de citações que os artigos de cada subárea recebeu ao longo do período investigado (2015-2019).

Figura SEQ Figura * ARABIC 23 - Data de publicação dos artigos por ano



Fonte: elaborado pela autora.

Os dados bibliométricos dos artigos analisados indicam que em Linguística o maior número de autocitação de um artigo é de 17 (dezessete). Esse mesmo artigo foi o mais citado, dentre os artigos do *corpus*, na base *SciELO*. Em Matemática Aplicada, o maior número de autocitação é de sete (07). O artigo mais citado nessa mesma plataforma teve apenas uma (01) autocitação.

A seguir, ilustramos como a autocitação se comportou em cada plataforma de base de dados, de acordo com a subárea em análise.

Figura SEQ Figura * ARABIC 24 - Demonstração de autocitação por base de dados



Fonte: elaborado pela autora.

O baixo índice de citação em ambas as plataformas de base de dados pode estar relacionado à publicação em língua portuguesa. Costa e Gazarian (2021)²⁸ mostram índices distintos para artigos de Linguística e de Imunologia, publicados nas mesmas bases, no mesmo período, mas em língua inglesa. Esse dado indica que a internacionalização do produto do conhecimento científico é padrão em lugares onde o inglês não é o idioma oficial, mantido por práticas em nível nacional e transnacional. As produções científicas que fogem a esse padrão não apresentam o impacto ou a visibilidade que outras publicadas em inglês, pois, como argumentam Lillis e Curry (2011), o estatuto global da língua inglesa, aliado ao aumento de publicações, implica uma pressão cada vez mais acentuada por publicação nessa língua. Ainda sob a perspectiva dos estudos de letramentos, a internacionalização da pesquisa científica reforça a ideia de que os letramentos são instituídos de acordo com as relações de poder de uma dada comunidade, no caso, a mundial. Essa ideia generalizada de mundialização “esconde” que variáveis distintas como idioma, mas também: periódico, financiadores, assunto do trabalho, reconhecimento dos autores no campo científico também atuam nessa promoção, não sendo suficiente a atividade de autocitação.

²⁸ Comunicação oral intitulada “Letramentos acadêmicos em artigos científicos em Ciências Biológicas e em Linguística, Letras e Artes: a questão da autocitação”, apresentada pelas mestrandas Juliana Renata Pereira da Costa e Ana Paula Garcia Gazarian, do Programa de Pós-graduação em Estudos Linguísticos da Unesp, no 68º Seminário do Grupo de Estudos Linguísticos do Estado de São Paulo (GEL), em São Carlos, São Paulo, no período de 05 a 09 de julho de 2021.

Considerações Finais

Nesta Dissertação de Mestrado, buscamos estudar padrões de autocitação em artigos científicos de duas grandes áreas de conhecimento, a saber, Ciências Humanas (subárea de Linguística) e Ciências Exatas (subárea de Matemática Aplicada). Levamos em consideração, dentre outros aspectos, a distinção entre culturas disciplinares e o “atravessamento” de demandas do contexto sócio-histórico da produção do conhecimento, em nível global, mediante a utilização de plataformas de bases de dados (no caso deste trabalho, *Web of Science* e *SciELO*) que promovem a circulação das pesquisas (dos pesquisadores, dos grupos de pesquisa, das instituições). Partiu-se da hipótese de que, por um lado, há prevalência das demais áreas de conhecimento em relação às Ciências Humanas, segundo uma dinâmica de relações de poder e autoridade em práticas letradas acadêmico-científicas; por outro, o processo de autocitação corrobora a contagem de citações em indicadores bibliométricos, com promoção de visibilidade em comunidades discursivas. Para o cumprimento desse objetivo geral, traçamos o cumprimento de dois objetivos específicos, levando-se em conta: (i) o estudo de padrões de autocitação em artigos científicos dessas duas subáreas, considerando o modo como esses artigos se autorreferenciam no processo de autocitação; (ii) o estudo da função da autocitação nos artigos dessas duas diferentes grandes áreas de conhecimento, dadas as suas atuações na construção da dimensão argumentativa do discurso científico.

O conjunto do material foi formado de 40 (quarenta) artigos científicos publicados em língua portuguesa, 20 (vinte) de cada subárea (Linguística e Matemática), 10 (dez) de cada uma das plataformas de bases de dados escolhidas (*Web of Science* e *SciELO*), com alto índice de citação segundo indicadores bibliométricos, no período de 2015 a 2019. Desse conjunto mais amplo de material, constitui-se um *corpus* formado de 27 (vinte e sete) artigos científicos das duas subáreas (16 de Linguística e 09 de Matemática Aplicada).

Para além de uma crítica do senso comum, segundo a qual a autocitação seria um modo artificial de se alcançar índices bibliométricos em bases de dados, nossa investigação relevou uma dinâmica complexa, constitutiva do funcionamento científico de diferentes subáreas do conhecimento. No quadro teórico constituído nesta Dissertação, buscamos discutir que há uma conjuntura que impulsiona a emergência desse funcionamento: demanda de produtividade contínua, avaliada por pares (mas não somente), sob pena de não haver

concessão de financiamento para pesquisa; exigência de originalidade da produção científica; criação de dispositivos de averiguação dessa originalidade, por meio de ferramentas digitais de detecção de plágio (ao mesmo tempo em que há franco desenvolvimento de Inteligência Artificial para dar suporte a esse produtivismo). O fenômeno da autocitação assim “responde” a essas condições de produção e distribuição do conhecimento científico.

No Capítulo 1 (“Fundamentação teórica”), assumimos o fenômeno da autocitação como consequência retórica do crescimento do produtivismo e da competitividade acadêmicas. À luz de pressupostos da perspectiva sociocultural dos letramentos, de maneira particularizada, dos letramentos acadêmicos, buscamos discutir formas pelas quais os autores se referem a si próprios e a trabalhos já realizados, considerando-se modos de referenciação no/do discurso científico. Não se ignorou que esse funcionamento é também associado a uma questão quantitativa de contagem de referências nas plataformas de bases de dados. Buscou-se, pois, explorar os modos de produção, organização e distribuição do conhecimento, num cenário de mundialização da pesquisa, em que a língua inglesa aparece como referência para a comunicação científica.

No Capítulo 2 (“Material e Métodos”), foram descritas e apresentadas justificativas para as escolhas: (i) das subáreas do conhecimento (Linguística e Matemática Aplicada), (ii) das plataformas de bases de dados (*Web of Science* e *SciELO*), (iii) do conjunto do material que permitiu a constituição do *corpus* e (iv) dos procedimentos metodológicos de organização e critérios para análise dos dados. Sobre a subáreas do conhecimento, observou-se que essa escolha se deu no âmbito da execução de projetos de pesquisa mais amplos, dos quais participamos ao longo do Curso de Mestrado. Outros trabalhos foram ou estão sendo realizados, com base em conjunto de material advindo de Ciências Biológicas, Ciências Contábeis, Direito, Educação, Física, Química, Saúde. A língua de publicação dos artigos científicos (língua portuguesa) também foi escolhida em meio a outras que estão sendo investigadas nesses projetos mais amplos (línguas espanhola, francesa e inglesa).

Ainda que o conjunto do material e o *corpus* possam ser considerados restritos para esse tipo de investigação, observamos que os critérios para sua constituição não foram nada evidentes, tendo sido construídos coletivamente ao longo dos anos 2020 e 2021, por meio do trabalho entre grupos de pesquisa sediados na Unesp (grupo de pesquisa “Práticas de leitura e escrita em contexto digital” – CNPq/Unesp), com participação da FURB, e na PUC Minas (grupo de pesquisa “Núcleo de Estudos em Linguagens, Letramentos e Formação” – NELLF/CNPq/PUC Minas). Mediante esse trabalho coletivo e com financiamento do Programa de Pós-graduação em Estudos Linguísticos da Unesp, tomamos conhecimento do

software MAXQDA. Foi feito um treinamento com profissional para sua utilização no tratamento dos dados. Os critérios de análise foram concebidos no âmbito dos projetos de pesquisa.

No Capítulo 3 (“Análise dos dados”), apresentamos inicialmente uma apreciação geral sobre a macroestrutura do artigo científico nas duas diferentes subáreas do conhecimento, observando, com base na composição do modelo matriz IMDR, que a cultura disciplinar na subárea de Linguística mantém a nomenclatura “Introdução”, sem variação, a exemplo do que ocorre com a Matemática Aplicada. Na subárea de Linguística, as demais seções são mantidas com mudança de título; na subárea de Matemática Aplicada, há fusão de seções, em que “resultados” e “discussão”/ “avaliação” aparecem juntos, o que não costuma ocorrer com a outra subárea avaliada. Essa organização era de nosso interesse na investigação dos padrões de autocitação nos artigos científicos das duas subáreas, dadas as suas atuações na construção da dimensão argumentativa do discurso científico.

Na análise das referências, classificamos as autocitações de acordo com os conceitos apresentados por Amante (2011), com base em Moravcsik (sem data): conceitual, evolutiva, confirmativa/negativa e justaposta (MORAVISCKI sem data *apud* AMANTE, 2011). Tanto na subárea de Linguística quanto na de Matemática Aplicada, as autocitações mais recorrentes tratam de um conceito ou de uma teoria empregada de forma indireta – evocada –, a partir dos trabalhos precedentes. As autocitações encontradas nessas subáreas têm como objetivo mobilizar ou expor algum conceito e são empregadas nas primeiras seções dos artigos, “Introdução” e “Fundamentação Teórica”. Inferimos que os autores, de ambas as subáreas, buscam citar seus trabalhos anteriores a fim de apoiar-se em conceitos já difundidos ou trabalhados por eles em outro momento, com o claro objetivo de dar suporte às afirmações atuais. A investigação mostrou que há uma maior tendência de os autores em Linguística citarem trabalhos de outros, em comparação aos da Matemática Aplicada.

Ao concluirmos o processo de análise, observamos, à luz do funcionamento discursivo da autocitação, em que o autor evoca conceitos difundidos por ele próprio em trabalhos anteriores, que o foco principal passa a ser a voz e o trabalho de quem autocita e não o de um terceiro. Esse movimento retórico pode evidenciar o processo de produção do conhecimento científico contínuo, produzido no âmbito das respectivas comunidades disciplinares. Esta Dissertação buscou, assim, mostrar no âmbito do funcionamento de duas diferentes subáreas, num contexto institucional e mundializado, o modo como se articulam as relações de poder engendradas na sociedade. Do ponto de vista do sujeito pesquisador, com base em trabalho já realizado, a autocitação sinaliza o histórico de pertencimento a correntes

teóricas; permite assim sustentar uma definição (BOCH; GROSSMANN, 2015) e aponta para a visibilidade na comunidade discursiva, em meio à competitividade (HYLAND, 2003).

O processo de investigação de artigos científicos em plataformas de bases de dados, em duas subáreas de conhecimento, apresentou, entretanto, um alcance: o do trabalho do pesquisador que terá de “ler” para além dos índices que podem ser localizados por sistema de busca automático, na interpretação dos modos de referência nos artigos. Observa-se que a continuidade dessa investigação demandaria *expertise* no domínio disciplinar (para além da superfície do texto), como mostra esta pesquisa, e terá muito a ganhar em parcerias interdisciplinares.

Referências

AMANTE, M. J. **Indicadores de visibilidade, impacto e influência científica: a auto-citação na análise e avaliação da investigação científica**, p.1-23, 2011. Disponível em: <https://repositorio.iscte-iul.pt/handle/10071/2329>. Acesso em: 20 mar. 2023.

ARAGÃO, R. M. L. de. Modelos de estruturação do artigo científico: retrato e discussão a partir de instruções aos autores da SciELO Brasil. **Cadernos de Letras da UFF**, v. 43, p. 153-163, 2011.

BICAS, H.; ROTHER, E.; BRAGA, M. E. Fatores de impacto, outros índices bibliométricos e desempenhos acadêmicos. **Arquivos Brasileiros de Oftalmologia**, v. 65, n. 2, p. 151-152, 2002.

BOCH, F; GROSSMANN, F. Referir-se ao discurso do outro: Alguns Elementos de Comparação entre Especialistas e Principiantes. **Scripta**, Belo Horizonte, v. 6, n. 11, p. 97-108, 2002.

BOCH, F; GROSSMANN, F. Sobre o uso de citações no discurso teórico: de constatações a proposições didáticas. **Letramento e formação universitária: formar para a escrita e pela escrita**. Campinas: Mercado de Letras, p. 283-307, 2015.

CAPES. **CAPES melhora ferramentas de avaliação da pós-graduação**. 2019. Disponível em: <https://www.gov.br/capes/pt-br/assuntos/noticias/capes-melhora-ferramentas-de-avaliacao-da-pos-graduacao>. Acesso em: 22 mar. 2023.

CLARES, L. M. **Mediação editorial na comunicação científica: um estudo de dois periódicos de humanidades**. Dissertação (Mestrado em Linguística) – Universidade Federal de São Carlos, São Carlos, 2017.

CLARIVATE ANALYTICS. **Research in Brazil: Funding excellence – Analysis prepared on behalf of CAPES by the Web of Science Group**. Disponível em: https://jornal.usp.br/wp-content/uploads/2019/09/ClarivateReport_2013-2018.pdf. Acesso em: 19 mar. 2023.

CLARIVATE ANALYTICS. **SciELO Citation Index Ajuda**. Disponível em: http://images.webofknowledge.com/WOKRS5132R4.2/help/pt_BR/SCIELO/hp_whatsnew_scielo.html. Acesso em: 19 mar. 2023.

CLARIVATE ANALYTICS. **Web of Science**. Disponível em: http://www-periodicos-capes-gov-br.ezl.periodicos.capes.gov.br/?option=com_plogin. Acesso em: 19 mar. 2023.

CONSELHO NACIONAL DE DESENVOLVIMENTO CIENTÍFICO E TECNOLÓGICO. **Dados abertos**. Disponível em: <http://memoria2.cnpq.br/web/guest/apresentacao1>. Acesso em: 19 mar. 2023.

CONSELHO NACIONAL DE DESENVOLVIMENTO CIENTÍFICO E TECNOLÓGICO. **Indicadores por grande área do conhecimento**. Disponível em: <http://memoria2.cnpq.br/web/guest/indicadores1>. Acesso em: 19 mar. 2023.

CORRÊA, M. L. G. As perspectivas etnográfica e discursiva no ensino da escrita: o exemplo de textos de pré-universitários. **Revista da ABRALIN**, v. 10, n. 4, p. 333-356, 2011. Disponível em: <https://revista.abralin.org/index.php/abralin/article/view/1115>. Acesso em: 19 mar. 2023.

COSTAS, R.; VAN LEEUWEN, T.; BORDONS, M. Self-citations at the meso and individual levels: effects of different calculation methods. **Scientometrics**, v. 82, n. 3, p. 517-519, 2010.

DAUNAY, B. Metáfrase e paráfrase: modalidades da apropriação do discurso de outrem na escrita acadêmica. **Linguagem em (Dis)curso**, v. 20, n. 2, p. 363-380, 2020.

DAUNAY, B.; DELCAMBRE, I. Les modalités énonciatives de la reprise du discours d'autrui dans les écrits de recherche et les écrits didactiques. **Scripta**, Belo Horizonte, n. 21/43, p. 37-64, 2017. Disponível em: <http://periodicos.pucminas.br/index.php/scripta/article/view/14232/12818>. Acesso em: 20 mar. 2023.

JESUS, J. F. H. de. **Uso dos operadores metadiscursivos em exemplares do gênero artigo científico produzidos por estudantes de graduação**. Tese (Doutorado em Letras) – Universidade do Estado do Rio Grande do Norte, Mossoró, 2020.

FISCHER, A.; FERREIRA, K. M.; SILVA, R. Escrita acadêmica em artigos científicos: autocitação em diferentes áreas disciplinares. **Revista on line de Política e Gestão Educacional**, Araraquara, v. 24, n. 3, p. 1257-1271, 2020.

GARCIA, F; ESTÊVÃO, C. Uma experiência de análise de conteúdo com uso de QDA's. **CIAIQ2016**, v. 4, p. 45-52, 2016.

GRUSZYNSKI, A.; GOLIN, C.; CASTEDO, R. Produção editorial e comunicação científica: uma proposta para edição de revistas científicas. **E-Compós**, v. 11, n. 2, p. 1-17, 2009. Disponível em: <https://www.e-compos.org.br/e-compos/article/view/238>. Acesso em: 21 mar. 2023.

HYLAND, K. Self-Citation and Self-Reference: Credibility and Promotion in Academic Publication. **Journal of the American Society for Information Science and Technology**, v. 54, n. 3, p. 251-259, 2003.

HYLAND, K. Disciplinary differences: Language variation in academic discourses. In: HYLAND, K.; BONDI, M. (Orgs.). **Academic discourse across disciplines**. Frankfurt: Peter Lang, 2006, p. 17-45.

HYLAND, K.; TSE, P. Metadiscourse in academic writing: a reappraisal. **Applied Linguistics**, v. 25, n. 2, p. 156-177, 2004.

- HYLAND, K.; JIANG, K. Changing patterns of self-citation: Cumulative inquiry or selfpromotion? **Text and Talk**, v. 38, n. 3, p. 365–387, 2018.
- IOANNIDIS, J. P. A. A generalized view of self-citation: Direct, co-author, collaborative, and coercive induced self-citation. **Journal of Psychosomatic Research**, v. 78, p. 7-11, 2015.
- KLEIMAN, A. B. (org.) **Os significados do letramento: uma nova perspectiva sobre a prática social da escrita**. Campinas: Mercado de Letras, 1995.
- KOMESU, F.; ASSIS, J. A. Artigos científicos publicados em periódicos brasileiros e franceses de alto impacto na subárea de Linguística: o que números de citação (não) mostram. In: LARANJEIRA, R. de M.; MIRANDA, F. D. S. S.; PARIS, L. G. (org.) **Letramentos Acadêmicos no Brasil: diálogos e mediações em homenagem a Raquel Salek Fiad**. 1ed. São Carlos: Pedro & João Editores, 2022, v. 1, p. 39-59.
- LEA, M. R.; STREET, B. V. Student writing in higher education: an academic literacies approach. **Studies in higher education**, v. 23, n. 2, p. 157-170, 1998.
- LEA, M.; STREET, B. O modelo de “letramentos acadêmicos”: teoria e aplicações. Tradução: Fabiana Komesu e Adriana Fischer. **Filologia e Linguística portuguesa**, v. 16, n. 2, p. 477-493, 2014. Disponível em: <https://www.revistas.usp.br/flp/article/view/79407>. Acesso em: 21 mar. 2023.
- LILLIS, T.; CURRY, M. **Academic writing in a global context: The politics and practices of publishing in English**. New York: Routledge, 2011.
- MACEDO, T. S. de; PAGANO, A. S. Análise de citações em textos acadêmicos escritos. **DELTA: Documentação de Estudos em Linguística Teórica e Aplicada**, v. 27, n. 2, p. 257-288, 2011.
- MARZIALE, M. H. P.; MENDES, I. A C. O fator de impacto das publicações científicas. **Revista Latino-Americana de Enfermagem**, v. 10, n. 4, p. 466-467, 2002.
- McGRATH, L.; KUTEEVA, M. Stance and engagement in pure mathematics research articles: Linking discourse features to disciplinary practices. **English for Specific Purposes**, v. 31, n. 3, p. 161-173, 2012.
- MELO RIBEIRO, H. C. Bibliometria: quinze anos de análise da produção acadêmica em periódicos brasileiros. **Biblios**, n. 69, p. 1-20, 2017. Disponível em: http://www.scielo.org.pe/scielo.php?script=sci_arttext&pid=S1562-47302017000400001&lng=es&nrm=iso. Acesso em: 21 mar. 2023.
- SALGADO, L.; CLARES, L. Mediação editorial em artigos científicos: um estudo de injunções e apagamentos nas humanidades. **Revista do GEL**, v. 14, n. 3, p. 29-58, 2017.
- SANTOS, R. N. M.; KOBASHI, N. Y. Bibliometria, Cientometria, Infometria: conceitos e aplicações. **Tendências da Pesquisa Brasileira em Ciência da Informação**, v. 2, n. 1, p. 155-172, 2009.

SCIENTIFIC ELECTRONIC LIBRARY ONLINE. Disponível em: <https://scielo.org>. Acesso em: 22 mar. 2023.

SILVA, J. A.; BIANCHI, M. L. P. Cientometria: a métrica da ciência. **Paidéia**, v. 11, p. 5-10, 2001.

SOARES, M. **Letramento**: um tema em três gêneros. Belo Horizonte: Autêntica, 2003.

STREET, B. **Literacy in theory and practice**. New York: Cambridge University Press, 1984.

STREET, B. **Literacy and Development**: Ethnographic Perspectives. London and New York: Routledge, 2001.

STREET, B. Dimensões “escondidas” na escrita de artigos acadêmicos. Tradução: Armando Silveiro e Adriana Fischer. **Perspectiva**, v. 28, n. 2, p. 541-567, 2010.

SWALES, J. **Genre Analysis**: English in Academic and Research Settings. Cambridge: Cambridge University Press, 1990.

TESSUTO, G. Generic structure and rhetorical moves in English-language empirical law research articles: Sites of interdisciplinary and interdiscursive cross-over. **English for Specific Purposes**, v. 37, p. 13-26, 2015.

THOMAZ, P. G.; ASSAD, R. S.; MOREIRA, L. F. P. Uso do Fator de impacto e do índice H para avaliar pesquisadores e publicações. **Arquivos Brasileiros de Cardiologia**, v. 96, n. 2, p. 90-93, 2011.

VERBI. **MAXQDA 2022 Manual**. Disponível em: <https://www.maxqda.com/help-mx22/welcome>. Acesso em: 21 mar. 2023.

WHITE, H. D. Authors as citers over time. In: **Journal of the American Society for Information Science and Technology**, v. 52, n. 2, p. 87-108, 2001.