

UNIVERSIDADE FEDERAL DE MINAS GERAIS
Faculdade de Letras
Programa de Pós-graduação em Estudos Linguísticos

Luanna Karoline de Souza Horta Mendonça

ESCRITA E AUTOPÓS-EDIÇÃO DE RESUMOS ACADÊMICOS EM LÍNGUA
ADICIONAL: uma análise de elementos retóricos em corpus comparável

Belo Horizonte
2025

Luanna Karoline de Souza Horta Mendonça

**ESCRITA E AUTOPÓS-EDIÇÃO DE RESUMOS ACADÊMICOS EM LÍNGUA
ADICIONAL: uma análise de elementos retóricos em corpus comparável**

Dissertação apresentada ao Programa de Pós-Graduação em Estudos Linguísticos da Universidade Federal de Minas Gerais como requisito parcial para obtenção do título de Mestra em Estudos Linguísticos.

Área de concentração: Linguística Aplicada
Linha de pesquisa: Estudos da Tradução
Orientador: Prof. Dr. Leonardo Pereira Nunes

Belo Horizonte
2025

M539e Mendonça, Luanna Karoline de Souza Horta.
Escrita e autopós-edição de resumos acadêmicos em língua adicional [manuscrito] : uma análise de elementos retóricos em corpus comparável / Luanna Karoline de Souza Horta Mendonça. – 2025.
1 recurso online (84 f. : il., tabs., maps. (algumas color.)) : pdf.

Orientador: Leonardo Pereira Nunes.

Área de concentração: Linguística Aplicada.

Linha de Pesquisa: Estudos da Tradução.

Dissertação (mestrado) – Universidade Federal de Minas Gerais, Faculdade de Letras.

Bibliografia: f. 82-84.

Exigências do sistema: Adobe Acrobat Reader.

1. Tradução e interpretação – Teses. 2. Tradução automática – Teses. 3. Redação acadêmica – Teses. 4. Linguística de corpus – Teses. 5. Resumos – Redação – Teses. I. Nunes, Leonardo Pereira. II. Universidade Federal de Minas Gerais. Faculdade de Letras. III. Título.

CDD : 418.02



UNIVERSIDADE FEDERAL DE MINAS GERAIS
FACULDADE DE LETRAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ESTUDOS LINGÜÍSTICOS

FOLHA DE APROVAÇÃO

Escrita e autopós-edição de resumos acadêmicos em língua adicional: uma análise de elementos retóricos em corpus comparável

LUANNA KAROLINE DE SOUZA HORTA MENDONÇA

Dissertação submetida à Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em Estudos Linguísticos, como requisito para obtenção do grau de Mestra em Estudos Linguísticos, área de concentração Linguística Aplicada, linha de pesquisa Estudo da Tradução.

Aprovada em 19 de dezembro de 2025, pela banca constituída pelos membros:

Prof. Dr. Leonardo Pereira Nunes - Orientador
UFMG

Profa. Dra. Norma Barbosa de Lima Fonseca
Colégio Militar de Belo Horizonte

Profa. Dra. Bárbara Malveira Orfanò
UFMG

Belo Horizonte, 19 de dezembro de 2025.



Documento assinado eletronicamente por **Leonardo Pereira Nunes, Professor do Magistério Superior**, em 29/12/2025, às 19:17, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Barbara Malveira Orfano, Professora do Magistério Superior**, em 08/01/2026, às 06:53, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Norma Barbosa de Lima Fonseca, Usuário Externo**, em 08/01/2026, às 11:50, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **4820585** e o código CRC **4654F4BD**.

AGRADECIMENTOS

Dizem que pesquisar é um ato solitário e, na maioria das vezes, é verdade. Mas um pesquisador não se cria sozinho; ele é o resultado de uma soma de vários componentes: seus parceiros, família, amigos, professores, colegas e experiências. Como pesquisadora em formação, não me criei sozinha e gostaria de começar agradecendo às pessoas que somam o que sou e me completam no que ainda me torno. Ao Gabriel Santos Mendonça, pelo apoio até nos momentos mais estressantes, pela ajuda para criar planilhas funcionais e por todo o amor. Aos meus amigos, que entenderam os momentos de ausência. À Daniele Auad e ao Rafael Nascimento, pelo acolhimento infinito na hora do desespero e pelo aconchego na hora do descanso, e à Letícia Guimarães, por todo auxílio acadêmico. Ao Leonardo Nunes, meu orientador, pelos ensinamentos; aos membros da banca, Profa. Dra. Bárbara Malveira Orfanò, Profa. Dra. Norma Barbosa de Lima Fonseca e Profa. Dra. Climene Fernandes Brito Arruda, pela disponibilidade; e à CAPES, pelo auxílio financeiro.

“The writing of history is a process of constant correction [...]”

W. G. Sebald

RESUMO

O inglês, enquanto língua franca da comunicação científica, impõe desafios significativos a pesquisadores não anglófonos, que precisam produzir e divulgar conhecimento em uma língua adicional. Nesse contexto, a tradução automática tem se consolidado como uma alternativa acessível para apoiar a escrita acadêmica, especialmente quando associada à autopós-edição, entendida como a revisão realizada pelo próprio autor após a geração do texto pela máquina. Dando continuidade às investigações de Nunes (2021), o presente estudo analisa a organização retórica em um *corpus* comparável composto por resumos acadêmicos escritos em inglês como língua adicional (L2) e por resumos autopós-editados nessa mesma língua a partir de traduções automáticas de textos originalmente escritos em português brasileiro (L1). A pesquisa fundamenta-se no modelo CARS (*Create a Research Space*), em sua versão expandida, e examina a distribuição de movimentos e passos retóricos em um *corpus* comparável composto por 36 resumos produzidos por doutorandos da Universidade Federal de Minas Gerais, pertencentes às áreas de Ciências Exatas e da Terra, Ciências Humanas e Ciências Sociais Aplicadas. Os textos foram produzidos em contexto empírico-experimental e analisados por meio de classificação manual e automática, com o auxílio da ferramenta INCEpTION. Os resultados indicam a predominância do Movimento 3, relacionado à apresentação de objetivos, métodos e resultados, em todos os modos de produção textual, enquanto os Movimentos 1 e 2, associados à contextualização do tema e à construção de nichos discursivos, apresentaram ocorrência limitada e distribuição irregular, variando conforme a área do conhecimento. A comparação entre textos escritos diretamente em L2 e textos traduzidos automaticamente e autopós-editados nessa língua revelou que esta última atuou principalmente como mecanismo de refinamento textual, favorecendo a fluidez e a clareza informacional, sem promover alterações substanciais na organização retórica global dos resumos. Os achados reforçam a necessidade de treinamento em termos de organização retórica para a produção textual em L2 e de letramento em gêneros acadêmicos, além de contribuírem para a compreensão da tradução automática como uma prática situada no contexto da escrita acadêmica contemporânea.

Palavras-chave: *corpus* comparável; escrita acadêmica em L2; tradução automática; autopós-edição; movimentos e passos retóricos.

ABSTRACT

English, as the dominant language of scientific communication, poses significant challenges for non-Anglophone researchers, who are required to read, conduct research, and publish in an additional language. In this context, machine translation has emerged as an accessible alternative to support academic writing, particularly when combined with self-post-editing, here understood as the revision carried out by the author after the machine-generated translation. Building on Nunes' 2021 research project, this study investigates the rhetorical organization in a comparable *corpus* composed of academic abstracts written in English as an additional language (L2) and of abstracts self-post-edited in the same language from machine translations of texts originally written in Brazilian Portuguese (L1). Grounded in the expanded CARS (*Create a Research Space*) model, the study examines the distribution of rhetorical moves and steps in a comparable *corpus* comprising 36 abstracts produced by doctoral students from the Federal University of Minas Gerais in the fields of Exact and Earth Sciences, Humanities, and Applied Social Sciences. The texts were produced in an empirical-experimental context and analyzed through both manual and automatic classification using the INCEpTION annotation tool. The results reveal a consistent predominance of Move 3, associated with the presentation of objectives, methods and results, across all modes of textual production, whereas Moves 1 and 2, related to topic contextualization and niche construction, occur less frequently and show irregular distribution across disciplinary areas. The comparison between texts written directly in L2 and machine-translated and self-post-edited texts in this language indicates that the latter primarily functions as a mechanism for textual refinement, enhancing fluency and informational clarity without substantially altering the overall rhetorical organization of abstracts. These findings highlight the importance of training for rhetorical organization within text production and of genre-based literacy so as to contribute to a broader understanding of machine translation as a practice within contemporary academic writing.

Keywords: comparable *corpus*, L2 academic writing; machine translation; self post-editing; rhetorical moves and steps.

LISTA DE ILUSTRAÇÕES

Figuras

Figura 1. O mapa do Estudos da Tradução.	17
Figura 2. Mapa de Estudos da Tradução Aplicados.....	18
Figura 3. Exemplo da abordagem geral de TAN.	24
Figura 4. Comparação entre RA e RC.	29
Figura 5. O modelo seminal de Swales.....	32
Figura 6. Modelo de Sanchez.....	34
Figura 7. Modelo expandido de Britto et al.	37
Figura 8. Ordens possíveis de realização das tarefas de produção textual	41
Figura 9. Planilha digital de controle - base de dados	44
Figura 10. Organização das pastas	45
Figura 11. Organização dos arquivos com seus códigos de nomenclatura.	45
Figura 12. Alinhamento dos textos.	45
Figura 13. <i>Corpus</i> classificado e validado no Google drive.	46
Figura 14. Interface dos projetos disponíveis no INCEpTION.	48
Figura 15. Interface dos conjuntos de etiquetas.....	49
Figura 16. Interface da Camada criada e suas propriedades.	50
Figura 17. Processo de classificação manual no INCEpTION.	52
Figura 18. Interface das configurações da Recomendação.	53
Figura 19. Interface das configurações de anotação.	53
Figura 20. Interface das preferências de anotação	54
Figura 21. Interface das classificações automáticas do INCEpTION.....	54

Tabelas

Tabela 1. Divisão geral dos resumos por tarefa e área de conhecimento.	42
Tabela 2. Distribuição das correspondências entre os assuntos versados em cada <i>subcorpus</i> por área do conhecimento.....	43
Tabela 3. Divisão dos Textos.....	51
Tabela 4. Dados gerais coletados.....	55
Tabela 5. Divisão dos textos em EL2 e PEL2.....	56
Tabela 6. Movimentos iguais em EL2 e PEL2.	60

Tabela 7. M1 e seus passos iguais.....	63
Tabela 8. M2 e seus passos iguais.....	64
Tabela 9. M3 e seus passos iguais.....	65
Tabela 10. Dados gerais da comparação entre TA e PEL2.....	67
Tabela 11. M1 e seus passos.	68
Tabela 12. M2 e seus passos.	70
Tabela 13. M3 e seus passos.	71
Tabela 14. Resultados gerais do IpN.	73
Tabela 15. Correspondência entre M1 e seus passos.....	74
Tabela 16. Correspondência entre M2 e seus passos.....	75
Tabela 17. Correspondência entre M3 e seus passos.....	76

LISTA DE ABREVIATURAS E SIGLAS

CARS	<i>Create a Research Space</i>
CET	Ciências Exatas e da Terra
CH	Ciências Humanas
CNPq	Conselho Nacional de Desenvolvimento Científico e Tecnológico
CSA	Ciências Sociais Aplicadas
EL1	Escrita na Primeira Língua
EL2	Escrita na Língua Adicional
FNI	Falante Nativo de Inglês
IFA	Inglês para Fins Acadêmicos
IFE	Inglês para Fins Específicos
ILF	Inglês como Língua Franca
ILFCA	Inglês como Língua Franca do Contexto Acadêmico
IPA	Inglês para Propósitos Acadêmicos
IpN	INCEpTION
JDK	<i>Java Development Kit</i>
JRE	<i>Java Runtime Environment</i>
JVM	<i>Java Virtual Machine</i>
LA	Língua Alvo
LC	Linguística de <i>Corpus</i>
LF	Língua Fonte
L1	Primeira Língua
L2	Segunda Língua / Língua Adicional

M1	Movimento 1
M2	Movimento 2
M3	Movimento 3
NFNI	Não Falante Nativo de Inglês
ONU	Organização das Nações Unidas
PEL2	Autopós-edição em L2
PT-BR	Português Brasileiro
RA	Resumo de Artigo
RC	Resumo de Conferência
SVM	<i>Support Vector Machine</i>
TA	Tradução automática
TAN	Tradução Automática Neural
TBC	Tradução Baseada em <i>Corpora</i>
TBE	Tradução Baseada em Exemplos
TBR	Tradução Baseada em Regras
TE	Tradução Estatística
UFMG	Universidade Federal de Minas Gerais
UFRGS	Universidade Federal do Rio Grande do Sul
UFSC	Universidade Federal de Santa Catarina
UnB	Universidade de Brasília

SUMÁRIO

1 INTRODUÇÃO	14
2 REFERENCIAL TEÓRICO	19
2.1 O inglês na ciência	19
2.1.1 Letramento acadêmico	21
2.2 Abordagens tradutórias	23
2.3 A autopós-edição.....	25
2.4 Gêneros e a escrita acadêmica	27
2.5 O modelo CARS e a análise de gêneros acadêmicos.....	30
2.6 Linguística de <i>Corpus</i>	38
3 METODOLOGIA	40
3.1 Metodologia de coleta.....	40
3.2 O desenho do <i>corpus</i>	42
3.3 Metodologia de análise do <i>corpus</i>	43
3.3.1 – Classificação automática – o INCEpTION.....	47
4 APRESENTAÇÃO E DISCUSSÃO DOS RESULTADOS	55
4.1 EL2 e PEL2 – Dados gerais	56
4.1.1 EL2 e PL2 – Correspondências.....	60
4.2 Tradução automática e sua autopós-edição.....	67
4.3 A classificação automática do INCEpTION.....	72
5 CONSIDERAÇÕES FINAIS.....	77
6 REFERÊNCIAS.....	82

1 INTRODUÇÃO

Durante a história, uma língua exerceu domínio quando se refere à criação e à propagação de ciência no Ocidente (Walter, 1996 apud Hamel, 2007) e, desde o pós-guerra, este espaço tem sido ocupado pelo inglês. Sua difusão como língua franca de publicações internacionais deve-se a alguns fatores como a Declaração de Bolonha de 1999, a criação da Organização das Nações Unidas (ONU), sediada em Nova York, e a invenção do computador e suas linguagens.

O uso do inglês nas publicações científicas impõe uma nova realidade para pesquisadores não nativos de países anglófonos que precisam produzir ciência em uma língua estrangeira. Para lidar com essa dificuldade, existem algumas opções, como aprender inglês, um processo que exige tempo, ou contratar tradutores profissionais, o que pode ser uma alternativa onerosa (Bower; Buitrago Ciro, 2019). Entretanto, o pós-guerra também proporcionou diversos avanços tecnológicos e um deles apresenta-se como alternativa viável e econômica: a tradução automática.

A tradução automática (TA) despertou o interesse do público em 1949 com a divulgação do Memorando Weaver, que, inspirado pelas técnicas de criptografia desenvolvidas durante a Segunda Guerra Mundial, sugeria a possibilidade de realizar traduções entre línguas por meio do computador. Desde então, o desenvolvimento da TA foi marcado por avanços e retrocessos. Entre os resultados considerados negativos desse período inicial destaca-se o relatório ALPAC, publicado em 1966, que avaliou criticamente as pesquisas em tradução automática da época e concluiu que os sistemas então disponíveis não atendiam às expectativas de qualidade, rapidez e viabilidade econômica, o que resultou na redução do financiamento e na interrupção de diversas pesquisas na área (Bower; Buitrago Ciro, 2019).

Apesar desse cenário desfavorável, experiências desenvolvidas em contextos específicos mostraram que a tradução automática podia funcionar de maneira operacional quando aplicada a domínios restritos. Um dos exemplos mais conhecidos é o sistema MÉTÉO, desenvolvido no Canadá na década de 1970 para a tradução de previsões meteorológicas do inglês para o francês. Seu desempenho positivo esteve diretamente relacionado às características do gênero textual tratado, marcado por forte padronização, vocabulário controlado e estruturas sintáticas recorrentes, condições que reduziram ambiguidades e

permitiram maior previsibilidade no processamento automático. Esse cenário abriu espaço para o desenvolvimento da TA a partir dos anos 2000 e suas diversas abordagens.

A primeira abordagem foi a Tradução Baseada em Regras (TBR), popular até o início dos anos 2000 e que, inicialmente, substituía palavra por palavra entre idiomas. Tal abordagem foi eficaz em domínios especializados, pois gerava traduções precisas quando as regras eram bem definidas. Porém, apresentava falhas gramaticais e, com o tempo, a complexidade e a rigidez da TBR levaram à adoção de métodos mais eficientes.

Um exemplo é a Tradução Baseada em *Corpora* (TBC), que apresenta dois tipos: i. a tradução baseada em exemplos (TBE), que utiliza um *corpus* paralelo e busca frases semelhantes para a tradução de novos textos e a ii. tradução estatística (TE), que usa *corpora* bilíngues para prever a probabilidade de qual seria a melhor tradução. Ambos apresentam falhas quando os *corpora* não são extensos o suficiente ou são de domínios diferentes, pois a TBC não é adaptável.

O modelo mais atual, a Tradução Automática Neural (TAN), gera traduções mais contextuais e naturais e utiliza *corpora* bilíngues. A TAN funciona codificando a frase original e decodificando para o idioma desejado. No entanto, ela precisa de um volume grande de dados e tem dificuldade ao trabalhar em domínios específicos, como a TE.

Os avanços apresentados pela TA, apesar de promissores e se manifestarem como uma alternativa para quem pesquisa e publica em inglês sem o domínio do idioma, não entregam os resultados necessários sozinhos. Como, então, utilizar a tradução automática para produzir e publicar em outro idioma? A resposta está em combinar a tradução automática com a autopós-edição¹ (Bower; Buitrago Ciro, 2019).

Neste contexto, pesquisas foram desenvolvidas para entender como a combinação de tradução automática com a autopós-edição auxiliam na escrita acadêmica (Bower; Buitrago Ciro, 2019). Tal foi a investigação conduzida por Nunes (2021), que propõe a elaboração de resumos acadêmicos escritos em inglês como língua adicional (L2) e em português como primeira língua (L1). Ainda, propõe a tradução automática dos resumos escritos em L1 e a posterior autopós-edição pelos seus respectivos autores.

¹ Aqui, autopós-edição será definida como a revisão feita pelo autor do texto após sua tradução automática.

Ainda que pesquisas sobre a combinação de tradução automática e autopós-edição como auxílio à escrita acadêmica estejam em andamento, poucas têm se dedicado a analisar elementos textuais de resumos. Ao submeter um trabalho para publicação ou apresentação, o resumo é o primeiro contato com o texto e é através dele que a decisão de continuar a ler ou não um trabalho acadêmico-científico é tomada. Tal fato torna, por exemplo, o artigo na íntegra ou a apresentação de uma contribuição em uma conferência algo interessante a ser explorado (Hyland, 2004).

Em consonância com Britto *et al.* (2021), que automatiza a identificação de padrões retóricos baseados no modelo CARS (*Create a Research Space*), trabalho seminal de Swales (1990), além das considerações propostas por Hyland (2004), o presente trabalho investiga os movimentos e passos retóricos em um *corpus* comparável composto por resumos escritos por doutorandos da Universidade Federal de Minas Gerais em inglês como L2 e por resumos autopós-editados nessa mesma língua a partir de textos desse gênero escritos em português como L1. Mais especificamente, o estudo objetiva:

- 1) averiguar em que medida os movimentos retóricos presentes nos textos escritos em inglês como língua adicional L2 (EL2) se aproximam ou se distanciam dos movimentos encontrados nos textos autopós-editados em inglês L2 (PEL2).
- 2) averiguar em que medida os passos retóricos presentes nos textos escritos em inglês como língua adicional L2 (EL2) se aproximam ou se distanciam dos passos encontrados nos textos autopós-editados em inglês L2 (PEL2).

Não obstante, pretende responder às seguintes perguntas:

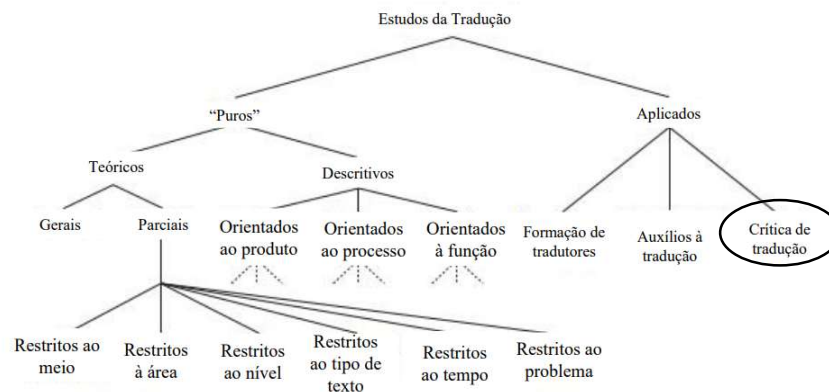
- 1) Quantos e quais são os movimentos retóricos presentes nos textos escritos em L2? Quantos e quais são os mesmos movimentos presentes nos textos autopós-editados em L2?
- 2) Quantos e quais são os passos retóricos presentes nos textos escritos em L2? Quantos e quais são os mesmos passos presentes nos textos autopós-editados em L2?

Ao sugerir um estudo que investiga práticas concretas de uso da linguagem em contexto acadêmico, considerando a produção de textos científicos em língua adicional mediada por tradução automática, o presente trabalho se insere na interface entre no campo da Linguística Aplicada, na medida em que aborda um problema recorrente da comunicação científica contemporânea. Ao tomar como objeto a escrita e a autopós-edição de resumos acadêmicos

produzidos por pesquisadores em formação, a pesquisa dialoga com questões diretamente relacionadas às condições de circulação do conhecimento em contextos não anglófonos e às demandas institucionais de internacionalização.

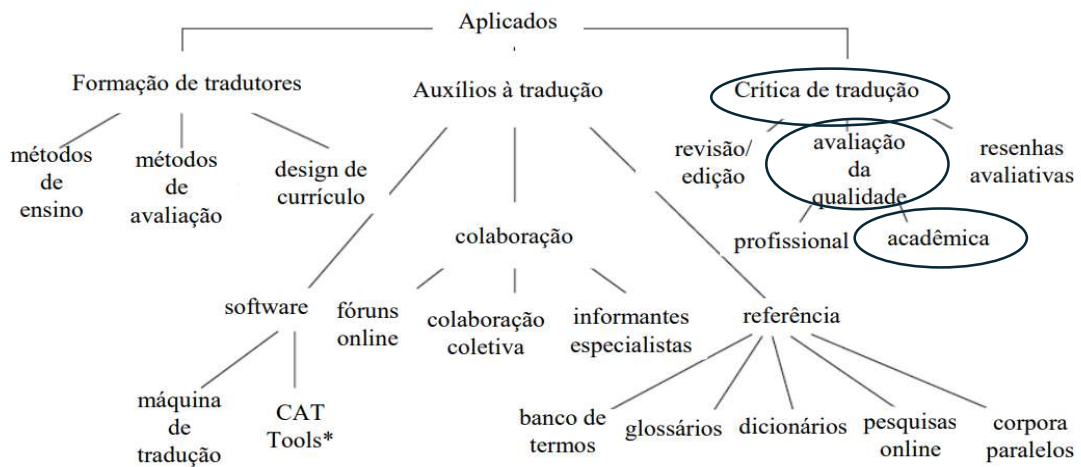
Ainda, a investigação se vincula aos Estudos da Tradução Aplicados, conforme delineados no mapa proposto por James Holmes (1972) (figura 1) e expandido por Jeremy Munday (2016) (figura 2), mais especificamente ao campo da crítica de tradução, entendida, neste estudo, como uma avaliação da qualidade tradutória ancorada em critérios situados de uso acadêmico. A análise proposta não se orienta por critérios prescritivos nem por concepções estritas de equivalência, mas pela adequação retórico-discursiva dos textos traduzidos e autopós-editados às convenções do gênero resumo acadêmico.

Figura 1. O mapa do Estudos da Tradução.



Fonte: Munday, 2016. Traduzido por Francieli S. Oliveira.

Figura 2. Mapa de Estudos da Tradução Aplicados.



Fonte: Munday, 2016. Traduzido por Francieli S. Oliveira.

Ao articular escrita acadêmica, tradução automática e análise de gêneros, a pesquisa contribui, do ponto de vista acadêmico, para a ampliação da compreensão sobre os efeitos dessas práticas na organização retórica de textos científicos e, do ponto de vista social, para uma reflexão mais crítica e informada sobre o uso da tradução automática por pesquisadores em formação.

Além da presente Introdução, este trabalho está organizado em mais quatro seções. Na seção seguinte, apresenta-se o referencial teórico que fundamenta o estudo. A terceira seção descreve os procedimentos metodológicos adotados na pesquisa. Na sequência, a quarta seção dedica-se à apresentação e à discussão dos resultados. Por fim, a quinta seção encerra o trabalho com as considerações finais.

2 REFERENCIAL TEÓRICO

2.1 O inglês na ciência

O inglês se consolidou como a língua franca da atualidade, ocupando uma posição central nas publicações acadêmicas e em conferências internacionais. Esse processo se relaciona tanto à posição histórica do Reino Unido como potência colonial quanto, sobretudo, à hegemonia dos Estados Unidos no contexto do pós-guerras, especialmente após a Segunda Guerra Mundial, quando o país assumiu protagonismo nos desenvolvimentos tecnológico e científico.

Todavia, a adoção do inglês como língua franca também pode ser compreendida a partir da intensificação dos processos de globalização e da necessidade de uma língua comum que facilitasse as interações em escala internacional, uma vez que, conforme observa Crystal (1997a, 1997b; apud Kaplan, 2001), no período atual há mais pessoas que utilizam o inglês como língua materna ou adicional do que qualquer outra língua, à exceção do chinês.

Embora o inglês como língua franca (ILF) apresente vantagens evidentes em termos de comunicação científica, a predominância de um único idioma gera impactos relevantes tanto para falantes nativos quanto para não falantes. A expressão *publish or perish* (publique ou pereça) circula amplamente no meio acadêmico e sintetiza a lógica de produtividade que orienta a carreira científica atual. Esse processo ocorre majoritariamente em inglês e, de acordo com De Swaan (2001), grande parte dos periódicos de maior visibilidade publica seus conteúdos nessa língua.

Esse cenário se configura como um obstáculo adicional para pesquisadores não falantes nativos de inglês (NFNI), uma vez que avaliações acadêmicas tendem a adotar parâmetros linguísticos associados ao falante nativo de inglês (FNI), priorizando aspectos gramaticais e lexicais rígidos e deixando em segundo plano os contextos socioculturais e acadêmicos nos quais o conhecimento é produzido (De Swaan, 2001; Bower; Buitrago Ciro, 2019; Jenkins, 2014).

Diante desse quadro, como lidar com a barreira linguística quando deixar de publicar não é uma possibilidade? Nesse sentido, Jenkins (2014) trata o inglês como a língua de ninguém e de todos, ao compreendê-lo como a língua franca do contexto acadêmico (ILFCA), cuja

prioridade é a comunicação entre interlocutores, e não a adesão a um discurso gramaticalmente perfeito.

Para Jenkins (2014), essa mudança de foco não depende apenas do uso individual da língua, mas está diretamente relacionada a transformações no âmbito institucional, especialmente no que diz respeito às políticas linguísticas e às práticas pedagógicas adotadas pelas universidades. A autora observa que, apesar de o inglês funcionar como língua franca no meio acadêmico, o ensino da escrita acadêmica ainda se orienta, em grande medida, por modelos normativos.

Nesse contexto, Jenkins distingue duas abordagens recorrentes no ensino do inglês em contextos acadêmicos. A primeira é o Inglês para Propósitos Acadêmicos (IPA), que enfatiza normas linguísticas consolidadas, gêneros acadêmicos estabilizados e modelos discursivos associados a tradições anglófonas. Trata-se de uma abordagem frequentemente institucionalizada, voltada à adequação do pesquisador aos padrões consolidados da escrita acadêmica internacional.

A segunda é o Inglês para Fins Específicos (IFE), que se orienta pelas necessidades comunicativas concretas de áreas disciplinares específicas, considerando os usos situados da língua e as demandas reais de estudantes e pesquisadores em seus contextos de atuação (Jenkins, 2014). As universidades buscam responder aos desafios linguísticos impostos pela centralidade do inglês por meio da adoção de abordagens baseadas em IPA, IFE ou em uma combinação entre ambas.

No contexto brasileiro, a Universidade Federal de Minas Gerais (UFMG) constitui um exemplo desse cenário ao ofertar, por meio da Faculdade de Letras, o conjunto de disciplinas de Inglês para Fins Acadêmicos (IFA). O nível inicial exige uma proficiência mínima equivalente ao nível B1, e as disciplinas são organizadas de forma progressiva, com exigência de aprovação no nível anterior ou realização de teste de nivelamento, priorizando práticas de leitura, escrita e oralidade em contextos acadêmicos.

Outras universidades públicas brasileiras também passaram a investir em ações voltadas ao ensino de inglês para fins acadêmicos, sobretudo no contexto das políticas de internacionalização. Estudos como o de Bergoli, Bocorny e Pires de Freitas (2022) mostram que instituições federais, como a Universidade Federal do Rio Grande do Sul (UFRGS), a Universidade Federal de Santa Catarina (UFSC) e a Universidade de Brasília (UnB), passaram

a ofertar cursos e programas de IFA, em geral vinculados a iniciativas nacionais, como o Idiomas sem Fronteiras.

Essas iniciativas indicam que o Inglês para Fins Acadêmicos tem sido incorporado pelas universidades como uma forma de responder às exigências da comunicação científica. Ainda assim, tanto o Inglês para Propósitos Acadêmicos quanto o Inglês para Fins Específicos apresentam limites quando se considera o acesso a esse tipo de formação. O IPA, ao manter normas associadas ao falante nativo, tende a exigir percursos formativos longos e investimento contínuo em ensino especializado, o que, em geral, dificulta seu alcance para pesquisadores não falantes nativos (Jenkins, 2014; Hyland, 2004).

O IFE, por sua vez, depende de planejamento curricular específico, análise de necessidades, formação docente constante e condições institucionais adequadas para sua implementação (Dudley-Evans; St John, 1998; Hyland, 2004). Em muitas universidades, essas exigências acabam limitando a oferta desse tipo de curso e favorecendo, em geral, pesquisadores que já possuem maior domínio da língua inglesa.

Pesquisadores não falantes nativos que buscam publicar seus trabalhos enfrentam, em geral, custos adicionais, sobretudo com tradução e revisão, já que muitos periódicos, nacionais e internacionais, exigem textos em língua inglesa. Parte da produção científica desenvolvida fora dos centros anglófonos, assim, circula menos, o que reduz a presença de diferentes perspectivas no debate científico.

É a partir dessa discussão que o letramento em tradução automática passa a ser considerado neste trabalho. Ao oferecer suporte à escrita acadêmica em língua inglesa, a tradução automática pode contribuir para ampliar o acesso às práticas de publicação científica, desde que utilizada de forma crítica e informada.

2.1.1 Letramento acadêmico

O conceito de letramento acadêmico permite compreender a escrita e a leitura no contexto universitário como práticas que vão além do domínio de normas linguísticas ou de técnicas de produção textual. Diferentemente de abordagens que tratam a escrita acadêmica como um conjunto de habilidades individuais, o letramento acadêmico a entende como uma

prática social situada, marcada por convenções disciplinares, expectativas institucionais e formas específicas de produção do conhecimento.

Lea e Street (2014) definem o letramento acadêmico como o conjunto de práticas de leitura e escrita próprias da universidade, práticas que variam conforme os gêneros acadêmicos, as áreas do conhecimento e as comunidades discursivas. Escrever academicamente, nessa perspectiva, envolve aprender a participar desses modos específicos de produção do saber, e não apenas dominar regras gramaticais ou formatos textuais. A escrita acadêmica se vincula, assim, às epistemologias disciplinares e às formas pelas quais o conhecimento é avaliado e reconhecido no meio científico.

A escrita acadêmica, assim, se relaciona às formas pelas quais o conhecimento é produzido e avaliado em diferentes áreas. Essa compreensão se apoia nos estudos de Street (1984, 2003), que entende o letramento como uma prática social situada e culturalmente construída. No ambiente universitário, essa perspectiva permite compreender que dificuldades na escrita acadêmica não decorrem, necessariamente, de déficits individuais, mas do acesso desigual às práticas de letramento valorizadas na academia. O letramento acadêmico envolve, portanto, a familiarização com gêneros específicos, modos de argumentação, formas de organização retórica e expectativas discursivas que variam conforme a área do conhecimento.

Quando a escrita acadêmica ocorre em língua adicional, essas práticas passam a envolver desafios adicionais. Em programas de pós-graduação fora das áreas de Letras, Linguística ou afins, a escrita acadêmica costuma ser aprendida no próprio fazer científico, sem uma formação sistemática voltada aos gêneros e às convenções da escrita acadêmica. Ainda que existam iniciativas institucionais voltadas ao ensino de inglês acadêmico, essas ações nem sempre contemplam as demandas concretas de escrita nas diferentes áreas do conhecimento.

A tradução automática passa a integrar as práticas de escrita acadêmica em língua inglesa como um recurso de apoio ao processo de escrita, seja na elaboração inicial do texto, seja em momentos de revisão e reformulação. Esse uso não se limita a uma etapa técnica do processo, mas interfere na organização textual, nas escolhas retóricas e na forma como o texto acadêmico é construído ao longo da escrita.

A escrita acadêmica mediada por tradução automática pressupõe a capacidade de integrar esse recurso às práticas de produção textual de forma crítica e informada. Isso envolve decidir quando e como recorrer à tradução automática, bem como avaliar seus efeitos sobre o

texto ao longo do processo de escrita, sem delegar à ferramenta o controle da autoria ou das escolhas discursivas (O'Brien; Simard; Goulet, 2018; Bowker, 2020).

Essa compreensão é central para a presente pesquisa, que analisa resumos produzidos em inglês como língua adicional e resumos autopós-editados nessa língua a partir de textos originalmente escritos em português com apoio de tradução automática. Considerar esses textos como resultado de práticas de escrita mediadas pela tradução automática permite interpretar as escolhas observadas no *corpus* à luz das condições concretas de produção da escrita acadêmica e das estratégias adotadas pelos pesquisadores ao longo do processo de escrita.

2.2 Abordagens tradutórias

Qun & Xiaojun (2015) definem a tradução automática como ou um subcampo da linguística computacional ou do processamento natural da linguagem que investiga o uso de *software* para traduzir textos ou falas de uma língua natural para outra. Todavia, traduzir não é só passar de uma língua fonte (LF) para uma língua alvo (LA), é algo complexo baseado em: i. regras, ii. *corpora* e iii. redes neurais.

A TA baseada em regras (TBR) utiliza as estruturas formais linguísticas e conhecimentos contextuais das duas línguas para criar os respectivos pares. São três tipos de TBR:

1. Modelo direto: que realiza a tradução usando uma transferência lexical onde a sentença da LF é diretamente traduzida para a LA. Tem como vantagem a simplicidade e funciona razoavelmente bem entre línguas sintaticamente semelhantes e como desvantagem a impossibilidade de comparar línguas sintaticamente muito diferentes.
2. Modelo interlíngual: utiliza a semântica e a sintaxe como intermediários do processo tradutório ao pegar uma sentença da LF, criar uma representação e depois transformar em um equivalente da LA. Tem como vantagem o emprego em sistemas multilíngues e como desvantagem a complexidade de sua aplicação.
3. Modelo de transferência: envolve análises nos âmbitos semântico, sintático, morfológico e contextual e cria sentenças na LA depois de analisar, transferir e sintetizar as informações provenientes da LF. Tem como vantagem a possibilidade de utilizar sentenças mais complexas e como desvantagem a necessidade de uma descrição detalhada das correspondências.

A TA baseada em *corpora* (TBC) emprega diferentes coletâneas de textos para encontrar as melhores correspondências e se divide em duas frentes: i. baseada em exemplos e ii. baseada em estatística.

1. Tradução baseada em exemplos (TBE) recorre a um *corpus* paralelo para encontrar pares de tradução preexistentes e separa as sentenças para exemplos já alinhados. Tem como desvantagem a falta de uma análise semântica e como vantagens a facilidade de aumentar o *corpus* e a produção de traduções de qualidade.
2. Tradução estatística (TE) aplica probabilidade para fazer a correspondência da LF para a LA, resultando na sentença mais provável. Tem como desvantagem necessidade de um *corpus* grande e bem elaborado, já que depende disso para criar uma equivalência de qualidade, apresenta como vantagem a não relação direta com uma língua e, por tanto, pode ser utilizada em outros idiomas.

Atualmente, o modelo de tradução empregado é o neural (TAN) e consiste, basicamente, em transformar sentenças da LF para LA utilizando redes neurais, adotando uma estrutura unificada. Isso significa que uma única rede neural aprende a relação entre as sentenças, o que acontece em duas fases: codificação e decodificação. Observe a figura 3 abaixo.

Figura 3. Exemplo da abordagem geral de TAN.



Fonte: Adaptado de Luong (2016).

Na codificação, a rede neural recebe a sentença, transforma as informações em um vetor que condensa o significado. Já na fase de decodificação, o vetor é utilizado para gerar a tradução. A TAN tem como vantagem a qualidade das traduções, já que utiliza dependências semânticas e sintáticas, fazendo mais do que uma tradução de palavra por palavra, apresenta como desvantagem a dificuldade de adaptação para domínios diferentes.

Para o presente trabalho, parte do *corpus* analisado é composto por resumos acadêmicos em português brasileiro (PT-BR) produzidos em condições empírico-experimentais que envolveram o uso de ferramentas de tradução automática para gerar textos em inglês (EN) e o emprego do arcabouço da análise de gêneros originalmente proposta por John Swales.

2.3 A autopós-edição

A tradução automática, apesar de seus avanços recentes, não produz, de forma autônoma, textos plenamente adequados para publicação acadêmica, o que torna necessária algum tipo de intervenção humana. O'Brien, Simard e Goulet (2018) demonstram que, embora os sistemas contemporâneos, especialmente os baseados em tradução automática neural, produzam traduções mais fluídas e contextualmente adequadas do que modelos anteriores, eles não reorganizam o texto de modo a atender plenamente às convenções discursivas e retóricas dos gêneros acadêmicos em língua inglesa (p. 255).

Nessas condições, as limitações observadas não se situam no plano computacional, mas no plano discursivo, o que faz com que a intervenção humana deixe de ser uma etapa secundária e passe a constituir parte integrante do processo de produção textual mediado por tradução automática. Tradicionalmente, a revisão do texto gerado por tradução automática é realizada por meio da pós-edição profissional, conduzida por tradutores especializados, cujo papel consiste em corrigir e ajustar o texto traduzido com base em critérios de correção linguística, fluidez, registro e adequação às normas da língua de chegada.

No entanto, no contexto específico da escrita acadêmica por pesquisadores não nativos de inglês, emerge a autopós-edição como uma prática distinta, na qual o próprio autor do texto revisa e ajusta o produto da tradução automática. O'Brien, Simard e Goulet definem esse processo como aquele em que os próprios autores são responsáveis pela pós-edição de seus textos, integrando a tradução automática ao processo de escrita acadêmica (2018, p. 238).

A diferença entre pós-edição profissional e autopós-edição não se limita ao agente responsável pela revisão, mas afeta a dinâmica da mediação tradutória. Na pós-edição profissional, o tradutor atua como mediador linguístico externo ao processo de produção do conhecimento, mobilizando competências tradutórias, estratégias de pesquisa terminológica e normas editoriais para adequar o texto à língua-alvo. Já na autopós-edição, o autor acumula os

papéis de escritor, tradutor e revisor, operando a partir de um conhecimento aprofundado do conteúdo científico que está sendo comunicado.

Essa diferença mostra-se particularmente relevante no contexto acadêmico, no qual a clareza conceitual e a estabilidade terminológica são centrais para a circulação do conhecimento. Escartín e Goulet (2021) demonstram que a autopós-edição surge precisamente em cenários nos quais pesquisadores desejam publicar em inglês, mas redigem seus textos inicialmente em sua língua materna e recorrem à tradução automática como apoio à escrita (p. 89). Nessa configuração, o objetivo não é substituir o tradutor profissional em todos os contextos possíveis, mas viabilizar a produção e a circulação de textos acadêmicos em inglês por autores que dominam o campo científico em questão, ainda que não possuam fluência plena na língua de publicação.

Um dos principais argumentos a favor da autopós-edição na escrita acadêmica reside no conhecimento de domínio. O estudo empírico de Parra Escartín e Goulet (2021) demonstra que autores não tradutores, ao atuarem como autopós-editores, são particularmente eficazes na identificação e correção de problemas terminológicos e semânticos no texto traduzido automaticamente, justamente em função de seu envolvimento direto com a pesquisa e de seu conhecimento especializado da área científica tratada (p. 103). As autoras observam que a maior parte das correções terminológicas é realizada pelos próprios pesquisadores, enquanto a revisão profissional tende a concentrar-se sobretudo em aspectos de fluidez e estilo (p. 104).

Esses achados indicam que, no contexto acadêmico, a adequação de um texto não depende exclusivamente de competências tradutórias formais, mas também da precisão conceitual e da consistência terminológica. Essa constatação permite problematizar a comparação entre autopós-edição e tradução profissional de forma mais cuidadosa. Tradutores profissionais, embora altamente qualificados do ponto de vista linguístico e metodológico, nem sempre possuem familiaridade profunda com o campo científico específico do texto traduzido.

Como apontam Parra Escartín e Goulet, revisores e tradutores externos podem não possuir conhecimento de domínio suficiente e precisam recorrer ao autor para esclarecer questões conceituais (2021, p. 91). Na autopós-edição, essa distância entre produção do conhecimento e mediação linguística é reduzida, uma vez que o próprio pesquisador valida as escolhas terminológicas, semânticas e argumentativas que estruturam o texto.

Além disso, a autopós-edição não se limita à correção linguística pontual. O'Brien, Simard e Goulet (2018) observam que, ao revisar a tradução automática, os autores frequentemente realizam reformulações discursivas, ajustes argumentativos e reorganizações textuais, integrando a revisão ao próprio processo de construção do texto acadêmico (p. 247). Esse movimento indica que a autopós-edição funciona como um processo de refinamento textual, no qual conteúdo, forma e intenção comunicativa são negociados de maneira integrada.

Nesse processo, a autopós-edição envolve diferentes níveis de revisão. O autor revisa aspectos linguísticos, como gramática e sintaxe, mas também aspectos discursivos, retóricos e conceituais, apoiando-se tanto no texto original quanto em seu conhecimento da área. O'Brien, Simard e Goulet (2018) demonstram que esse processo não impacta negativamente a qualidade final do texto e, em alguns casos, resulta em menor número de erros normalizados quando comparado à escrita direta em inglês como língua estrangeira (p. 252).

Dessa forma, no contexto da escrita acadêmica, a combinação entre tradução automática e autopós-edição pode ser compreendida como uma prática híbrida, situada entre escrita e tradução. Trata-se de uma estratégia que não pressupõe fluência plena em inglês, mas que se apoia no conhecimento especializado do autor para garantir precisão conceitual, consistência terminológica e adequação argumentativa. Ao permitir que o pesquisador atue como agente central do processo tradutório, a autopós-edição se configura como uma alternativa viável e funcional para a circulação internacional do conhecimento científico, especialmente em contextos nos quais o acesso a serviços profissionais de tradução é limitado ou inviável.

2.4 Gêneros e a escrita acadêmica

Um texto pode ser definido como uma instância concreta da linguagem em uso, seja falada ou escrita, que manifesta as propriedades comunicativas dos gêneros aos quais pertence. Como parte do comportamento linguístico natural, revela as escolhas e estratégias discursivas adotadas em uma situação específica, considerando o contexto, o interlocutor e o propósito. Já o conceito de gênero tem evoluído de forma significativa durante os anos. Inicialmente entendido como um grupo de textos que compartilhavam formas, estilos e conteúdo, se adaptou para compreender as mudanças nos meios de comunicação.

Para a linguística aplicada, gênero é usado para se referir a diferentes tipos de eventos comunicativos – escritos ou falados – associados a aspectos e contextos específicos. O contexto, todavia, não é algo que existe separado do texto e, conseqüentemente, do gênero. Enquanto Marcuschi (2002) defende que os gêneros estão em constante transformação já que são moldados pela atividade comunicativa, Flowerdew (2015) afirma que os gêneros são, tipicamente, utilizados por um seleto grupo de pessoas. Para ele, gênero está relacionado com poder uma vez que concede acesso a, por exemplo, informação e educação.

De acordo com Madianou e Miller (2012), a sociedade vive o chamado “ambiente polimídia”, que possibilita a escolha da ferramenta mais adequada para o propósito comunicativo. Isto inclui tecnologias mais antigas e atuais, faladas ou escritas, síncronas ou assíncronas, criando um ambiente multimodal de comunicação.

Essa multimodalidade impacta diretamente a forma como os gêneros se manifestam e se adaptam. No contexto da TA, entender e reconhecer os gêneros – antigos e novos – é fundamental para garantir uma adequação comunicativa e funcional das traduções. É nesse cenário que se insere a escrita acadêmica, uma prática discursiva fortemente regulada por convenções de gênero e por expectativas específicas da comunidade científica.

Escrever para fins acadêmicos, como argumentado anteriormente, não é tarefa corriqueira. Modelos mais antigos como o CARS se apresentam como ajuda, mas não são suficientes e, para escrever sobre um tema, primeiro é importante entender o gênero sobre o qual se apoia. O primeiro contato que o leitor tem com o texto, depois do título, é o resumo – gênero que será tratado no presente trabalho. Hyland (2004) apresenta o resumo de pesquisa como fator fundamental do sistema acadêmico, pois são eles que introduzem, de forma sucinta, o trabalho desenvolvido.

É através do resumo que o autor não só sintetiza o conteúdo do artigo, mas também desperta interesse e mostra relevância do estudo. Para capturar a atenção do leitor, é necessário construir um equilíbrio de clareza de informação, demonstrar domínio do tema ao usar terminologias de forma correta e referências que demonstrem domínio das informações retratadas. Todavia, nem todo resumo é igual.

Swales e Feak (2009) abordam as transformações que os resumos de artigo e de conferência são gêneros distintos que enfrentam desde sua concepção até sua versão final e como suas características textuais variam, mas seguem protocolos rigorosos. O resumo de artigo

(RA) se faz de extrema importância devido ao aumento exponencial de publicações científicas e por exercer o papel de apresentar o artigo completo para o leitor. O RA ideal é dividido em cinco partes e cada uma possui um movimento: i. introdução; ii. objetivo; iii. método; iv. resultados e v. conclusão.

Enquanto o RA funciona como uma apresentação de um trabalho maior, resumo de conferência (RC) é um texto independente elaborado com o intuito de colocar seu autor em conferências e promover seu trabalho (2009, p. 43) e, em geral, não possui uma estrutura padrão e se organiza em parágrafos contínuos. Todavia, estudos recentes sugerem uma organização de seis partes, mesmo que alguns não apresentem todas as alternativas e onde o sexto movimento aparece como forma de reforço (Swales; Feak, 2009, p. 45): i. descrição do tópico de pesquisa; ii. justificativa; iii. metodologia; iv. resumo das principais descobertas; v. destaque dos resultados; e vi. observações adicionais. A seguir, encontra-se uma comparação entre RA e RC.

Figura 4. Comparação entre RA e RC.

	RA	RC
Movimento 1	Introdução: fornecer informações gerais	Descrição do tópico de pesquisa: estabelecer a relevância do tema
Movimento 2	Objetivo: explicar a finalidade do estudo	Justificativa: explicar a escolha do tema
Movimento 3	Método: descrever a metodologia aplicada	Metodologia: descrever a metodologia aplicada
Movimento 4	Resultados: apresentar as descobertas	Resumo das principais descobertas: apresentar as descobertas
Movimento 5	Conclusão: destacar a relevância dos resultados e possíveis aplicações	Destaque dos resultados: destacar a relevância dos resultados e possíveis aplicações
Movimento 6	Não se aplica	Observações adicionais: abordar implicações teóricas ou práticas, limitações e possíveis caminhos futuros

Elaborar um resumo não é apenas uma questão de técnica, mas sim uma prática que reflete a dinâmica da comunicação científica. Em um mundo onde a produção de conhecimento se expande rapidamente, a habilidade de condensar informações relevantes em uma síntese clara e envolvente se torna crucial. O resumo não apenas antecipa o conteúdo do trabalho, mas também atua como uma ponte entre o autor e o leitor, algo que facilita o acesso à pesquisa e promove a disseminação do saber.

Ao compreender a importância do resumo, os pesquisadores não apenas aprimoram sua capacidade de comunicação, mas também contribuem para a construção de um ambiente acadêmico mais colaborativo e acessível.

2.5 O modelo CARS e a análise de gêneros acadêmicos

Em seu livro *Genre Analysis: English in Academic and Research Settings*, John Swales apresenta seu conceito de gênero, de forma geral, como um conjunto de eventos comunicativos que compartilham algumas características tanto formais quanto informais e aponta três fatores para analisar gêneros: estrutura textual, contexto social e propósito comunicativo. Entender que os gêneros não são estruturas rígidas e que, sim, evoluem com o tempo, levou John Swales a criar um modelo que permitisse entender como textos acadêmicos e científicos se estruturam, o Modelo CARS (*Create A Research Space*).

O modelo CARS é composto por três movimentos – que Swales nomeia movimentos retóricos e os elucida como partes estruturais que apresentam funções distintas dentro de um gênero e colaboram para o objetivo geral do texto. Cada movimento retórico possui um ou mais passos retóricos. São eles:

1. Movimento 1 - Estabelecendo um território: momento em que o autor apresenta o contexto e a relevância de seu trabalho. Inclui os seguintes passos:

1.1 - passo 1 - reivindicar centralidade: declaração do tópico de interesse e descrição do escopo, e/ou;

1.2 - passo 2 - fazer generalizações: alegações gerais sobre o tema, e/ou;

1.3 - passo 3 - revisar itens de pesquisas passadas: revisão de pesquisas anteriores, trazendo proposições relevantes e delineando o que já se sabe sobre o tema.

2. Movimento 2 - Estabelecendo um nicho: momento em que o autor demonstra um ponto a ser explorado e inclui um dos quatro passos:

2.1a - passo 1A - contrapor ideias/resultados;

2.1b - passo 1B - indicar uma lacuna;

2.1c - passo 1C - levantar uma questão;

2.1d - passo 1D - continuar uma tradição;

3. Movimento 3 - Ocupando um nicho: é momento em que o autor ocupa um lugar no tópico abordado e pode conter os seguintes passos:

3.1a passo 1A - delinear os propósitos: mostrar o que se espera conseguir com o estudo, ou;

3.1b passo 1B - anunciar a pesquisa: apresentar os objetivos e hipóteses da pesquisa;

3.2 passo 2 - anunciar as principais descobertas;

3.3 passo 3 - indicar a estrutura do artigo.

Os passos não são obrigatórios e não precisam, necessariamente, seguir a ordem aqui apresentada, mas segui-los é uma forma de garantir clareza de informações. Na figura 5, encontra-se um resumo dos movimentos e seus respectivos passos.

Figura 5. O modelo seminal de Swales

Movimento 1 - Estabelecendo um território	
Passo 1	reivindicar centralidade e/ou
Passo 2	fazer generalizações e/ou
Passo 3	revisar itens de pesquisas passadas
Movimento 2 - Estabelecendo um nicho	
Passo 1A	contrapor ideias/resultados ou
Passo 1B	indicar uma lacuna ou
Passo 1C	levantar uma questão ou
Passo 1D	continuar uma tradição
Movimento 3 - Ocupando um nicho	
Passo 1A	delinear propósitos ou
Passo 1B	anunciar pesquisa
Passo 2	anunciar principais descobertas
Passo 3	indicar estrutura do artigo

Fonte: elaboração própria.

No artigo *Applicability and variation of Swales' CARS Model to Applied Linguistics Article Abstracts*, Sanchez (2018) desenvolveu uma investigação empírica voltada à análise da organização retórica de resumos de artigos da área de Linguística Aplicada. Para isso, aplicou o modelo CARS a um *corpus* de resumos acadêmicos, com o objetivo de examinar como os movimentos e passos originalmente propostos apareciam nesses textos e até que ponto o modelo descrevia adequadamente as estruturas observadas.

A análise mostrou que, embora os movimentos do modelo CARS estivessem presentes nos resumos, eles não se limitavam à construção do espaço de pesquisa. Os dados indicaram que os resumos traziam, de forma recorrente, informações sobre os procedimentos metodológicos adotados, a apresentação dos principais resultados e comentários finais sobre a pesquisa, como conclusões, implicações ou possíveis aplicações do estudo. Esses elementos apareceram de maneira sistemática no *corpus* analisado e tiveram papel central na organização do resumo.

Ao analisar esses textos com base no modelo CARS, Sanchez (2018) observou que esse conjunto de informações acabava sendo tratado dentro do Movimento 3, responsável por ocupar o nicho de pesquisa. Nesse movimento, passavam a ser incluídas informações de natureza diferente, como a apresentação dos objetivos do estudo, a descrição da metodologia, o relato dos resultados e a indicação das contribuições da pesquisa. Para o autor, essa forma de organização dificultava a separação entre essas diferentes funções e comprometia a precisão da descrição oferecida pelo modelo original.

A análise também mostrou que esse padrão se repetia ao longo do *corpus*, indicando que não se tratava de casos isolados. Diante disso, Sanchez concluiu que o modelo CARS, em sua formulação original, não era suficiente para descrever de forma precisa a organização dos resumos acadêmicos da área de Linguística Aplicada, o que o levou a propor sua ampliação com base nos resultados obtidos. Assim, Sanchez sugeriu a inclusão de mais três movimentos:

1. Movimento 4 - Descrevendo a metodologia: inclusão de uma exposição metodológica da pesquisa.
2. Movimento 5 - Resumindo os resultados: apresentação de um breve anúncio dos resultados principais.
3. Movimento 6 - Discutindo a pesquisa: composto por três possíveis passos.
 - 3.1 – Tirar conclusões: que apresenta um breve apanhado sobre as descobertas, e/ou;

3.2 – Indicar implicações: onde o autor conecta o estudo com o contexto de forma ampla, e/ou;

3.3 – Indicar aplicações: autor menciona as aplicações práticas da pesquisa.

Figura 6. Modelo de Sanchez.

Movimento 1- Estabelecendo um território	
Passo 1	reivindicar centralidade e/ou
Passo 2	fazer generalizações e/ou
Passo 3	revisar itens de pesquisas passadas
Movimento 2 - Estabelecendo um nicho	
Passo 1A	contrapor ideias/resultados ou
Passo 1B	indicar uma lacuna ou
Passo 1C	levantar uma questão ou
Passo 1D	continuar uma tradição
Movimento 3 - Ocupando um nicho	
Passo 1A	delinear propósitos ou
Passo 1B	anunciar pesquisa
Passo 2	anunciar principais descobertas
Passo 3	indicar estrutura de RA
Movimento 4 - Descrevendo a metodologia	
Movimento 5 - Resumindo os resultados	
Movimento 6 - Discutindo a pesquisa	
Passo 1	tirar conclusões
Passo 2	indicar implicações
Passo 3	indicar aplicações

Fonte: elaboração própria.

Sanchez (2018) mostrou que o modelo CARS, em sua formulação original, não descrevia de forma suficientemente precisa a organização retórica de resumos acadêmicos da área de Linguística Aplicada, o que levou à proposição de novos movimentos com base nos padrões observados no *corpus* analisado. Em outra frente de investigação, ainda que motivada por questões distintas, estudos posteriores passaram a discutir não apenas os limites descritivos do modelo, mas também sua utilização em análises conduzidas com apoio de procedimentos automáticos.

É nesse contexto que se insere a pesquisa desenvolvida por Britto et al. (2021). O estudo partiu do crescimento expressivo da produção científica e do impacto desse aumento sobre os processos de avaliação e revisão de artigos acadêmicos, em um cenário no qual o número de revisores humanos não acompanhou a expansão das publicações. Diante dessa realidade, os autores voltaram-se à possibilidade de empregar métodos automáticos como apoio à análise da organização retórica dos textos científicos.

Para isso, Britto et al. (2021) recorreram ao modelo CARS como base para a análise retórica, mas o adaptaram de modo a torná-lo compatível com procedimentos de classificação automática. A proposta envolveu o uso de técnicas de aprendizado de máquina, em especial o *Support Vector Machine* (SVM), aplicado à identificação de movimentos e passos retóricos no nível da sentença. Esse tipo de abordagem exigiu que as categorias do modelo fossem definidas de forma mais clara e sistemática, de modo que pudessem ser reconhecidas por algoritmos de classificação.

A reformulação do modelo concentrou-se, assim, no detalhamento dos passos associados a cada movimento, com o objetivo de reduzir a sobreposição entre funções retóricas distintas. No modelo original, diferentes tipos de informação tendem a ser agrupados em categorias mais amplas, o que não representa um problema em análises qualitativas, mas se torna um obstáculo quando a análise depende da distinção entre funções específicas no nível da sentença. Ao explicitar essas funções, a versão proposta por Britto et al. (2021) permitiu uma identificação mais precisa dos movimentos e passos, condição necessária para o uso de métodos automáticos.

Além disso, o modelo reformulado preservou a organização geral do CARS, mantendo seus movimentos centrais, o que garantiu continuidade com a tradição da análise de gêneros acadêmicos. Ao mesmo tempo, o maior detalhamento dos passos tornou o modelo mais

adequado a estudos que envolvem automação, sem exigir a criação de categorias novas ou alheias ao referencial teórico original. A partir dessa reformulação, o modelo passa a contemplar os seguintes movimentos e passos:

1. Movimento 1 - Estabelecendo o território: introdução do contexto e relevância do tema e é composto por quatro passos:

1.1 - passo 1 - estabelecer a importância do tópico para a disciplina;

1.2 - passo 2 - estabelecer a importância do tópico para o mundo ou a sociedade;

1.3 - passo 3 - estabelecer a importância do tópico como um problema a ser abordado;

1.4 - passo 4 - referir-se aos trabalhos passados para estabelecer o que já se sabe.

2. Movimento 2 - Estabelecendo um nicho: demonstração da necessidade de mais pesquisa e identificação do que falta e/ou pode ser aprimorado e é feito em dois passos:

2.1 - passo 1 - indicar e destacar inadequações, fraquezas, controvérsias e resultados negativos do campo de estudo;

2.2 - passo 2 - identificar uma lacuna no conhecimento, uma falta ou escassez de uma pesquisa anterior.

3. Movimento 3 - Ocupando um nicho: explicação de como preencher a lacuna identificada no movimento anterior enquanto apresenta a contribuição do estudo e é realizado em sete passos:

3.1 - passo 1 - destacar o foco, objetivo, propósito ou argumento da pesquisa atual;

3.2 - passo 2 - definir as questões ou hipóteses levantadas pela pesquisa;

3.3 - passo 3 - descrever o *design* e métodos das pesquisas;

3.4 - passo 4 - explicar a relevância ou motivos pelo interesse na pesquisa;

3.5 - passo 5 - descrever as limitações da pesquisa;

3.6 - passo 6 - delinear a estrutura de um capítulo, artigo, dissertação ou tese;

3.7 - passo 7 - explicar palavras-chave, fornecendo definições para os conceitos utilizados.

Figura 7. Modelo expandido de Britto et al.

Movimento 1 - Estabelecendo o território	
Passo 1	estabelecer a importância do tópico para a disciplina
Passo 2	estabelecer a importância do tópico para o mundo ou a sociedade
Passo 3	estabelecer a importância do tópico como um problema a ser abordado
Passo 4	referir-se aos trabalhos passados para estabelecer o que já se sabe
Movimento 2 - Estabelecendo um nicho	
Passo 1	identificar e destacar inadequações, fraquezas, controvérsias e resultados negativos do campo de estudo
Passo 2	identificar uma lacuna no conhecimento, uma falta ou escassez de uma pesquisa anterior
Movimento 3 - Ocupando um nicho	
Passo 1	destacar o foco, objetivo, propósito ou argumento da pesquisa atual
Passo 2	definir as questões ou hipóteses levantadas pela pesquisa
Passo 3	descrever o design e métodos da pesquisa
Passo 4	explicar a relevância ou motivos pelo interesse na pesquisa
Passo 5	descrever as limitações da pesquisa
Passo 6	delinear a estrutura de um capítulo, artigo, dissertação ou tese
Passo 7	explicar palavras-chave

Fonte: elaboração própria.

No âmbito desta pesquisa, essa proposta mostrou-se adequada às escolhas analíticas realizadas. O *corpus* analisado é reduzido e a análise foi conduzida no nível da sentença, o que torna particularmente importante a utilização de categorias bem definidas. O detalhamento dos passos facilitou a identificação das funções retóricas nos resumos e permitiu a aplicação de recursos de automatização no processo de anotação, por meio do INCEPTION, sem a necessidade de ajustes adicionais no modelo.

Dessa forma, para a presente pesquisa, a adoção da versão do modelo CARS proposta por Britto et al. (2021) decorreu não apenas de sua relevância teórica, mas de sua adequação às condições concretas desta pesquisa, possibilitando uma análise mais controlada da organização retórica dos resumos acadêmicos e uma integração coerente entre o modelo adotado e os procedimentos metodológicos empregados.

2.6 Linguística de *Corpus*

Considerando que a presente investigação se baseia na análise de conjuntos de textos, a Linguística de *Corpus* (LC) assume, neste trabalho, o papel de suporte para a análise e a descrição da linguagem em uso a partir de dados textuais organizados em *corpora*. Nessa abordagem, o *corpus* corresponde ao ponto de partida da investigação, sendo tratado como uma coletânea de textos selecionados e organizados de acordo com critérios definidos pelo pesquisador, em consonância com os objetivos do estudo (Bonelli, 2010).

A construção de um *corpus* envolve decisões relacionadas à seleção, organização e delimitação dos textos que o compõem. Esses critérios incluem aspectos como língua, gênero textual, domínio discursivo e contexto de produção, sendo fundamentais para assegurar a comparabilidade dos dados analisados. As escolhas realizadas nessa etapa incidem diretamente sobre os padrões que podem ser observados nos dados e, conseqüentemente, sobre as interpretações que deles decorrem.

A Linguística de *Corpus* é compreendida, neste trabalho, como uma abordagem amplamente utilizada para a observação empírica da linguagem em uso, podendo ser articulada a diferentes quadros teóricos no âmbito da Linguística Aplicada, incluindo os Estudos da Tradução. Seu desenvolvimento relaciona-se aos avanços tecnológicos que viabilizaram o armazenamento e o processamento de grandes quantidades de dados textuais, ampliando as possibilidades de análise sistemática da linguagem (McCarthy; O’Keeffe, 2010).

No âmbito da LC, os *corpora* podem ser organizados de acordo com a relação estabelecida entre os textos que os constituem. Entre os tipos mais recorrentes, destacam-se os *corpora* paralelos, multilíngues e comparáveis, cuja distinção assume relevância particular em pesquisas que envolvem tradução.

Os *corpora* paralelos são compostos por textos produzidos em uma língua e suas respectivas traduções para outra língua. Esses textos são organizados de forma alinhada, geralmente por palavras, sentenças ou segmentos, estabelecendo uma relação direta a partir do sentido que compartilham. Essa configuração permite a observação de padrões de equivalência, escolhas lexicais e marcas linguísticas recorrentes no processo tradutório.

Os *corpora* multilíngues reúnem textos originalmente produzidos em duas ou mais línguas, sem que haja uma relação direta de tradução entre eles. Esses textos são organizados segundo critérios de comparabilidade, como gênero textual ou domínio discursivo, possibilitando observar como funções comunicativas semelhantes são realizadas em diferentes línguas.

Já os *corpora* comparáveis reúnem textos escritos originalmente em uma língua e textos traduzidos para essa mesma língua, sem correspondência direta entre documentos específicos. Essa configuração possibilita comparar aspectos como estilo, vocabulário e organização textual de produções originais e traduzidas, desde que os critérios de coleta assegurem a comparabilidade em termos de gênero textual e função comunicativa.

No campo dos Estudos da Tradução, os *corpora* comparáveis têm se mostrado particularmente produtivos, pois permitem investigar padrões associados à linguagem traduzida sem recorrer ao alinhamento direto entre textos. Em pesquisas que envolvem tradução automática e autopós-edição, esse tipo de *corpus* possibilita examinar em que medida textos traduzidos automaticamente e posteriormente revisados se aproximam ou se distanciam de textos produzidos diretamente na língua de chegada.

Os avanços tecnológicos associados à Linguística de *Corpus* também contribuíram para a ampliação dos tipos de dados que podem compor um *corpus*, incluindo, de forma ampla, materiais de natureza falada e multimodal. Esse desenvolvimento reforçou a LC como uma abordagem recorrente na Linguística Aplicada, ao oferecer meios para a análise empírica da linguagem em diferentes contextos de uso.

No presente trabalho, a Linguística de *Corpus* fornece o suporte metodológico para a análise de resumos acadêmicos produzidos em diferentes modos de produção textual. A adoção de um *corpus* comparável possibilita observar padrões retóricos em textos escritos em inglês como língua adicional e em textos traduzidos automaticamente e autopós-editados nessa mesma

língua, contribuindo para a descrição dessas produções como práticas reais de escrita acadêmica mediadas pela tradução automática.

3 METODOLOGIA

3.1 Metodologia de coleta

A presente pesquisa utiliza um *corpus* composto por textos produzidos no âmbito do projeto de Nunes (2021), em sessões de coleta realizadas no Laboratório de Tradução (LETRA) da Faculdade de Letras da UFMG. Entre junho de 2022 e abril de 2024, participantes foram convidados ao laboratório para executar tarefas de escrita e de autopós-edição em um ambiente controlado, com instruções padronizadas e monitoramento das condições de produção textual. Todo o procedimento foi previamente aprovado pelo Comitê de Ética em Pesquisa².

Os textos foram produzidos por doze doutorandos vinculados a programas de pós-graduação da UFMG e distribuídos em três grandes áreas conforme classificação do CNPq³: Ciências Sociais Aplicadas (CSA), Ciências Exatas e da Terra (CET) e Ciências Humanas (CH).

Para compor a amostra, os sujeitos precisaram atender aos seguintes critérios:

1. ser brasileiro (a);
2. estar, ao menos, no segundo ano do doutorado;
3. desenvolver pesquisa em uma das três áreas;
4. ter, como L1, o português brasileiro;
5. leitura e escrita em inglês (L2) em nível intermediário autodeclarado.

A exigência da realização de um teste para verificar o nível de proficiência do inglês foi considerada. Todavia, se mostrou inviável devido à dificuldade logística no recrutamento dos participantes (Guimarães, 2025 p.45).

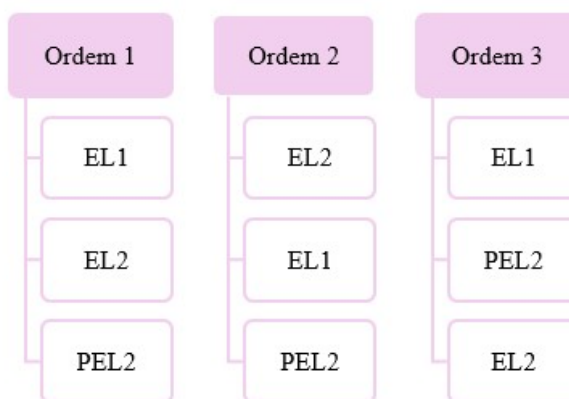
² Número do Certificado de Apresentação de Apreciação Ética (CAAE): 40708520.1.0000.5149

³ Link para a classificação: <https://lattes.cnpq.br/documents/11871/24930/TabeladeAreasdoConhecimento.pdf>

No que diz respeito às tarefas, cada participante realizou três: i. produção de um resumo em português brasileiro (EL1), ii. escrita de um resumo em inglês (EL2), iii. a autopós-edição em inglês (PEL2) da tradução automática do resumo escrito em L1. O tradutor automático utilizado para gerar os textos foi o DeepL⁴ e a autopós-edição foi executada na ferramenta Translog-II, um programa desenvolvido para investigar o processo de tradução e seus vários aspectos.

De acordo com Jakobsen, “o Translog fornece ao pesquisador e ao educador uma ferramenta para pesquisar e trabalhar com o processo de tradução⁵” (2006, p.105). Ainda, durante a tarefa de autopós-edição, os participantes puderam consultar dicionários e lançar mão de mecanismos de busca, mas sem a utilização de sistemas de tradução automática. A sequência dessas tarefas foi aleatorizada (Figura 8) de modo a reduzir efeitos de familiarização, permitir a observação de possíveis influências na ordem de realização das tarefas e preservar ao máximo a validade ecológica dos dados.

Figura 8. Ordens possíveis de realização das tarefas de produção textual



Fonte: adaptado de Guimarães (2025).

⁴ Endereço eletrônico da iniciativa: <https://www.deepl.com/pt-BR/translator>

⁵ Traduzido pela autora. Trecho original: Translog provides the research and the educator a tool to work with and research the process of translation.

Os participantes produziram resumos acadêmicos para dois congressos: i. um nacional e ii. um internacional, seguindo as seguintes normas (Nunes, 2021, p. 10):

- a) a língua do resumo (PT-BR ou EN), conforme o evento;
- b) texto entre 100 e 250 palavras;
- c) ausência de figuras, tabelas, gráficos ou palavras-chave;
- d) organização do texto em objetivos, métodos, resultados e possíveis contribuições para a área.

3.2 O desenho do *corpus*

O *corpus* analisado neste estudo é do tipo comparável e monolíngue, sendo composto por trinta e seis textos produzidos no âmbito da pesquisa de Nunes (2021), conforme descrito anteriormente. Para cumprir os objetivos da pesquisa, foram escrutinados apenas os resumos originalmente escritos em inglês (EL2), os textos de saída da máquina, ou seja, traduzidos automaticamente (TA) para o inglês e as autopós-edições em inglês (PEL2), conforme a tabela 1.

Tabela 1. Divisão geral dos resumos por tarefa e área de conhecimento.

Área	EL2	TA	PL2	Total
CET	4	4	4	12
CH	5	5	5	15
CSA	3	3	3	9
Textos	12	12	12	36

Fonte: adaptado de Guimarães (2025).

É importante ressaltar que os conteúdos de alguns resumos escritos diretamente em inglês L2 não corresponderam aos dos traduzidos automaticamente (TA) e, conseqüentemente, aos textos autopós-editados (PEL2). Isso se deve ao fato de que alguns textos originais em português-brasileiro (EL1) versavam sobre tópicos e assuntos diversos. Contudo, apesar dessa discrepância, todos os textos foram considerados para análise por serem comparáveis, já que são do mesmo gênero (resumo de conferência), foram produzidos sob as mesmas condições experimentais e seguiram os mesmos critérios de compilação. Ainda, foram analisados separadamente.

Tabela 2. Distribuição das correspondências entre os assuntos versados em cada *subcorpus* por área do conhecimento.

	EL2 = TA/PEL2	EL2 ≠ TA/PEL2
CET	4	0
CH	3	2
CSA	2	1

Fonte: elaboração própria.

3.3 Metodologia de análise do *corpus*

Após a coleta, extraiu-se os textos dos arquivos XLM gerados pelo programa Translog-II a partir do rastreador ocular Tobii T60, sendo eles inseridos em uma planilha digital⁶ que funciona como base de controle e checagem dos dados, incluindo informações sobre a área do conhecimento, o modo de produção, a ordem da tarefa e a identificação anonimizada dos participantes (Figura 9).

⁶ Aqui, usa-se a Planilha Google.

Figura 9. Planilha digital de controle - base de dados

ORDEM	CÓD.	ÁREA	TA	PE L2	EL
ORDEM 1	BPC0551	CET	The kinetic study of th	The kinetic study of the	The study of thermal de

Fonte: captura de tela realizada pela pesquisadora.

Com o intuito de permitir análises distintas e a utilização de ferramentas de processamento automático, os resumos foram armazenados em arquivos de texto simples (.txt) com codificação UTF-8. A nomeação seguiu um código específico, composto por quatro elementos: i. ordem de realização das tarefas (1, 2 ou 3); ii. sigla da área de conhecimento (CSA, CET ou CH); iii. o modo de produção do texto (escrita direta L2, traduções automáticas para L2 e autopós-edições dos textos em L2) e; iv. um código alfanumérico que identifica o participante. Dessa forma, todos os arquivos foram rotulados da seguinte forma: “1-CH-PEL2-EJT0905.”

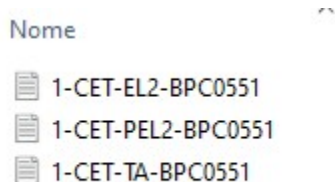
Os textos foram, inicialmente, organizados em pastas divididas por áreas, no formato .txt e com a nomeação específica, tanto das pastas divididas por áreas e com os arquivos já com os códigos de nomenclatura (Figuras 10 e 11). Em seguida, foram transferidos para um arquivo .xlxs do programa Microsoft Excel[®] de modo apresentar os dados em uma estrutura de planilha, com linhas e colunas, que facilitasse o seu manuseio e a sua visualização. A partir dessa configuração, realizou-se o alinhamento das sentenças no nível de equivalência de sentido: sentenças que expressavam o mesmo conteúdo/tópico foram agrupadas na mesma linha, enquanto, nos casos em que não se identificou um segmento correspondente em determinado *subcorpus*, a célula correspondente foi mantida em branco e em linha única (Figura 12).

Figura 10. Organização das pastas



Fonte: captura de tela realizada pela pesquisadora.

Figura 11. Organização dos arquivos com seus códigos de nomenclatura.



Fonte: captura de tela realizada pela pesquisadora.

Figura 12. Alinhamento dos textos.

	A	B	C
1	EL2	TA	PEL2
2		The kinetic study of the thermal decomposition of	The kinetic study of the thermal decomposition of drugs
3	The study of thermal decomposition of		
4	This treatment isn't right because the	However, the traditional practice of this method	However, the traditional practice of this method
5	Normally, this studies are made with	Traditionally, the data is translated empirically,	Traditionally, the data is translated empirically, through
6		However, it is possible to perform such studies	However, it is possible to perform such studies through
7	The objective of this work is construct a	The objective of this work is to build artificial neural	The objective of this work is to build artificial neural
8	In this point, the networks who are made,	The networks architected so far have produced	The networks architected so far have produced results
9	This work is very useful in drug industry	The present study has direct contributions in the	The present study has direct contributions in the

Fonte: captura de tela realizada pela pesquisadora.

A próxima etapa envolveu a anotação de cada segmento quanto ao movimento e ao passo retórico com base no modelo CARS expandido que fundamenta a análise. Essa anotação foi registrada diretamente na planilha, por meio de colunas específicas, permitindo visualizar,

para cada linha, tanto o conteúdo textual quanto a função retórica da sentença no contexto do resumo acadêmico.

Após a primeira classificação feita pela pesquisadora, os arquivos em formato .xlxs (planilha do Microsoft Excel) foram transferidos para o Google Drive, mais especificamente, o Google Planilhas. Isso aconteceu para facilitar a validação dos dados (Figura 13), posteriormente realizada por um linguista. O último processo de classificação foi realizado, em parte, automaticamente pelo programa escolhido INCEpTION.

Figura 13. *Corpus* classificado e validado no Google drive.

	A	B	C	D	E	F
1	EL2	M/P	TA	M/P	PEL2	M/P
2			The kinetic study of the thermal decomposition of drugs and polymers is routinely done through the use of linear equations that aim to treat experimental data obtained through thermal analyzers.	M1/P1	The kinetic study of the thermal decomposition of drugs and polymers is routinely done through the use of linear equations that aim to treat experimental data obtained through thermal analyzers.	M1/P1
3	The study of thermal decomposition of drugs and polymers is very useful to predict they time of life in several conditions.	M1/P1				
4	This treatment isn't right because the errors associates are large them the results.	M2/P1	However, the traditional practice of this method produces errors of large orders of magnitude, making the results inaccurate.	M2/P1	However, the traditional practice of this method produces errors of large orders of magnitude, making the results inaccurate.	M2/P1
5	Normally, this studies are made with matematics models using of singularity and empirical form.	M2/P1	Traditionally, the data is translated empirically, through trial and error, and thermal decomposition models whose equations describe the experimental data most accurately are accepted as correct.	M2/P1	Traditionally, the data is translated empirically, through trial and error, and thermal decomposition models whose equations describe the experimental data most accurately are accepted as correct.	M2/P1
6			However, it is possible to perform such studies through new kinetic models, based on the traditional ones, and one way to design these models is through artificial neural networks.	M2/P2	However, it is possible to perform such studies through new kinetic models, based on the traditional ones, and one way to design these models is through artificial neural networks.	M2/P2
7	The objective of this work is construct a neural network capable of use differents mathematical models to predict the thermal decomposition of drugs and polymers.	M3/P1	The objective of this work is to build artificial neural networks in MATLAB language that produce accurate thermal decomposition kinetic models with negligible order of magnitude (smaller than 10 ⁻⁴) residual errors, for thermal and kinetic analysis of the thermal decomposition of drugs and polymers.	M3/P1	The objective of this work is to build artificial neural networks in MATLAB language that produce accurate thermal decomposition kinetic models with residual errors with acceptable order of magnitude (smaller than 10 ⁻⁴), for thermal and kinetic analysis of the thermal decomposition of drugs and polymers.	M3/P1
8	In this point, the networks who are made, are capable of minimize the error in one million times.	M3/P1	The networks architected so far have produced results with error up to a million times smaller than traditional methods.	M3/P1	The networks architected so far have produced results with errors a million times smaller than traditional methods.	M3/P1
9	This work is very useful in drug industry and automobilistic industry for predict the shelf life and decomposition time of polymers use in extreme condition, respectfully.	M3/P4	The present study has direct contributions in the automotive and pharmaceutical industries, predicting with great accuracy the decomposition time of automotive polymers under extreme conditions and the shelf life of pharmaceuticals, respectively.	M3/P4	The present study has direct contributions in the automotive and pharmaceutical industries, predicting with great accuracy the decomposition time of automotive polymers under extreme conditions and the shelf life of pharmaceuticals, respectively.	M3/P4
10						

Fonte: captura de tela realizada pela pesquisadora.

3.3.1 – Classificação automática – o INCEpTION

O INCEpTION é, de acordo com seu guia de utilização⁷:

“[...] um ambiente de anotação textual texto útil para vários tipos de tarefas. As anotações são geralmente utilizadas para fins linguísticos e/ou de aprendizado de máquina. INCEpTION é uma aplicação web na qual vários usuários podem trabalhar no mesmo projeto e pode conter vários projetos de anotação ao mesmo tempo.⁸”

Ainda:

“[...] pode ser ampliado e adaptado às necessidades individuais. Em geral, oferece elementos predefinidos como bases de conhecimento, camadas e conjuntos de etiquetas, como um ponto de início, mas permite que você os modifique ou crie novos aspectos. Você pode, por exemplo, integrar uma base de conhecimento de sua escolha; criar e modificar bases predefinidas [...] construir recomendações personalizados para te sugerir anotações [...]”⁹

Todavia, a execução do INCEpTION exige que o Java esteja instalado na versão 11 ou superior. Trata-se de uma linguagem de programação e plataforma orientada ao produto que possui três componentes necessários para seu funcionamento: i. o *Java Runtime Environment* (JRE), ambiente que permite executar programas em Java; ii. o *Java Development Kit* (JDK), o conjunto de ferramentas utilizadas e; iii. *Java Virtual Machine* (JVM), o software que faz o código Java funcionar.

⁷ Disponível em: <https://inception-project.github.io/releases/38.5/docs/user-guide.html>

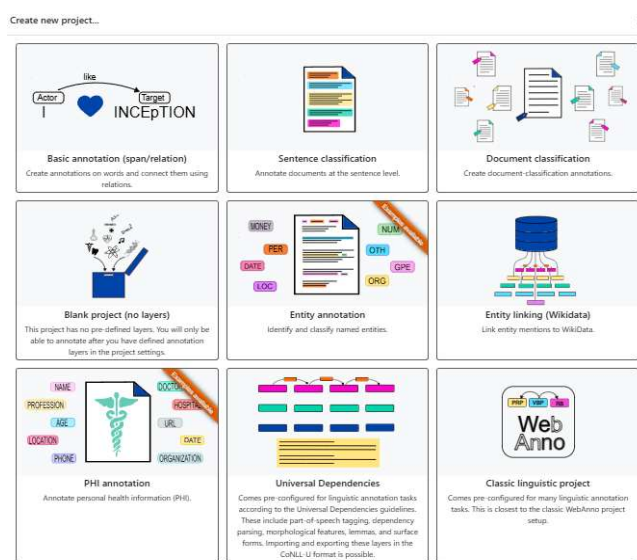
⁸ Traduzido pela pesquisadora. Trecho original: “INCEpTION is a text-annotation environment useful for various kinds of annotation tasks on written text. Annotations are usually used for linguistic and/or machine learning concerns. INCEpTION is a web application in which several users can work on the same annotation project and it can contain several annotation projects at a time. [...]”.

⁹ Traduzido pela pesquisadora. Trecho original: “In addition, INCEpTION is extendable and adaptable to individual requirements. Often, it provides predefined elements, like knowledge bases, layers and tagsets to give you a starting point but you can also modify them or create your own from scratch. You may for example integrate a knowledge base of your choice; create and modify custom knowledge bases; create and modify custom layers and tagsets to annotate your individual task; build custom so-called recommenders which automatically suggest annotations to you so you will work quicker and easier; and much more.’

De modo geral, o JDK reúne as ferramentas usadas para escrever e compilar programas Java, convertendo o código em *bytecode*¹⁰, que é então enviado ao JRE e que, com suas bibliotecas e com a JVM, interpreta o código e faz o programa funcionar em diferentes sistemas. Além do Java, o INCEption requer algumas condições específicas do sistema para sua operação como a utilização dos navegadores Chrome ou Safari¹¹.

Após atender a todos os requisitos, a instalação do programa e a criação da conta de administrador, a primeira etapa foi a criação de um projeto. O programa oferece algumas opções configuradas previamente para diferentes tipos de anotação (Figura 14) e, para o presente trabalho, a escolha foi por um Projeto em Branco (*Blank Project*) que não apresenta camadas pré-definidas.

Figura 14. Interface dos projetos disponíveis no INCEption.



Fonte: captura de tela realizada pela pesquisadora.

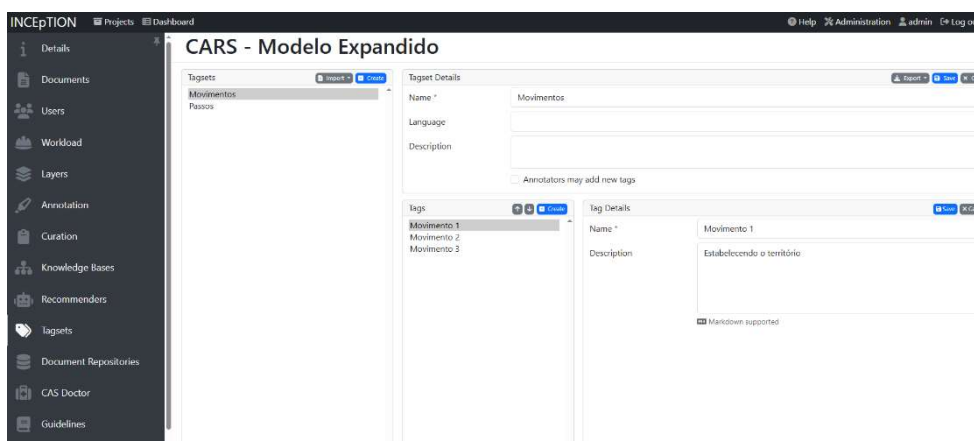
Com o novo projeto “CARS – Modelo Expandido”, deu-se início às configurações iniciais necessárias para o tipo de análise aqui proposta. A primeira etapa consistiu na criação de dois Conjuntos de Etiquetas (*tagsets*) e de suas respectivas Etiquetas (*tags*). Um Conjunto

¹⁰ *Bytecode* é uma forma intermediária de código que é gerada a partir de linguagens de programação de alto nível, como Java, C# e Python. Esse código não é diretamente executável pelo hardware do computador, mas sim por uma máquina virtual, como a Java Virtual Machine (JVM). O bytecode é projetado para ser portátil, permitindo que o mesmo código seja executado em diferentes plataformas sem a necessidade de recompilação. Fonte: <https://flammadesign.com.br/glossario/o-que-e-bytecode-entenda-sua-importancia/>

¹¹ Outros navegadores podem funcionar, todavia são testados com menos frequência pelos desenvolvedores.

de Etiqueta corresponde ao grupo de categorias utilizadas na anotação, enquanto cada Etiqueta representa uma categoria específica dentro desse grupo. Assim, foram criados dois conjuntos: Movimentos, que recebeu as etiquetas M1, M2 e M3, e Passos, que recebeu as etiquetas P1, P2, P3 etc., todas acompanhadas de suas respectivas descrições.

Figura 15. Interface dos conjuntos de etiquetas.



Fonte: captura de tela realizada pela pesquisadora.

Para que um conjunto de etiquetas funcione, é necessário que este esteja vinculado a uma Camada (*layer*), mais especificamente, a uma determinada configuração (*feature*) de uma Camada. Para tanto, criou-se a camada Classificação M/P. Como optou-se por um projeto em branco, a camada criada foi configurada manualmente, uma vez que nenhum elemento estrutural é fornecido previamente. Esta envolve três aspectos fundamentais: i. as propriedades, que definem as informações básicas; ii. os tipos, que determinam sua estrutura e forma de anotação e; iii. os comportamentos, que estabelecem as regras de funcionamento e interação durante o processo de anotação.

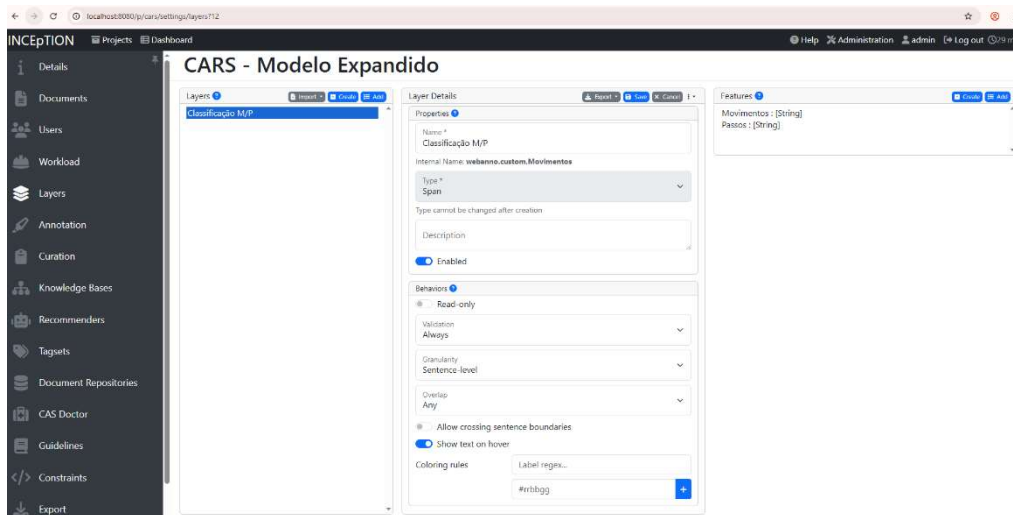
As propriedades reúnem informações essenciais para a identificação e funcionamento. Elas incluem o nome, seu tipo e uma breve descrição que orienta o anotador e aparece como dica na interface. Também é possível ativar ou desativar a camada conforme a necessidade do projeto. Os tipos definem a forma estrutural da anotação que será realizada. No INCEpTION, há três tipos principais: *span*, que marca segmentos contínuos de texto; *relation*, que estabelece conexões entre dois spans; e *chain*, que organiza spans em sequências encadeadas.

A escolha do tipo determina a lógica de funcionamento da camada e o modo como as anotações poderão ser criadas, visualizadas e interpretadas dentro do projeto. Os comportamentos estabelecem as regras de uso e interação da camada durante a anotação. Eles controlam aspectos como possibilidade de edição, exibição do texto, exigência de validação ao finalizar um documento, nível mínimo de granularidade permitido para criar sequências textuais (*spans*), permissões de sobreposição e a possibilidade de cruzar fronteiras de sentença, entre outros.

Para o presente trabalho, as configurações foram:

- 1) Nome: Classificação M/P
- 2) Tipo: *Span* (segmento contínuo de texto)
- 3) Validação: *sempre* (verifica se as anotações são válidas sempre que o documento é marcado como finalizado)
- 4) Granularidade: nível da sentença (anotações dentro do limite da sentença)
- 5) Sobreposição: sempre
- 6) Característica: *string* (texto livre)

Figura 16. Interface da Camada criada e suas propriedades.



Fonte: captura de tela realizada pela pesquisadora.

A próxima etapa envolveu a escolha dos textos com a finalidade de criar uma base de conhecimento para realizar as classificações automáticas sem criar um *overfitting*, fenômeno que acontece quando a máquina trabalha com poucos dados e o modelo tende a “decorar” em

vez de “entender”. Aprender com poucos dados de treino é difícil porque o modelo não dispõe de exemplos suficientes para identificar quais características são realmente relevantes para a tarefa. Como consequência, a ferramenta pode acabar reproduzindo detalhes superficiais ou irrelevantes, em vez de aprender padrões mais robustos, capazes de se aplicar a novos dados.

Para maximizar os resultados da classificação automática e reduzir a influência de padrões estilísticos ou temáticos, os três tipos de textos (EL2, TA e PEL2) foram definidos como domínios discursivos. Essa escolha se baseia na compreensão de que cada modo de produção constitui uma origem textual distinta, gerando características linguísticas próprias.

As áreas de conhecimento (CET, CH e CSA), por sua vez, foram classificadas como macrodomínios disciplinares, pois representam conjuntos mais amplos de práticas científicas que influenciam o estilo argumentativo, o vocabulário técnico e, possivelmente, a organização retórica dos textos. Essa distinção permite separar variações decorrentes do processo de produção (domínios discursivos) daquelas associadas ao campo científico (macrodomínios disciplinares) (Choi e Ko, 2025, p.2).

A escolha dos textos para treinamento do INCEpTION seguiu três critérios metodológicos: i textos que contemplassem o maior número possível de movimentos e passos, assegurando representatividade do gênero; ii inclusão de ao menos uma produção de cada participante, preservando diversidade de participantes; e iii manter uma distribuição equilibrada entre os macrodomínios disciplinares, evitando vieses decorrentes sobre ou sub-representação de áreas específicas.

Estas especificações resultaram em 18 resumos, divididos por todas as áreas, participantes e categorias.

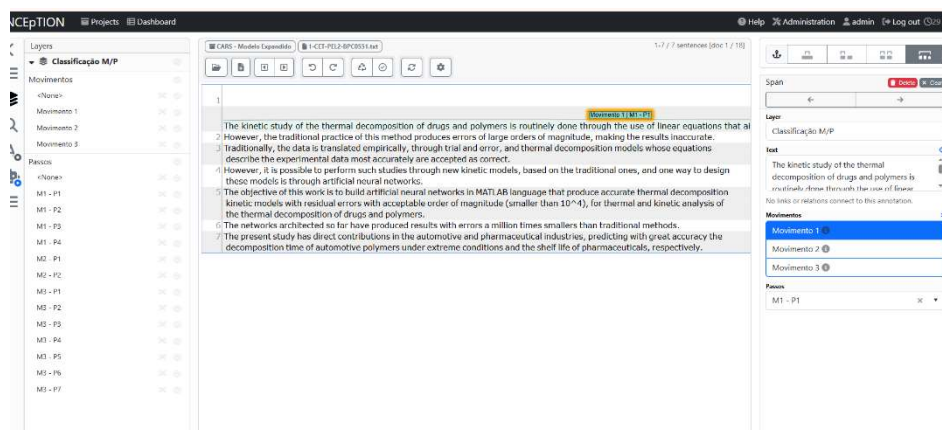
Tabela 3. Divisão dos Textos.

Área	EL2	TA	PL2	Total
CET	3	2	2	
CH	2	2	4	
CSA	1	1	1	
	6	5	7	18

Fonte: elaborado pela pesquisadora.

Após a definição dos textos, os arquivos .txt foram importados para a plataforma no formato texto simples – uma sentença por linha, e classificados manualmente.

Figura 17. Processo de classificação manual no INCEpTION.



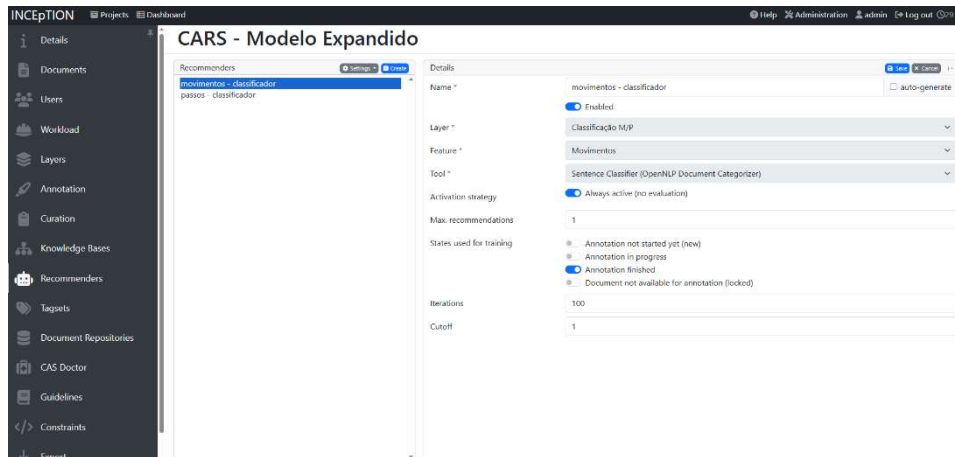
Fonte: captura de tela realizada pela pesquisadora.

Com a base de dados estabelecida e o modelo treinado, o último passo envolveu a criação e configuração do Recomendadores (*Recommenders*), ferramenta disponibilizada pelo INCEpTION, que aprende através das anotações feitas manualmente e sugere classificações automáticas para outras sentenças. Dois Recomendadores distintos foram criados para automatizar o processo de anotação.

O primeiro, nomeado movimentos – classificador, foi definido para reconhecer as sentenças associadas a esse domínio. Sua camada foi associada à de Classificação M/P, enquanto sua Característica ao de Movimentos. Como Ferramenta (*tool*), optou-se pelo Classificador de Sentença (*Sentence Classifier – OpenNLP Document Categorizer*).

O segundo, nomeado de Passos – classificador também teve sua Camada ligada à de Classificação M/P, mas sua Característica foi relacionada a de Passos. A ferramenta também foi Classificador de Sentença. A opção de Ativação de Estratégia (*Activation strategy*) manteve-se sempre ativada. Em conjunto, ambos oferecem um sistema de recomendação que equilibra precisão e abrangência, fortalecendo a consistência das anotações.

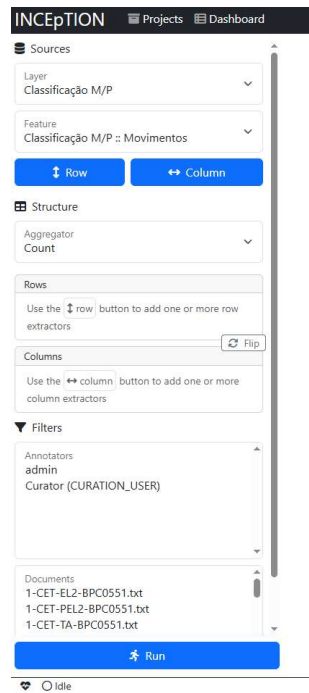
Figura 18. Interface das configurações da Recomendação.



Fonte: captura de tela realizada pela pesquisadora.

Por fim, os 18 textos restantes foram importados e as configurações de anotação alteradas (Figuras 19 e 20) para permitir a classificação automática das sentenças remanescentes pelo INCEpTION (Figura 21).

Figura 19. Interface das configurações de anotação.



Fonte: captura de tela realizada pela pesquisadora.

Figura 20. Interface das preferências de anotação

Fonte: captura de tela realizada pela pesquisadora.

Figura 21. Interface das classificações automáticas do INCEption.

Fonte: captura de tela realizada pela pesquisadora.

Com todos os passos metodológicos descritos, a próxima seção apresenta as análises e a discussão dos resultados obtidos.

4 APRESENTAÇÃO E DISCUSSÃO DOS RESULTADOS

Este capítulo apresenta os achados obtidos a partir da classificação retórica dos resumos. Convém, antes de apresentar os resultados específicos, descrever o *corpus* de forma geral, para uma contextualização mais ampla dos dados. O *corpus*, em geral, gerou trezentas e noventa e sete sentenças (397), que foram analisadas e divididas de acordo com a área, ordem de execução das tarefas e modo de produção. Essa distinção permitiu que os resultados melhor refletissem os critérios de compilação, evitando que variações, como o número de participantes, afetassem a validade dos resultados.

Tabela 4. Dados gerais coletados.

Subcorpus	Sentenças		Movimento 1		Movimento 2		Movimento 3		Total	
	absoluto	porcentagem	absoluto	porcentagem	absoluto	porcentagem	absoluto	porcentagem	absoluto	porcentagem
CET	134	33,75%	44	33%	12	9%	78	58%	134	100%
CH	165	41,56%	21	13%	14	8%	130	79%	165	100%
CSA	98	24,69%	32	33%	0	0%	66	67%	98	100%
Ordem 1	105	26,45%	6	6%	10	10%	89	85%	105	100%
Ordem 2	163	41,06%	52	32%	12	7%	99	61%	163	100%
Ordem 3	129	32,49%	38	29%	4	3%	87	67%	129	100%
EL2	85	21,41%	18	21%	8	9%	59	69%	85	100%
PEL2	92	23,17%	23	25%	7	8%	62	67%	92	100%
TA	86	21,66%	22	26%	7	8%	57	66%	86	100%
INCEpTION	134	33,75%	34	25%	4	3%	96	72%	134	100%

Fonte: elaboração própria.

Com as distribuições gerais do *corpus* apresentado, foi possível retomar os objetivos que sustentam o percurso analítico do capítulo.

A organização segue diretamente os objetivos da pesquisa: i. examinar como os movimentos retóricos presentes nos textos escritos em inglês (EL2) se aproximam ou se distanciam dos presentes nos textos que foram autopós-editados (PEL2) após tradução automática e, ii. comparar, com o mesmo foco, a distribuição dos passos retóricos. Ainda, considerando a relevância metodológica do processo de produção textual, investigou-se também os movimentos e passos presentes nos textos traduzidos automaticamente (TA) e o desempenho da classificação automática realizada pelo INCEpTION.

4.1 EL2 e PEL2 – Dados gerais

A apresentação dos resultados começa pelos dados referentes aos textos em EL2 e PEL2. A comparação entre esses dois *subcorpora* oferece uma visão inicial das escolhas feitas pelos participantes e a influência da tradução automática e da autopós-edição na organização retórica dos resumos.

Tabela 5. Divisão dos textos em EL2 e PEL2.

EL2 e PEL2	Movimento 1		Movimento 2		Movimento 3		Total de Movimentos	
	absoluto	porcentagem	absoluto	porcentagem	absoluto	porcentagem	absoluto	total
CET	21	33%	7	11%	36	56%	64	100%
CH	8	11%	8	11%	59	79%	75	100%
CSA	12	32%	0	0%	26	68%	38	100%
Ordem 1	2	4%	7	15%	39	81%	48	100%
Ordem 2	22	32%	6	9%	41	59%	69	100%
Ordem 3	17	28%	2	3%	41	68%	60	100%
EL2	18	21%	8	9%	59	69%	85	100%
PEL2	23	25%	7	8%	62	67%	92	100%

Fonte: elaboração própria.

A análise dos movimentos retóricos revelou a predominância do Movimento 3 (M3), demonstrando que os autores concentram seus esforços na apresentação de objetivos, métodos, resultados e contribuições.

Nas três áreas analisadas é possível observar que a organização retórica apresenta tanto regularidades quanto particularidades significativas. Na área de Ciências Exatas e da Terra (CET), percebe-se a presença tanto quanto expressiva do Movimento 1 (M1), que é responsável por contextualizar, apresentar o tema e justificar sua relevância. O M1, combinado com o M3, indica uma maior afinidade dos participantes à contextualização inicial e à justificativa do estudo descrito no resumo, enquanto a baixa recorrência do Movimento 2 (M2) sugere uma dificuldade em estabelecer um nicho discursivo.

Nas Ciências Humanas (CH), a dominância do M3 (79%) em relação aos M1 e M2, ambos com 11%, sugere um estilo de escrita onde o foco recai sobre a explicitação do que foi realizado, em detrimento das estratégias de contextualização mais amplas ou das elaborações de problematizações.

Já a área de Ciências Sociais Aplicadas (CSA) apresenta um padrão singular onde, embora o M1 permaneça expressivo com 32%, e o M3 continue dominante, com 68%, o M2 não aparece em nenhuma produção.

Essa ausência reforça uma transição direta entre introdução e apresentação dos estudos, sem construção explícita do nicho. Essa falta pode ser interpretada como indicador de insegurança teórica, limitação de repertório discursivo ou, ainda, como reflexo de práticas pedagógicas que priorizam a descrição do estudo em detrimento da argumentação sobre sua relevância.

As análises apontam para uma organização retórica que muda de acordo com as especificidades disciplinares. A CET favorece sequências mais extensas e estruturadas, enquanto a CH mostra um maior equilíbrio entre os movimentos. A CSA apresenta produções mais concisas e com pouca variação de movimentos. Essas características evidenciam que, embora sigam padrões gerais semelhantes, cada campo mobiliza recursos retóricos de formas distintas conforme suas práticas discursivas.

As ordens de produção acrescentam uma dimensão processual que corrobora a compreensão desses padrões. Na Ordem 1, a concentração elevada do M3 (81%), acompanhada de um M1 bastante reduzido (4%) e de um M2 ainda incipiente (15%), sugere que os participantes favorecem a transmissão da informação central apresentando os objetivos e os métodos, sem investir na construção do contexto e/ou do nicho.

Na Ordem 2, percebe-se um equilíbrio retórico mais evidente entre M1 (32%) e M3 (59%), o que demonstra que, ao retornar à tarefa, os participantes revelam uma maior consciência retórica e ampliam sua capacidade de situar o estudo em um contexto discursivo mais abrangente.

Entretanto, a Ordem 3 retoma a predominância do M3, que aparece com 68% de ocorrência, enquanto o M1 cai para 28%, e o M2 permanece residual (3%). Essa oscilação pode refletir tanto uma simplificação retórica quanto o efeito de sobreposição entre as etapas de pós-edição e escrita final em EL2, que tende a concentrar o texto no conteúdo principal.

Os dados indicam que a ordem das tarefas exerceu pouca influência na estrutura retórica dos resumos, atuando como elemento de refinamento. Isso significa que produzir textos em diferentes sequências evidencia ajustes pontuais na densidade informacional e na distribuição dos movimentos, sem alterar a lógica organizacional. Assim, independente da ordem, as tarefas de escrita e de autopós-edição convergem para estruturas semelhantes, indicando coerência entre as versões produzidas.

No que diz respeito aos modos de produção aqui analisados, EL2 e PEL2, as diferenças encontradas permitem observar o impacto da tecnologia e da revisão humana no processo de escrita. A análise dos textos produzidos diretamente em inglês como L2 (EL2) realizada a partir da identificação e quantificação dos movimentos e passos retóricos (M1, M2 e M3), com esses dados contabilizados e comparados no Excel por meio de fórmulas para o cálculo de frequências absolutas e relativas, permitiu não apenas compreender o repertório retórico empregado, mas também mapear estratégias de simplificação, omissões, preferências sintáticas e escolhas discursivas.

Em termos gerais, os resultados sugerem que os textos em EL2 apresentam relativa estabilidade organizacional, com predominância marcada do Movimento 3 (M3), uso moderado do Movimento 1 (M1) e aparição limitada do Movimento 2 (M2). Essa distribuição indica que, na escrita inicial em língua adicional, os participantes tendem a priorizar sequências discursivas diretas e expositivas, ao mesmo tempo em que evitam movimentos que exigem formulações argumentativas mais complexas.

O M1, responsável por situar o tema e introduzir o território discursivo, é encontrado de forma restrita e irregular. Com exceção da CET, que tem uma maior expressão, ele surge de forma condensada e com pouca laboração contextual. Sua baixa incidência pode vir de dois fatores i. as exigências linguísticas que esse movimento requer como uma maior densidade lexical; e ii. a característica do gênero por si só que, devido a limitação de palavras (normalmente entre 150 e 200), incentiva a escolha de informações essenciais, o que reduz a extensão de contextualização.

O M2 é o mais escasso entre as produções em EL2, e sua baixa frequência indica que os participantes evitam sequências argumentativas, já que exigem maior precisão linguística e domínio das articulações argumentativas da escrita acadêmica em inglês. Quando aparece é reduzido a afirmações genéricas sobre a relevância do estudo ou sua necessidade de aprofundar determinado problema.

Em síntese, os textos em EL2 caracterizam-se por uma economia discursiva que, combinada com a predominância de M3, denota uma estratégia para minimizar os riscos linguísticos. Assim, os participantes escrevem resumos centrados em informações essenciais e evitando formulações que exigem maior fluência e domínio de outro idioma.

Os textos traduzidos automaticamente e autopós-editados (PEL2) revelam produções textuais com maior estabilidade organizacional, fluidez e uma menor dispersão estrutural entre as áreas. Assim como no EL2, o domínio do M3 reafirma que, mesmo com o suporte da tradução automática, os participantes continuam preferindo um foco nos elementos centrais. Entretanto, diferentemente dos textos escritos diretamente em L2, os resumos trazem o M3 de forma expandida, mais completa e articulada, o que indica que a influência do português como L1, mediada pela tradução automática, não apenas interfere no texto final, mas contribui positivamente para a construção retórica, especialmente no que diz respeito à expansão e articulação do M3.

Esse fenômeno é mais perceptível na área de Ciências Humanas (CH), na qual ocorreu um aumento no número total de elementos retóricas entre EL2 e PEL2. Isso sugeriu um maior conforto dos participantes ao revisar a saída da tradução automática para que eles escrevessem com mais segurança para aperfeiçoar o conteúdo e não só formular frases gramaticalmente aceitáveis. Assim, o PEL2 não alterou a hierarquia retórica dominante, mas permitiu que os participantes articulassem mais o movimento que melhor dominavam.

No Movimento 1 (M1) foi possível observar pequenos acréscimos nos textos em que os participantes introduzem contextualizações ou justificativas que não estavam presentes na EL2. Esses acréscimos, apesar de discretos, indicam que a autopós-edição permitiu aos autores incluir informações que poderiam ter sido omitidas por causa do esforço cognitivo necessário na escrita em EL2. O Movimento 2 (M2) permanece como o menos utilizado, reforçando que a autopós-edição, mesmo contribuindo para a clareza textual, não necessariamente estimulou a adoção de estratégias retóricas consideradas mais sofisticadas.

Em suma, os dados indicam que a PEL2 não transforma a escrita acadêmica, mas oferece um apoio viável para pesquisadores que com limitações de proficiência em inglês.

4.1.1 EL2 e PL2 – Correspondências

Após uma análise geral dos movimentos encontrados nos textos em EL2 e PEL2, a próxima etapa foi destinada a refinação dos dados. A comparação entre os textos, a partir dos movimentos que permanecem iguais e dos que não são preservados, oferece uma visão mais abrangente dos efeitos que a tradução automática exerceu na organização retórica dos resumos, além de mostrar como os participantes negociam, adaptam ou mantêm as escolhas discursivas. A tabela 6 apresenta os movimentos que permaneceram iguais no EL2 e no PEL2.

Tabela 6. Movimentos iguais em EL2 e PEL2.

EL2 = PEL2	Movimento 1		Movimento 2		Movimento 3		Total de Movimentos	
	absoluto	porcentagem	absoluto	porcentagem	absoluto	porcentagem	absoluto	total
CET	8	31%	3	12%	15	58%	26	100%
CH	2	13%	2	13%	12	75%	16	100%
CSA	0	0%	0	0%	5	100%	5	100%
Ordem 1	0	0%	2	14%	12	86%	14	100%
Ordem 2	5	38%	2	15%	6	46%	13	100%
Ordem 3	5	25%	1	5%	14	70%	20	100%

Fonte: elaboração própria.

Aqui, é importante esclarecer que o *corpus* foi dividido para um melhor entendimento dos dados. *A priori*, foram analisados nove textos em EL2 com correspondência em PEL2 (também composto por nove textos). Os que não apresentam correspondência semântica foram estudados à parte e estão apresentados posteriormente.

Ao isolar apenas os movimentos idênticos entre EL2 e PEL2, a tabela 6 demonstra a estrutura retórica preservada entre as versões e permite verificar se as áreas e as ordens influenciaram os resultados.

A área com o maior número absoluto de movimentos preservados é a Ciências Exatas e da Terra (CET), com 26 movimentos distribuídos relativamente de forma proporcional. Com 31% de M1, os elementos introdutórios que contextualizam o problema e/ou enunciam o campo de pesquisa permanecem, enquanto os 12% de M2, ainda que modestos, sugerem que, quando as estratégias de delimitação da lacuna são empregadas, elas permanecem coerentes o suficiente

para serem mantidas na autopós-edição e, com 58% de equivalência, o M3 reafirma a natureza descritiva e metodológica dos textos de CET.

Com apenas 16 movimentos mantidos, a área de Ciências Humanas (CH) parece passar por mudanças substanciais ao longo do processo de reescrita. O M3 detém 75% dos movimentos mantidos enquanto os M1 e M2 representam 13% cada, reforçando a ideia de que resumos dessa área são construídos como um espaço de exposição direta do estudo, com foco em descrever objetivos, métodos e resultados.

A área com o menor número de movimentos preservados, a Ciências Sociais Aplicadas (CSA), oferece um quadro distinto e marcado por uma simplicidade estrutural. Com 29% de M1 e 71% de M3, demonstra uma tendência a concisão e ao foco informativo, ao passo que a ausência de M2 sugere a falta de construção de nichos argumentativos tanto na EL2 quanto na PEL2.

Ao observar as três áreas conjuntamente, nota-se que, embora a quantidade total de movimentos preservados varie bastante, o padrão estrutural se mantém constante, com o M3 sendo o mais preservado. Já os movimentos introdutórios e de construção do nicho aparecem com menor frequência, sugerindo uma maior dependência das preferências estilísticas e do domínio linguístico.

Além da análise por área, a Tabela 6 permitiu explorar como a ordem das tarefas influenciou nos padrões relacionados à progressão do trabalho do participante e ao modo como ele articulou suas escolhas retóricas.

Os resultados referentes a Ordem 1 mostram um padrão retórico marcado pela centralidade do M3 e pela inexistência do M1 entre os movimentos preservados. Ao considerar os 14 movimentos existentes, 86% correspondem ao Movimento 3, confirmando seu papel como elemento mais estável, ao passo que a ausência do Movimento 1 evidencia que as sentenças de introdução, contextualização e delimitação temática não se mantiveram entre os textos. O Movimento 2 (M2) apareceu de forma tímida, com apenas duas ocorrências, indicando a ausência de construções relativas ao nicho da pesquisa.

Assim, o conjunto dos movimentos mantidos em EL2 e PEL2 na Ordem 1 demonstra que os textos se organizaram, em sua maioria, a partir do núcleo informativo, ao mesmo tempo que os movimentos mais complexos tenderam a ser modificados ou completamente apagados.

A Ordem 2, entretanto, apresenta uma distribuição mais equilibrada entre os movimentos preservados. Entre os 13 movimentos, 38% correspondem ao M1, indicando introduções e contextualizações mais sólidas, que parecem resistir às transformações entre EL2

e PEL2. Esse resultado sugere que, nessa ordem, os participantes preservam com maior consistência os elementos que situam o estudo e orientam a leitura.

Com 15% de M2 e 46% de M3, observa-se que essa ordem mantém não apenas as informações que descrevem objetivos, métodos e resultados, mas também parte das estratégias de construção do nicho. Esse equilíbrio indica que, ao retornar ao texto após alternar entre L1 e L2, o participante tende a estabilizar escolhas retóricas mais amplas, preservando tanto a apresentação do tema quanto segmentos que posicionam o estudo em um espaço discursivo mais definido.

A presença proporcionalmente mais elevada de M1 também sugere maior consistência na abertura textual, possivelmente em decorrência do apoio proporcionado pela formulação inicial em L1 antes da autopós-edição. Assim, a Ordem 2 se destaca como a configuração que mais preserva diversidade retórica entre as versões, articulando de maneira relativamente estável elementos introdutórios, a problematização do tema e a descrição do estudo.

A Ordem 3, por sua vez, apresenta um padrão intermediário entre a concisão observada na Ordem 1 e o equilíbrio característico da Ordem 2. Entre os 19 movimentos preservados, 68% correspondem ao Movimento 3, o que indica que as informações relacionadas aos objetivos, procedimentos e resultados continuam sendo as mais estáveis ao longo das três tarefas de produção textual.

O Movimento 1 aparece em 26% dos casos, mostrando que parte das estratégias de contextualização e apresentação do tema foi mantida, ainda que em proporção mais modesta do que na Ordem 2. Já o Movimento 2 ocorre apenas uma vez (5%), revelando que construções associadas à delimitação da lacuna ou à justificativa do estudo são pouco consistentes entre as versões e tenderam a ser modificadas ou descartadas durante a reescrita e a autopós-edição.

Assim, a Ordem 3 sinaliza um processo de preservação parcial em que apenas alguns elementos da abertura são mantidos, enquanto a maior parte da estrutura se reorganiza em torno da descrição do estudo. Esse padrão sugere que, mesmo após múltiplas etapas de reformulação, os participantes tendem a priorizar informações diretamente relacionadas ao conteúdo da pesquisa.

Ao analisar as três ordens de maneira integrada, observou-se que, embora cada sequência produza um padrão distinto de preservação dos movimentos, a dominância do M3 permanece constante. Por outro lado, os movimentos que introduzem o tema e constroem o

nicho, M1 e M2, respectivamente, tendem a ocorrer de forma menos frequente, o que demonstra que tais movimentos não se estabilizam com a mesma facilidade, permanecendo vulneráveis às transformações promovidas pela escrita e pela autopós-edição.

Como proposto, os passos também foram destrinchados e examinados. A Tabela 7 apresenta os dados do Movimento 1 e seus passos.

Tabela 7. M1 e seus passos iguais.

EL2 = PEL2	Movimento 1 - Estabelecendo o território									
	Passo 1		Passo 2		Passo 3		Passo 4		Total de Movimentos	
	absoluto	porcentagem	absoluto	porcentagem	absoluto	porcentagem	absoluto	porcentagem	absoluto	total
CET	2	25%	0	0%	1	13%	5	63%	8	100%
CH	0	0%	1	50%	0	0%	1	50%	2	100%
CSA	0	0%	0	0%	0	0%	0	0%	0	0%
Ordem 1	0	0%	0	0%	0	0%	0	0%	0	0%
Ordem 2	0	0%	0	0%	1	20%	4	80%	5	100%
Ordem 3	2	40%	1	20%	0	0%	2	40%	5	100%

Fonte: elaboração própria.

Na área de CET, o Movimento 1 é o mais frequente no conjunto de textos, aparecendo oito vezes. A maior parte dessas ocorrências é do Passo 4, que representa 63%, enquanto o Passo 1 aparece de forma moderada, com 25%. Os números vão ao encontro com o padrão descrito por Swales (1990), de que a função crucial da introdução é estabelecer limites no espaço de pesquisa e indicar que os textos da área focam em delimitar o território da pesquisa.

Na área de CH, o Movimento 1 aparece em apenas duas ocorrências, divididas igualmente entre o Passo 2 e o Passo 4, o que indicam ênfase, sobretudo, na justificativa de relevância e na delimitação temática. Além disso, a coincidência entre passos nas duas versões sugere que o pós-editor reconheceu o papel argumentativo desempenhado pelo trecho original, reconstituindo-o de forma adequada. Entretanto, área de CSA não apresentou qualquer ocorrência de Movimento 1, indicando a ausência explícita de estratégias de marcação de território.

Os achados, no geral, são corroborados e esperados. Hyland (2004) afirma que a escrita científica se baseia em estruturas retóricas padronizadas para garantir clareza e precisão, o que vai ao encontro com os padrões encontrados nos textos da Ciências Exatas e da Terra. Para o referido autor, textos de Ciências Humanas tendem a indicar sua relevância de forma mais

explícita nas introduções, e textos de áreas aplicadas, como a Ciências Sociais Aplicadas, normalmente priorizam propósito em detrimento do contexto em suas convenções retóricas.

Em relação às ordens de produção, a Ordem 1 não apresenta qualquer ocorrência de todos os passos do Movimento 1, indicando que essa configuração retórica não mobilizou estratégias formais de estabelecimento do território no *corpus*. A Ordem 2 é caracterizada por alta estabilidade estrutural, com 80% das ocorrências no Passo 4 e 20% no Passo 3, e a Ordem 3 apresenta o perfil mais completo e equilibrado do *corpus*, distribuindo igualmente os passos 1 e 4 (40% cada) e concentrando 20% no Passo 3.

A Tabela 8 exibe os resultados encontrados no Movimento 2 e seus respectivos passos.

Tabela 8. M2 e seus passos iguais.

EL2 = PEL2	Movimento 2 - Estabelecendo um nicho					
	Passo 1		Passo 2		Total de Movimentos	
	absoluto	porcentagem	absoluto	porcentagem	absoluto	total
CET	3	100%	0	0%	3	100%
CH	1	50%	1	50%	2	100%
CSA	0	0%	0	0%	0	0%
Ordem 1	2	100%	0	0%	2	100%
Ordem 2	1	50%	1	50%	2	100%
Ordem 3	1	100%	0	0%	1	100%

Fonte: elaboração própria.

No *subcorpus* de CET, o Movimento 2 aparece exclusivamente no Passo 1, com 100% das ocorrências. Isso significa que os textos dessa área utilizam o Movimento 2 unicamente para indicar a falta de robustez em pesquisas anteriores. Assim, quando o mesmo passo aparece em EL2 e PEL2, isso indica que a tradução automática combinada com a autopós-edição preservou a função comunicativa mais característica da área: a construção de uma justificativa cognitiva baseada na identificação de insuficiências ou de fronteiras ainda não transpostas.

Na área de CH, o Movimento 2 aparece com uma distribuição variada com 50% no Passo 1 e 50% no Passo 2. Isso indica que textos dessa área criaram seu nicho tanto pela

identificação de lacunas quanto pela problematização ou contestação do conhecimento existente. A área de CSA não apresentou ocorrências do Movimento 2 na tabela. Isso indica que os textos desse *subcorpus* não construíram explicitamente nichos retóricos por meio de lacunas ou de contestações.

A Ordem 1 apresentou 2 ocorrências, ambas no Passo 1. Isso significa que essa ordem construiu seu nicho exclusivamente ao apontar limitações na literatura existente e que quando EL2 e PEL2 apresentam o mesmo passo, isso indica que a tradução preservou a essência da ação retórica.

A Ordem 2 apresentou distribuição equilibrada entre os dois passos, revelando que essa ordem utilizou duas formas distintas de criar nicho: i. indicar problemas na literatura (Passo 1) e, ii. explicitar lacunas de pesquisas anteriores (Passo 2).

A Ordem 3 apresentou apenas um movimento, localizado exclusivamente no Passo 1, o que significa que essa ordem estabeleceu seu nicho exclusivamente pela identificação de problemas, demonstrando que a ação retórica central foi preservada, mesmo que a ordem apresentasse apenas uma ocorrência.

Em relação ao Movimento 3 e seus passos, é possível conferir os dados na Tabela 9.

Tabela 9. M3 e seus passos iguais.

EL2 = PEL2	Movimento 3 - Ocupando um nicho												Total de Movimentos	
	Passo 1		Passo 2		Passo 3		Passo 4		Passo 5	Passo 6	Passo 7	absoluto		
	absoluto	porcentagem	absoluto	porcentagem	absoluto	porcentagem	absoluto	porcentagem	absoluto	absoluto	absoluto			
CET	5	36%	0	0%	4	29%	5	36%	0	0	0	14	100%	
CH	5	42%	0	0%	2	17%	5	42%	0	0	0	12	100%	
CSA	3	60%	1	20%	1	20%	0	0%	0	0	0	5	100%	
Ordem 1	7	58%	1	8%	1	8%	3	25%	0	0	0	12	100%	
Ordem 2	2	29%	0	0%	2	29%	3	43%	0	0	0	7	100%	
Ordem 3	4	33%	0	0%	4	33%	4	33%	0	0	0	12	100%	

Fonte: elaboração própria.

Na área de CET, o Movimento 3 é amplamente distribuído entre o Passo 1 (36%), o Passo 3 (29%) e o Passo 4 (36%), revelando que os textos dessa área ocupam o nicho principalmente ao anunciar objetivos, descrever procedimentos e/ ou metodologia e apresentar resultados e/ou a relevância dos achados. A ocorrência do Passo 1, 3 e 4 em EL2 e PEL2

demonstrou que foram mantidas as funções comunicativas, preservando a clareza, o propósito e a previsibilidade estrutural do texto.

Os dados de CH mostram que o Movimento 3 se concentra principalmente nos Passos 1, com 42%, e 4, também com 42%, com menor presença do Passo 3, com 17%. Esses passos revelam uma ocupação do nicho ao anunciar objetivos, ao definir como a investigação seria conduzida e ao apresentar os resultados. A coincidência no Passo 1 significa que a apresentação dos objetivos foi mantida, no Passo 4, que a reportagem dos achados e/ou da relevância da pesquisa foram preservadas, e no Passo 3, que o detalhamento metodológico manteve seu papel.

Em CSA, o Movimento 3 apresentou predominância absoluta do Passo 1 (60%), seguido por ocorrências menores do Passo 2 (20%) e do Passo 3 (20%), sem presença do Passo 4. Isso indica que os textos dessa área majoritariamente ocuparam o nicho anunciando os objetivos do estudo. A coincidência no Passo 1 revela que o objetivo do estudo foi mantido, no Passo 2, que as questões e hipóteses de pesquisa foram preservadas, e no Passo 3, que a apresentação metodológica permaneceu a mesma.

A Ordem 1 mostra predominância do Passo 1, com pequenas ocorrências dos Passos 2 e 3, além de um pequeno aumento no Passo 4. Isso indica uma estrutura retórica que ocupa o nicho principalmente anunciando objetivos, mas também aventando hipóteses, detalhando os métodos de investigação e explicitando resultados.

A Ordem 2 e sua distribuição entre os Passos 1, 3 e 4 mostrou que essa categoria ocupou o nicho principalmente anunciando objetivos e apresentando resultados, com menor presença de outros detalhes.

A Ordem 3 distribuiu o Movimento 3 entre os Passos 1, 3 e 4, demonstrando um padrão equilibrado, semelhante ao encontrado nas Ciências Humanas e Ciências Exatas e da Terra. Essa ordem ocupa o nicho ao anunciar objetivos, detalhar a metodologia e reportar resultados, mantendo a progressão argumentativa e a estrutura retórica dos textos nas diferentes áreas e ordens analisadas.

A próxima análise envolveu os dados obtidos da tradução automática (TA) e seus produtos autopós-editados (PEL2) pelos participantes.

4.2 Tradução automática e sua autopós-edição

A Tabela 10 apresenta a distribuição geral dos três movimentos retóricos nos *subcorpora* analisados, considerando conjuntamente os dados de tradução automática (TA) e de tradução automática com a autopós-edição (PEL2).

Tabela 10. Dados gerais da comparação entre TA e PEL2.

TA = PEL2	Movimento 1		Movimento 2		Movimento 3		Total de Movimentos	
	absoluto	porcentagem	absoluto	porcentagem	absoluto	porcentagem	absoluto	total
CET	9	31%	4	14%	16	55%	29	100%
CH	5	14%	3	8%	28	78%	36	100%
CSA	8	38%	0	0%	13	62%	21	100%
Ordem 1	1	4%	3	13%	20	83%	24	100%
Ordem 2	11	32%	3	12%	18	56%	32	100%
Ordem 3	10	35%	1	3%	19	62%	30	100%

Fonte: elaboração própria.

De forma geral, observou-se um padrão altamente consistente de organização retórica, caracterizado pela predominância do Movimento 3 (M3), seguido pelo Movimento 1 (M1), enquanto o Movimento 2 (M2) apresentou ocorrência significativamente menor. Ao comparar os dados por áreas, é possível observar os padrões distintos que orientam a análise dos resultados.

Na CET, o M3 concentrou mais da metade das ocorrências, enquanto o M1 apresentou participação relevante e o M2 apareceu de forma residual, indicando que os textos dessa área privilegiaram a ocupação direta do nicho discursivo, precedida por uma contextualização funcional do território, com pouca ênfase na construção explícita de lacunas.

Na área de CH, o M3 foi ainda mais dominante, representando a maior parte das ocorrências, enquanto o M1 apareceu de forma secundária e o M2 manteve frequência reduzida. Esse resultado sugere que os textos dessa área concentraram seus esforços retóricos na apresentação, validação e interpretação das contribuições, com menor investimento na problematização explícita do campo.

Na CSA, observou-se a ausência do Movimento 2, enquanto os Movimentos 1 e 3 se distribuíram de forma complementar, com predominância do Movimento 3. Esse padrão

indicou que os textos dessa área não construíram explicitamente nichos discursivos por meio da identificação de lacunas ou contestação do conhecimento existente, concentrando-se principalmente na ocupação do nicho e na contextualização mínima do território.

Ao observar os dados por ordem, nota-se que a Ordem 1 apresenta forte predominância do M3, com presença bastante reduzida dos M1 e M2. Isso indica a priorização da ocupação imediata do nicho, com pouca contextualização prévia ou problematização.

A Ordem 2 apresentou um padrão distinto, com maior equilíbrio entre os Movimentos 1 e 3, embora o M3 continuasse predominante. Essas características sugerem uma combinação de forma mais explícita da contextualização do território com a ocupação do nicho.

Já a Ordem 3 retomou a predominância do M3, acompanhada por uma presença relevante do M1 e ocorrência mínima do M2. Esse padrão sugere que a ocupação do nicho permanece como ação retórica central, com contextualização complementar do território.

De forma geral, os dados indicam que a organização macroestrutural dos textos é altamente estável e que a autopós-edição não promove redistribuições significativas dos movimentos retóricos. A seguir, são apresentados os resultados detalhados da análise, com foco nos movimentos retóricos e seus passos. A Tabela 11 apresenta os dados do Movimento 1.

Tabela 11. M1 e seus passos.

TA = PEL2	Movimento 1 - Estabelecendo o território									
	Passo 1		Passo 2		Passo 3		Passo 4		Total de Movimentos	
	absoluto	porcentagem	absoluto	porcentagem	absoluto	porcentagem	absoluto	porcentagem	absoluto	total
CET	3	22%	0	0%	1	22%	5	56%	9	100%
CH	0	0%	2	40%	0	0%	3	60%	5	100%
CSA	0	0%	7	88%	1	13%	0	0%	8	100%
Ordem 1	1	100%	0	0%	0	0%	0	0%	1	100%
Ordem 2	0	0%	3	13%	2	12%	6	75%	11	100%
Ordem 3	2	20%	6	60%	0	0%	2	20%	10	100%

Fonte: elaboração própria

Na área de CET, o Movimento 1 distribui-se entre os Passos 1, 3 e 4, com clara predominância do Passo 4 (56%). Esse padrão apontou para uma tendência de estabelecimento do território discursivo por meio de um enquadramento ou de uma consolidação através da

literatura existente sobre um tema, em detrimento de uma progressão introdutória focada na importância de um tópico estudado.

Ainda que menos frequentes, os Passos 1 e 3 indicam o recurso a estratégias complementares, voltadas tanto à apresentação inicial do tópico quanto à caracterização mais ampla do campo.

Na área de CH, observa-se um comportamento ainda mais concentrado, já que o M1 se realiza exclusivamente no Passo 4. Tal configuração sugere uma forma altamente condensada de estabelecimento do território, centrada apenas na consolidação ou no enquadramento do campo discursivo, sem recorrer a estratégias graduais de introdução ou delimitação temática.

Já na área de CSA, identifica-se um comportamento distinto. O M1 ocorre majoritariamente no Passo 2 (88%), com apenas um caso no Passo 3, o que aponta para a preferência por estratégias que privilegiam a delimitação do foco específico ou a relevância do objeto de estudo para o mundo, em detrimento da explicitação de problemas a serem abordados ou de menção à literatura da área.

No que diz respeito às ordens, a Ordem 1 apresentou uma única realização do M1, restrita ao Passo 1, indicando que, quando essa ordem estabeleceu o território, isso ocorreu exclusivamente externando-se a importância de um tema para a disciplina, sem articulação com outros passos.

A Ordem 2, por sua vez, caracterizou-se por maior diversidade na realização do M1, com predominância do Passo 4 (75%), seguida pelos Passos 2 e 3, o que revela a combinação de estratégias de delimitação temática via referências a trabalhos anteriores e de enquadramento final do território discursivo.

Por fim, a Ordem 3 apresentou maior recorrência do Passo 2 (62%), acompanhada de presenças pontuais dos Passos 1, 3 e 4, sugerindo uma tendência a estabelecer o território principalmente por meio da delimitação do foco ou da caracterização específica do campo, ainda que outras estratégias sejam ocasionalmente mobilizadas.

Em relação aos passos do Movimento 2, a Tabela 12 apresenta os números de forma detalhada.

Tabela 12. M2 e seus passos.

TA = PEL2	Movimento 2 - Estabelecendo um nicho					
	Passo 1		Passo 2		Total de Movimentos	
	absoluto	porcentagem	absoluto	porcentagem	absoluto	total
CET	3	75%	1	25%	4	100%
CH	1	33%	2	67%	3	100%
CSA	0	0%	0	0%	0	0%
Ordem 1	2	67%	1	33%	3	100%
Ordem 2	1	33%	2	67%	3	100%
Ordem 3	1	100%	0	0%	1	100%

Fonte: elaboração própria.

Na área de CET, o Movimento 2 concentra-se predominantemente no Passo 1, que correspondeu a 75% dos casos observados, enquanto o Passo 2 apareceu de forma menos expressiva, representando 25%. Esse padrão sugere que os textos dessa área tendem a construir o nicho retórico sobretudo por meio da identificação de controvérsias, limitações ou inadequações na literatura existente, recorrendo apenas de maneira secundária à menção de lacunas ou escassez de estudos em determinado domínio do conhecimento.

Em CH, por sua vez, observa-se uma distribuição mais equilibrada entre os passos que compõem o M2, com 33% dos casos associados ao Passo 1 e 67% ao Passo 2, o que indica que os textos dessa área articulam a construção do nicho tanto pela identificação de lacunas quanto pela problematização ou contestação do conhecimento previamente estabelecido. Já na área de CSA, não se registraram realizações do M2, o que sugere que os textos dessa área não constroem explicitamente nichos retóricos por meio da identificação de lacunas ou da contestação do conhecimento existente.

Além disso, o fato de esse movimento não se manifestar nem na tradução automática nem na autopós-edição reforça a interpretação de que tal ausência está relacionada às convenções do gênero e às práticas discursivas da área, e não a efeitos decorrentes do processo tradutório.

No que diz respeito às ordens, observa-se que a Ordem 1 apresenta duas realizações do M2, ambas vinculadas ao Passo 1, indicando que, nessa configuração, a construção do nicho ocorre exclusivamente por meio da identificação de questões e limitações no conhecimento disponível.

A Ordem 2, em contraste, revela um comportamento inverso, uma vez que o Passo 2 se mostra predominante, representando 67% dos casos, em relação ao Passo 1, o que aponta para uma tendência de estabelecer o nicho prioritariamente por meio da indicação de fendas no conhecimento existente, em vez de se restringir à identificação de limitações de trabalhos anteriores.

Por fim, a Ordem 3 apresenta apenas uma realização do M2, localizada exclusivamente no Passo 1, o que sugere que, quando o nicho é construído nessa ordem, ele se fundamenta unicamente na identificação de problemas na literatura, sem o desdobramento em outras estratégias retóricas.

As informações do Movimento 3 são apresentadas na Tabela 13.

Tabela 13. M3 e seus passos.

TA = PEL2	Movimento 3 - Ocupando um nicho												Total de Movimentos absoluto total	
	Passo 1		Passo 2		Passo 3		Passo 4		Passo 5	Passo6	Passo 7			
	absoluto	porcentagem	absoluto	porcentagem	absoluto	porcentagem	absoluto	porcentagem	absoluto	absoluto	absoluto			
CET	5	31%	0	0%	5	31%	6	38%	0	0	0	16	100%	
CH	8	29%	1	4%	7	25%	12	43%	0	0	0	28	100%	
CSA	5	38%	2	15%	4	31%	2	15%	0	0	0	13	100%	
Ordem 1	8	40%	2	10%	3	15%	7	35%	0	0	0	20	100%	
Ordem 2	5	27%	0	0%	6	33%	7	38%	0	0	0	18	100%	
Ordem 3	5	26%	1	4%	7	39%	6	30%	0	0	0	19	100%	

Fonte: elaboração própria.

Na área de CET, o Movimento 3 manifestou-se de forma distribuída entre os Passos 1, 3 e 4, sendo este último o mais recorrente, com 38% dos registros. Esse arranjo sugere que os textos dessa área tendem a desenvolver o fechamento do espaço retórico por meio da explicitação de resultados, articulando essa estratégia com a de externar as implicações, aplicações ou consolidação dos achados, bem como a de fazer referência aos procedimentos metodológicos adotados.

Na área de CH, identificou-se uma concentração ainda mais acentuada no Passo 4, que respondeu por 43% das realizações do M3, seguido pelos Passos 1 e 3. Tal configuração aponta para uma orientação discursiva voltada à interpretação, validação ou discussão das implicações dos resultados, que assume papel central no momento de fechamento do espaço discursivo.

Na área de CSA, por sua vez, o M3 apresenta um perfil menos concentrado, com distribuição similar entre os Passos 1 e 3 (38% e 31%, respectivamente), acompanhada de ocorrências mais discretas dos Passos 2 e 4. Esse comportamento indica que, nessa área, o

fechamento do nicho se construiu em maior grau a partir da explicitação dos objetivos e da descrição de procedimentos, e em menor grau através do levantamento de hipóteses e da exposição e validação dos achados. Contudo, não houve priorização de um único tipo de estratégia.

No que diz respeito às ordens, observou-se que a Ordem 1 privilegiou os Passos 1 e 4, o que revela uma dinâmica de fechamento que combina a apresentação dos objetivos com os resultados e sua validação e síntese final.

A Ordem 2, em contraste, concentrou-se sobretudo nos Passos 4 e 3, com 38% e 33%, respectivamente, o que caracteriza um fechamento discursivo fortemente orientado à validação e à descrição dos resultados, em detrimento de movimentos mais introdutórios ou procedimentais.

Por fim, a Ordem 3 apresenta uma distribuição equilibrada entre os Passos 3 e 4, acompanhada de presença menos expressivas dos Passos 1 e 2. Esse padrão sugere que essa ordem tende a encerrar o percurso retórico principalmente por meio da descrição e da consolidação dos resultados, recorrendo de forma secundária a outras estratégias.

4.3 A classificação automática do INCEpTION

O *subcorpus* aqui analisado, composto por 111 movimentos retóricos, apresenta uma distribuição assimétrica entre os três movimentos identificados. Observa-se uma predominância consistente do Movimento 3 (M3), responsável por 72,9% das ocorrências, o que indica que a organização retórica dos textos se concentra majoritariamente no desenvolvimento do conteúdo principal. Em contrapartida, o Movimento 2 (M2) apresenta frequência reduzida (3,6%), configurando-se como um elemento pouco recorrente no *subcorpus*, enquanto o Movimento 1 (M1) ocupa uma posição intermediária, com 23,42% das ocorrências.

Tabela 14. Resultados gerais do IpN.

INCEpTION	Movimento 1		Movimento 2		Movimento 3		Total	
	absoluto	porcentagem	absoluto	porcentagem	absoluto	porcentagem	absoluto	porcentagem
EL2 = EL2 IpN	9	20%	1	2%	36	78%	46	100%
PEL2 = PEL2 IpN	9	28%	0	0%	23	72%	32	100%
TA = TA IpN	8	24%	3	9%	22	67%	33	100%

Fonte: elaboração própria.

Ao observar os modos de produção textual, é possível constatar uma diversidade maior na distribuição dos movimentos. EL2 se diferencia pela presença expressiva do M3 (78%), percentual significativamente superior ao registrado nos demais textos, o que aponta para uma organização retórica específica nesse modo de produção, assim como em PEL2 e TA.

É importante ressaltar que a aplicação da classificação do INCEpTION (IpN) não implicou qualquer modificação no número de sentenças nem no volume total de dados analisados, mantendo-se constantes os quantitativos de cada *subcorpus*. Trata-se de uma reclassificação analítica das mesmas sentenças, cujo efeito incide exclusivamente sobre a atribuição dos movimentos e de seus respectivos passos, que podem ser reorganizados ou reinterpretados a partir desse critério.

Nesse processo, não houve necessidade de alinhamento semântico entre sentenças, uma vez que os textos foram comparados internamente. Assim, as diferenças observadas entre os textos com e sem IpN dizem respeito à distribuição interna dos movimentos e passos, e não a alterações no número de sentenças.

Com as informações gerais, as análises detalhadas dos movimentos e seus passos foram agrupadas por modo de produção textual. Essa decisão foi tomada para estabelecer um cenário de análise controlado, no qual as divergências entre as classificações manual e automática não são consequência da ausência de correspondência semântica, mas do desempenho do INCEpTION. Com base nisso, os resultados possibilitam examinar os padrões segundo os quais as classificações manual e automática se aproximam ou se distanciam, preservando a validade ecológica do procedimento, uma vez que incidem sobre o mesmo material textual.

A partir dos dados apresentados, a discussão aborda os movimentos retóricos e seus respectivos passos e a Tabela 15 apresenta os números do Movimento 1.

Tabela 15. Correspondência entre M1 e seus passos.

INCEpTION	Movimento 1 - Estabelecendo o território									
	Passo 1		Passo 2		Passo 3		Passo 4		Total de Movimentos	
	absoluto	porcentagem	absoluto	porcentagem	absoluto	porcentagem	absoluto	porcentagem	absoluto	total
EL2 = EL2 IpN	0	0%	0	0%	1	14%	6	86%	7	100%
PEL2 = PEL2 IpN	1	11%	1	11%	2	22%	5	56%	9	100%
TA = TA IpN	0	0%	7	88%	0	0%	1	12%	8	100%

Fonte: elaboração própria.

No recorte EL2 (textos originalmente escritos em inglês), observa-se uma forte concentração no Passo 4 (86%), seguido pelo Passo 3 (14%), sem qualquer ocorrência dos Passos 1 e 2. Esse padrão é altamente coerente com aquele observado anteriormente na classificação manual dos textos em inglês. A predominância do Passo 4 sugere que o INCEpTION foi capaz de identificar corretamente o passo mais recorrente e prototípico do Movimento 1 nesse tipo de texto.

A ausência dos demais passos não indica necessariamente falha do sistema, mas parece refletir uma característica estrutural do próprio *corpus*, no qual esses passos são pouco mobilizados. Assim, nesse modo de produção, o desempenho do INCEpTION pode ser considerado satisfatório e alinhado à análise humana, sobretudo no reconhecimento das estratégias retóricas dominantes.

Em TA, os dados revelam uma concentração expressiva do Passo 2, que representa 88% das ocorrências, com apenas 12% atribuídos ao Passo 4. Esse padrão coincide com o observado na classificação manual dos textos traduzidos automaticamente, conforme indicado em análises anteriores. Essa correspondência sugere que, em textos de tradução automática, os marcadores linguísticos associados ao Passo 2 apresentam maior incidência, o que favorece sua identificação tanto por anotadores humanos quanto pelo sistema automático.

Nesse sentido, o INCEpTION demonstra alto grau de precisão nesse modo de produção, sendo capaz de replicar quase integralmente o padrão identificado manualmente. Ao mesmo tempo, a baixa diversidade de passos indica que o desempenho positivo do sistema está associado a um perfil retórico mais simplificado desses textos, o que reduz a ambiguidade classificatória.

No caso de PEL2, os dados indicam um cenário mais complexo para a avaliação do INCEpTION, já que a presença de todos os passos do Movimento 1 resulta em uma organização retórica mais variada, distinta daquela identificada em EL2 e TA. Isso corresponde ao mesmo conjunto de movimentos e passos considerado na classificação manual, em que não há predominância de um único passo.

No processo de classificação automática, essa configuração mais diversa pode impor maior dificuldade ao INCEpTION, sobretudo na identificação de passos menos frequentes, como os Passos 1 e 2. Ainda assim, o Passo 4 permanece predominante, o que sugere que sua identificação é favorecida por sua maior frequência no texto, mesmo diante de uma distribuição mais diversificada dos passos.

A seguir, apresentam-se os dados do Movimento 2 e de seus passos.

Tabela 16. Correspondência entre M2 e seus passos.

INCEpTION	Movimento 2 - Estabelecendo um nicho					
	Passo 1		Passo 2		Total de Movimentos	
	absoluto	porcentagem	absoluto	porcentagem	absoluto	total
EL2 = EL2 IpN	1	100%	0	0%	1	100%
PEL2 = PEL2 IpN	0	0%	0	0%	0	0%
TA = TA IpN	1	33%	2	67%	3	100%

Fonte: elaboração própria.

Ao observar a Tabela 16, é possível notar que nos textos EL2, foi identificada uma única ocorrência do Movimento 2, realizada exclusivamente por meio do Passo 1. Esse padrão indica que o M2 apresenta baixa incidência e ausência de variação interna o que reflete diretamente na classificação, tanto manual quanto automática, nas quais o movimento aparece de forma pontual. Nesse contexto, o desempenho do INCEpTION não é impactado por ambiguidades classificatórias, considerando que se trata de uma ocorrência pontual.

Em PEL2, o M2 não apresenta ocorrências com correspondência entre a classificação manual e a automática no conjunto analisado. Já nos textos traduzidos automaticamente (TA), observam-se três casos do M2 com correspondência entre as classificações manual e

automática, distribuídos entre os Passos 1 e 2, o que indica alinhamento na atribuição dos passos.

Na sequência, apresentam-se os dados do Movimento 3 e de seus passos.

Tabela 17. Correspondência entre M3 e seus passos.

INCEpTION	Movimento 3 - Ocupando um nicho												Total de Movimentos	
	Passo 1		Passo 2		Passo 3		Passo 4		Passo 5	Passo 6	Passo 7	absoluto		
	absoluto	porcentagem	absoluto	porcentagem	absoluto	porcentagem	absoluto	porcentagem	absoluto	absoluto	absoluto			
EL2 = EL2 IpN	8	42%	0	0%	5	26%	6	32%	0	0	0	19	100%	
PEL2 = PEL2 IpN	6	33%	0	0%	6	33%	6	33%	0	0	0	18	100%	
TA = TA IpN	6	32%	1	5%	4	21%	8	42%	0	0	0	19	100%	

Fonte: elaboração própria.

Os dados da Tabela 17 mostram que, nos textos EL2, o Movimento 3 apresenta uma concentração no Passo 1 (42%), seguido pelo Passo 4 (32%) e pelo Passo 3 (26%), sem registros dos demais passos. Esse padrão coincide com o identificado na classificação manual, indicando correspondência entre as duas classificações. A predominância do Passo 1 sugere que o INCEpTION foi capaz de acompanhar a distribuição mais frequente dos passos do M3, e a ausência dos demais passos não aponta uma limitação do sistema.

Nos textos traduzidos automaticamente (TA), observa-se um padrão semelhante ao de EL2, com predominância do Passo 4, seguido pelos Passos 1 e 3, além de uma ocorrência pontual do Passo 2. Esse resultado também se mostra coerente com a classificação manual, indicando alinhamento entre as classificações. A presença residual do Passo 2 amplia a diversidade de passos, mas não altera o padrão geral de concentração do movimento, o que sugere que o INCEpTION foi capaz de replicar adequadamente a distribuição observada manualmente.

No caso de PEL2, o M3 apresenta uma distribuição equilibrada entre os Passos 1, 3 e 4 sem prevalência de um único Passo 2. Esse padrão corresponde ao mesmo conjunto de passos identificado na classificação manual, indicando correspondência entre as duas classificações quanto à composição do movimento. No processo de classificação automática, essa distribuição mais equilibrada pode fazer com que um passo se destaque menos, mas não compromete o reconhecimento do M3 como um todo, uma vez que o INCEpTION mantém a identificação dos mesmos passos observados manualmente.

5 CONSIDERAÇÕES FINAIS

Esta dissertação reportou uma investigação dos movimentos e passos retóricos em um *corpus* comparável composto por resumos escritos por doutorandos da Universidade Federal de Minas Gerais em inglês como língua adicional (L2) e por resumos autopós-editados nessa mesma língua a partir de textos desse gênero originalmente escritos em português como primeira língua (L1). Para responder ao objetivo geral da pesquisa, o estudo formulou perguntas voltadas a identificar quantos e quais movimentos e passos retóricos ocorrem em cada modo de produção, bem como em que medida esses elementos se aproximam ou se distanciam quando considerados a área do conhecimento e a ordem de realização das tarefas.

Ao assumir essa perspectiva, a pesquisa se ancora na compreensão da escrita acadêmica como prática social situada e considera a tradução automática não apenas como ferramenta tecnológica, mas como parte constitutiva das condições contemporâneas de produção do texto científico. De modo geral, os resultados sugerem que a organização retórica dos resumos analisados apresenta uma base estrutural relativamente consistente, ainda que marcada por variações relevantes. Essas variações parecem relacionar-se principalmente à área do conhecimento e ao modo de produção textual e, em menor medida, à ordem de realização das tarefas.

Observa-se, de forma recorrente, que os participantes tendem a mobilizar estratégias retóricas voltadas à apresentação direta do conteúdo central de suas pesquisas. Em todos os conjuntos de textos analisados, verifica-se a predominância do Movimento 3, responsável pela exposição de objetivos, procedimentos metodológicos, resultados e contribuições, enquanto os Movimentos 1 e 2 apresentam ocorrência mais limitada e distribuição menos sistemática ao longo do *corpus*.

O Movimento 1, quando mobilizado, tende a desempenhar uma função introdutória mais restrita, frequentemente delimitando literatura do campo de estudo do objeto investigado, mas com pouca expansão contextual. Já o Movimento 2 mostrou-se pouco recorrente e, na maioria dos casos, realizado de forma implícita ou genérica, sem a explicitação clara de lacunas, controvérsias ou limitações do conhecimento existente. Esses resultados podem ser compreendidos à luz das descrições clássicas do modelo CARS, que apontam os Movimentos 1 e 2 como particularmente sensíveis ao domínio discursivo e ao grau de familiaridade do autor

com as práticas retóricas da comunidade científica, especialmente quando a escrita ocorre em língua adicional.

A análise dos passos retóricos reforça essa interpretação ao indicar que a diversidade de passos é significativamente maior no Movimento 3 do que nos demais movimentos. Enquanto os passos associados à apresentação de objetivos, métodos e resultados são amplamente mobilizados pelos participantes, aqueles vinculados à contextualização ampla do campo ou à construção explícita de nichos discursivos aparecem de forma mais restrita. Estudos sobre escrita acadêmica em língua adicional sugerem que esse tipo de assimetria está frequentemente associado à dificuldade de mobilizar recursos discursivos de natureza avaliativa e argumentativa, que exigem não apenas conhecimento linguístico, mas também domínio das convenções disciplinares e das expectativas retóricas do gênero resumo acadêmico.

A comparação entre os textos produzidos em EL2 e em PEL2 indica que a tradução automática seguida de autopós-edição não promove alterações substantivas na organização retórica global dos resumos. Em ambos os modos de produção, observam-se padrões semelhantes de distribuição dos movimentos, com destaque para a centralidade do Movimento 3 e a presença reduzida dos Movimentos 1 e 2. Esses achados sugerem que a autopós-edição atua predominantemente como um recurso de refinamento textual, contribuindo para a fluidez e a clareza informacional, mas sem incidir de forma significativa sobre escolhas retóricas mais amplas, que parecem depender sobretudo do conhecimento de gênero e do letramento acadêmico dos autores.

No que se refere à análise dos dados por área do conhecimento, foi possível responder à pergunta de pesquisa relacionada à variação retórica conforme o domínio disciplinar. Nas Ciências Exatas e da Terra, observou-se maior incidência do Movimento 1, bem como uma distribuição mais equilibrada dos passos do Movimento 3, refletindo práticas discursivas que valorizam a delimitação temática, a explicitação metodológica e a apresentação sistemática de evidências empíricas. Esse padrão sugere uma maior estabilização das convenções retóricas associadas ao gênero nesse domínio.

Nas Ciências Humanas, a organização retórica mostrou-se fortemente concentrada no Movimento 3, com menor diversidade de passos e baixa recorrência do Movimento 1. Esse resultado aponta para uma escrita orientada prioritariamente à exposição das contribuições do estudo, com menor investimento em estratégias de contextualização ampla do campo. Já nas Ciências Sociais Aplicadas, destacou-se a ausência do Movimento 2 e uma ocupação do espaço

discursivo centrada nos objetivos e na aplicabilidade dos resultados, indicando especificidades na forma como o espaço de pesquisa é construído nesse domínio.

Esses padrões permitem compreender que a área do conhecimento exerce influência significativa sobre a forma como os movimentos e passos retóricos são realizados. Ainda que a centralidade do conteúdo informativo seja comum aos diferentes domínios, a variação disciplinar molda a seleção, a frequência e a combinação dos recursos retóricos mobilizados, reforçando a compreensão da escrita acadêmica como prática situada e socialmente construída.

Quanto à ordem de realização das tarefas, os dados indicam um impacto limitado sobre a organização retórica dos textos. Embora tenham sido observadas variações pontuais na distribuição de movimentos e passos, não se identificaram reorganizações estruturais profundas associadas à sequência de produção textual. Esse resultado sugere que as convenções do gênero resumo e as práticas discursivas disciplinares exercem influência mais determinante sobre a organização retórica do texto do que o percurso processual específico adotado pelos participantes.

Além da análise dos movimentos e passos retóricos nos textos EL2 e PEL2, esta pesquisa também incluiu a comparação entre pares de tradução automática e autopós-edição. Ainda que o foco central tenha sido a organização retórica, os resultados indicaram que a autopós-edição contribui para a melhoria da fluidez textual, especialmente no que se refere à coerência e à clareza informacional, sem modificar de forma substantiva a estrutura global do conteúdo produzido.

A comparação entre a classificação manual e a classificação automática realizada pelo INCEpTION revelou uma correspondência razoável, sobretudo no reconhecimento dos movimentos mais frequentes. No entanto, a classificação automática apresentou limitações na identificação de passos menos recorrentes e de funções discursivas mais sutis, que exigem interpretação contextual e conhecimento retórico, aspectos que se mostraram mais adequadamente capturados pela análise manual.

Os achados desta pesquisa devem ser interpretados à luz de algumas limitações. O *corpus* analisado é relativamente restrito e vincula-se a uma única instituição, o que limita a generalização dos resultados. Ademais, o contexto empírico-experimental de produção textual, embora controlado, não reproduz integralmente as condições reais de produção acadêmica,

como prazos editoriais, processos de avaliação por pares e pressões institucionais associadas à publicação científica.

Outro aspecto a considerar refere-se ao nível de proficiência em inglês dos participantes, que foi autodeclarado, o que impede análises mais precisas sobre sua influência nas escolhas retóricas observadas. Além disso, o processo de anotação manual dos movimentos e passos, embora teoricamente fundamentado e validado, envolve julgamentos interpretativos e, como tal, está sujeito a vieses analíticos.

Apesar dessas limitações, esta pesquisa contribui para os debates na Linguística Aplicada e nos Estudos da Tradução ao evidenciar como a escrita acadêmica em língua adicional se organiza retoricamente em contextos mediados pela tradução automática e pela autopós-edição. Os resultados indicam que esses recursos tecnológicos tendem a atuar como instrumentos de apoio e refinamento textual, mas não garantem, por si só, a ampliação do repertório retórico mobilizado pelos autores.

Ao analisar de forma integrada os movimentos e passos retóricos em textos produzidos em EL2 e em PEL2, o estudo reforça que a tradução automática, mesmo combinada à intervenção do autor, não substitui o conhecimento discursivo necessário à mobilização de estratégias retóricas mais complexas, particularmente aquelas associadas à contextualização do campo e à construção explícita de nichos discursivos. Nesse sentido, os achados apontam para a relevância de ações formativas que articulem letramento acadêmico, estudos de gênero e letramento em tradução automática.

Os resultados também apontam para possibilidades de investigações futuras que aprofundem a análise da organização retórica de resumos acadêmicos em língua adicional, especialmente no que se refere à baixa recorrência dos Movimentos 1 e 2 e à centralidade do Movimento 3 nos diferentes modos de produção textual. A ampliação do *corpus* e a consideração de outros contextos institucionais podem favorecer uma compreensão mais refinada da influência das áreas do conhecimento na realização dos movimentos e passos retóricos.

Além disso, estudos futuros podem explorar de forma mais detalhada a relação entre tradução automática, autopós-edição e estratégias retóricas de natureza avaliativa. A análise da classificação automática realizada pelo INCEpTION, em diálogo com a anotação manual, também se apresenta como um caminho relevante para compreender os limites dessas

ferramentas na identificação de nuances discursivas e para subsidiar seu aprimoramento em pesquisas de análise de gêneros acadêmicos.

Conclui-se, assim, que a análise dos movimentos e passos retóricos em resumos acadêmicos permite compreender não apenas padrões recorrentes de organização textual, mas também as condições concretas em que esses textos são produzidos no contexto da comunicação científica contemporânea. Ao considerar a escrita acadêmica em língua adicional como uma prática social situada e mediada por tecnologias, este estudo evidencia que a tradução automática e a autopós-edição integram o cotidiano de pesquisadores em formação, especialmente em contextos institucionais não anglófonos marcados por demandas crescentes de internacionalização.

Nesse cenário, a mediação tecnológica não se apresenta como solução automática para desafios retóricos e discursivos, mas como um recurso que amplia possibilidades de produção textual, ao mesmo tempo em que expõe limites relacionados ao acesso desigual às práticas de letramento acadêmico e ao conhecimento de gênero. Ao articular análise de gêneros, tradução automática e escrita acadêmica, a pesquisa reforça a necessidade de abordagens formativas que considerem essas práticas de modo crítico e contextualizado, contribuindo para a circulação mais ampla do conhecimento científico sem desconsiderar as condições linguísticas, institucionais e sociais em que ele é produzido.

6 REFERÊNCIAS

- ALVES, F. A triangulação como opção metodológica em pesquisas empírico-experimentais em tradução. In: PAGANO, Adriana Silvino (org.). Metodologias de pesquisa em tradução. Série Estudos Linguísticos, v. 3. Belo Horizonte: FALE-UFMG, 2001. p. 69-92.
- BERGOLI, P. P. .; BOCORNY, A. E. P. .; PIRES DE FREITAS, A. . L. Inglês para fins acadêmicos em uma instituição de ensino superior pública brasileira: uma análise de necessidades. *The Specialist*, [S. l.], v. 43, n. 2, p.93-121, 2022. DOI: 10.23925/2318-7115.2022v43i2a7.
- BERBER-SARDINHA, T. *Linguística de Corpus*. Barueri, SP: Manole, 2004.
- BERGOLI, Talita; BOCORNY, Aline; PIRES DE FREITAS, Luciana Maria Almeida. O ensino de inglês para fins acadêmicos em universidades federais brasileiras: demandas, desafios e perspectivas. *Revista Brasileira de Linguística Aplicada*, Belo Horizonte, v. 22, n. 3, p. 1–25, 2022.
- BONELLI, E. *Corpus linguistics at work*. Amsterdam; Philadelphia: John Benjamins, 2010.
- BOWER, L. Machine translation literacy instruction for translators and other language professionals. *The Interpreter and Translator Trainer*, v. 14, n. 2, p. 185–201, 2020.
- BOWER, L.; BUITRAGO CIRO, J. *Machine Translation and Global Research: Towards Improved Machine Translation Literacy in the Scholarly Community*. ed. Emerald Publishing Limited, 2019.
- BRITTO, F. A. et al. Comparing Supervised Machine Learning techniques for Genre Analysis in Software Engineering research Articles. In *Proceedings of the International Conference on Recent Advances in Natural Language Processing, 2021*, 63–72. Available on: <https://aclanthology.org/2021.ranlp-1.8/>
- CHOI, K; KO, Y. Robust fine-tuning for low-resource NLP: combining adversarial and metric-based learning to mitigate overfitting. *Expert Systems With Applications*, v. 288, 2025, p. 127737. DOI: <https://doi.org/10.1016/j.eswa.2025.127737>.
- CRYSTAL, D. The future of Englishes. In: GRADDOL, David; MEINHOF, Ulrike (org.). *English in a changing world*. London: AILA, 1997b.
- DE SWAAN, A. *Words of the world: the global language system*. Cambridge: Polity Press, 2001.
- DUDLEY-EVANS, T.; ST JOHN, M.J. *Developments in ESP: a multi-disciplinary approach*. 1. ed. London: Cambridge University Press, 1998.
- ESCARTÍN, C. P.; GOULET, MC. Self-post-editing of machine translation: an exploratory study of revision behavior. *Translation, Cognition & Behavior*, Amsterdam, v. 4, n. 1, p. 85–113, 2021. DOI: 10.1075/tcb.00047.esc
- FLOWERDEW, J. Some thoughts on English for Research Publication Purposes (ERPP) and related issues. *Language Teaching*, v. 48, n. 2, p. 250-262, 2015. DOI: 10.1017/S0261444812000523.
- GUIMARAES, L.G. *Escrita, tradução automática e autopós-edição de resumos acadêmicos: uma análise textual com base em taxonomia de erros*. Dissertação (Pós-graduação em Estudos

Linguísticos). Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, 2025.

HALLIDAY, M. A. K.; MATTHIESSEN, C. Halliday's introduction to functional grammar. Milton Park, Abingdon, Oxon: Routledge, 2014.

HAMAEL, R.E. (2007) The Dominance of English in the International Scientific Periodical Literature and the Future of Language Use in Science. *AILA Review*, 20, 53-71. <https://doi.org/10.1075/aila.20.06ham>.

HOLMES, J.S. The name and nature of translation studies. In: VENUTI, L. (ed.). *The translation studies reader*. 1. ed. London; New York: Routledge, 2000. p. 172-185.

HUTCHINS, J. Machine translation: A concise history. *Computer aided translation: Theory and practice*, v. 13, n. 29-70, 2007.

HYLAND, K. *Disciplinary Discourses: social interactions in academic writing*. ed. The University of Michigan, 2004.

INCEPTION PROJECT. INCEPTION user guide. Disponível em: <https://inception-project.github.io/releases/38.5/docs/user-guide.htm>

JAKOBSEN, A. L. Research methods in translation—Translog. In: *Computer Key-Stroke Logging and Writing*. Brill, 2006. p. 95-105.

JENKINS, J. *English as a Lingua franca in the International University: the politics of academic English language policy*. ed. Routledge, 2014.

JUN, L; WANG, A. ZHANG, C. Integrating Curriculum Design Theory into ESP Course Construction: Aviation English for Aircraft Engineering. *Open Journal of Modern Linguistics*, Vol.4 No.2, April 28, 2014.

KAPLAN, R. *The Oxford handbook of applied linguistics*. Oxford: Oxford University Press, 2001.

KRINGS, H. P. *Repairing texts: Empirical investigations of machine translation post-editing processes*. Kent State University Press, 2001.

LEA, M.R.; STREET, B.V. The academic literacies model: theory and applications. *Theory into Practice*, v. 45, n. 4, p. 368–377, 2006.

LUONG, M. T. *Neural machine translation*. 2016. Tese Stanford University, Stanford, 2016.

MADIANOU, M.; MILLER, D. *Migration and new media: transnational families and polymedia*. London: Routledge, 2012.

MARCUSCHI, L. A. Gêneros textuais: definição e funcionalidade. In: DIONÍSIO, A. P.; MACHADO, A. R.; BEZERRA, M. A. (org.). *Gêneros textuais & ensino*. Rio de Janeiro: Lucerna, 2002. p. 19–36.

McCARTHY, M.; O'KEEFFE, A. Historical perspective: what are corpora and how have they evolved? In: O'KEEFFE, A.; McCARTHY, M. (org.). *The Routledge handbook of corpus linguistics*. London: Routledge, 2010. p. 3–13.

MUNDAY, J. *Introducing Translation Studies: Theories and Applications*. 4. ed. London: Routledge, 2016.

NUNES, L. P. *A produção textual acadêmica em distintos domínios: investigando a escrita e a pós-edição em língua inglesa sob a perspectiva linguística e cognitiva*. Projeto de pesquisa (Pós-graduação em Estudos Linguísticos). 29 p. Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, 2021.

O'BRIEN, S.; SIMARD, M.; GOULET, MJ. Measuring the cognitive effort of post-editing: an eye-tracking study. *Machine Translation*, v. 32, n. 1–2, p. 1–30, 2018.

PYM, A. Exploring Translation Theories. *Cadernos de Tradução*. Tradução de Eduardo César Godarth, Yéo N'gana e Bernardo Sant'Anna. , v. 36, n. 3, p. 214–268, 2016.

QUN, L.; XIAOJUN, Z. Machine translation. In: SIN-WAI, C. (ed.). *The Routledge Encyclopedia of Translation Technology*. London; New York: Routledge, 2015. p. 105-119.

SANCHEZ, J. A. Applicability and variation of Swales' Cars Model to Applied Linguistics Article Abstracts. *ELIA* v. 18, 2018, pp. 213-240.

ST. JOHN, M. J. English for specific purposes. In: BYRAM, M. (org.). *Routledge encyclopedia of language teaching and learning*. London: Routledge, 1998. p. 298–300.

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

STREET, B. V. What's "new" in new literacy studies? Critical approaches to literacy in theory and practice. *Current Issues in Comparative Education*, v. 5, n. 2, p. 77–91, 2003.

SWALES, J. M. *Genre Analysis: English in academic and research settings*. ed. Cambridge University Press, 1990.

SWALES, J. M.; FEAK, C. B. *Abstracts and the Writing of Abstracts*. ed. The University of Michigan, 2009.

TOURY, G. *Descriptive Translation Studies and beyond*. Amsterdam: Benjamins, 1995.