

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
INSTITUTO DE LETRAS E LINGUÍSTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ESTUDOS LINGUÍSTICOS

STHEFANY KAMILLA ALVES

**TRADUÇÃO AUTOMÁTICA NO BRASIL: (DES)AUTOMATIZANDO
NÚMEROS E DISCURSOS DOS ÚLTIMOS 30 ANOS**

UBERLÂNDIA

2022

STHEFANY KAMILLA ALVES

**TRADUÇÃO AUTOMÁTICA NO BRASIL: (DES)AUTOMATIZANDO
NÚMEROS E DISCURSOS DOS ÚLTIMOS 30 ANOS**

Defesa de Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Estudos Linguísticos do Instituto de Letras e Linguística da Universidade Federal de Uberlândia, como requisito parcial para a obtenção do título de mestre em Estudos Linguísticos.

Área de Concentração: Estudos em Linguística e Linguística Aplicada

Linha de pesquisa: Linguagem, ensino e sociedade

Orientador: Profa. Dra. Marileide Dias Esqueda

UBERLÂNDIA

2022



UNIVERSIDADE FEDERAL DE UBERLÂNDIA
Coordenação do Programa de Pós-Graduação em Estudos
Linguísticos
Av. João Naves de Ávila, nº 2121, Bloco 1G, Sala 1G256 - Bairro Santa Mônica,
Uberlândia-MG, CEP 38400-902
Telefone: (34) 3239-4102/4355 - www.ileel.ufu.br/ppgel - secppgel@ileel.ufu.br



ATA DE DEFESA - PÓS-GRADUAÇÃO

Programa de Pós-Graduação em:	Estudos Linguísticos				
Defesa de:	Dissertação de mestrado				
Data:	Vinte e nove de setembro de dois mil e vinte e dois	Hora de início:	09:00	Hora de encerramento:	12:00
Matrícula do Discente:	11922ELI023				
Nome do Discente:	Sthefany Kamilla Alves				
Título do Trabalho:	Tradução automática no Brasil: (des)automatizando números e discursos dos últimos 30 anos				
Área de concentração:	Estudos em linguística e linguística aplicada				
Linha de pesquisa:	Linguagem, ensino e sociedade				
Projeto de Pesquisa de vinculação:	Ensino e aprendizagem de Tradução: pressupostos epistemológicos, teóricos e metodológicos				

Reuniu-se, por videoconferência, a Banca Examinadora, designada pelo Colegiado do Programa de Pós-graduação em **Estudos Linguísticos**, assim composta: Professores Doutores: [Silvana Maria de Jesus \(UFU\)](#); [Roberto Carlos de Assis \(UFPB\)](#); e [Marileide Dias Esqueda \(UFU\)](#), orientadora da candidata.

Iniciando os trabalhos a presidente da mesa, Dra. Marileide Dias Esqueda, apresentou a Comissão Examinadora e a candidata, agradeceu a presença do público, e concedeu à Discente a palavra para a exposição do seu trabalho. A duração da apresentação da Discente e o tempo de arguição e resposta foram conforme as normas do Programa.

A seguir a senhora presidente concedeu a palavra, pela ordem sucessivamente, às examinadoras, que passaram a arguir a candidata. Ultimada a arguição, que se desenvolveu dentro dos termos regimentais, a Banca, em sessão secreta, atribuiu o resultado final, considerando a candidata:

Aprovada.

Esta defesa faz parte dos requisitos necessários à obtenção do título de Mestre.

O competente diploma será expedido após cumprimento dos demais requisitos, conforme as normas do Programa, a legislação pertinente e a regulamentação interna da UFU.

Nada mais havendo a tratar foram encerrados os trabalhos. Foi lavrada a presente ata que após lida e achada conforme foi assinada pela Banca Examinadora.



Documento assinado eletronicamente por **Roberto Carlos de Assis, Usuário Externo**, em 29/09/2022, às 11:25, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Silvana Maria de Jesus, Professor(a) do Magistério Superior**, em 29/09/2022, às 11:33, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Marileide Dias Esqueda, Professor(a) do Magistério Superior**, em 29/09/2022, às 11:42, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site https://www.sei.ufu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **3959225** e o código CRC **7D10EB17**.

Referência: Processo nº 23117.073563/2022-29

SEI nº 3959225

Ficha Catalográfica Online do Sistema de Bibliotecas da UFU
com dados informados pelo(a) próprio(a) autor(a).

A474 Alves, Sthefany Kamilla, 1988-
2022 TRADUÇÃO AUTOMÁTICA NO BRASIL: (DES)AUTOMATIZANDO
NÚMEROS E DISCURSOS DOS ÚLTIMOS 30 ANOS [recurso
eletrônico] / Sthefany Kamilla Alves. - 2022.

Orientadora: Marileide Dias Esqueda.
Dissertação (Mestrado) - Universidade Federal de
Uberlândia, Pós-graduação em Estudos Linguísticos.
Modo de acesso: Internet.
Disponível em: <http://doi.org/10.14393/ufu.di.2022.603>
Inclui bibliografia.
Inclui ilustrações.

1. Linguística. I. Esqueda, Marileide Dias, 1973-,
(Orient.). II. Universidade Federal de Uberlândia. Pós-
graduação em Estudos Linguísticos. III. Título.

CDU: 801

Bibliotecários responsáveis pela estrutura de acordo com o AACR2:
Gizele Cristine Nunes do Couto - CRB6/2091
Nelson Marcos Ferreira - CRB6/3074

AGRADECIMENTOS

À minha orientadora, Marileide Dias Esqueda, por estar sempre apontando o melhor caminho a seguir, sempre disposta a compartilhar e agregar conhecimento, por dar voz aos pensamentos de seus alunos e, acima de tudo, inspirar. Ao ser exemplo de profissional que é, ela se torna uma inspiração. Agradeço pela seriedade com que ela encarou cada documento, cada tópico dessa análise e, por consequência, cada linha desse texto. Agradeço pela paciência que comigo teve, pois em momentos da pandemia da Covid-19 não foi nada fácil.

Aos meus companheiros de pós-graduação, que sempre me estimularam a seguir com a minha pesquisa, participar de grupos de pesquisa e eventos e pela companhia. Com vocês, não me senti só!

Aos parceiros do grupo de pesquisa GETTEC (Grupo de Estudo e Pesquisas em Tradução, Tecnologias, Ensino e Cienciometria), especialmente a Rayanne Silva Barbosa e Flávio de Sousa Freitas, parceiros de estudos, fontes de motivação, conselhos e conversas enriquecedoras.

Ao PPGEL, por realizar um sonho. Fonte de orgulho!

Aos professores do PPGEL, que cada um, à sua maneira, colaborou para meu crescimento pedagógico e pessoal.

À FAPEMIG, pelo apoio concedido à realização dessa pesquisa.

Por fim, ao Wender, meu companheiro, que com muita paciência, me apoiou, incentivou, me deu força nos momentos em que eu pensei que não ia conseguir e é a pessoa que nunca duvidou da minha capacidade em realizar meus sonhos.

E aos meus amigos e familiares, que estavam comigo nos momentos em que eu mais precisava de apoio e que me alegraram em momentos difíceis. Obrigada por todas as palavras de incentivo.

RESUMO

As primeiras pesquisas em tradução automática foram inauguradas no final da década de 1940, com a teoria de quebra de código linguístico do matemático Warren Weaver e, a partir dessa possibilidade, o tema vem sendo investigado até os dias atuais. Contudo, esses movimentos de pesquisa não foram contínuos e sofreram desencorajamentos, principalmente com o relatório do ALPAC (*Automatic Language Processing Advisory Committee*), publicado em 1966, que concluía que a tradução automática era imprecisa, cara, e andava a passos lentos. O relatório desacelerou as pesquisas por quase uma década, até que o tema voltou a ser investigado em meados de 1970, com pesquisas que se dividiram em dois eixos principais de: os estudos teóricos a respeito da Linguística no âmbito dos sistemas de tradução automática e os estudos que tratavam de testes computacionais, com abordagens empíricas, voltadas à construção desses sistemas (HUTCHINS, 2005, 2015; STUPIELLO, 2010). Com a evolução tecnológica e popularização da internet, essa subárea da Linguística Computacional ganhou notoriedade mundial e se tornou um ponto de investigação relevante para os Estudos da Tradução. Considerando a importância da produção científica sobre tradução automática para os Estudos da Tradução (GILE, 2015), este trabalho investigou, por meio de indicadores cienciométricos (VANTI, 2002; ESQUEDA, 2020), aliados a um viés qualitativo, a produção científica sobre tradução automática em cenário brasileiro. Com isso, demonstra quais discussões vêm sendo realizadas no país, nos últimos 30 anos – *i.e.*, de 1990 a 2020, e o impacto de tais discussões para a formação de tradutores brasileiros. Foram coletados todos os artigos de periódicos brasileiros dos Estudos da Tradução que tratavam do tema: “tradução automática”. Assim, essa busca nos títulos, resumos ou palavras-chave das publicações resultaram em 17 artigos ao longo de todas as edições de 14 periódicos verificados. Em virtude do número reduzido de publicações, o cenário de busca foi expandido e as mesmas palavras-chave foram submetidas à plataforma de busca Google Acadêmico, compondo um *corpus* de pesquisa com 87 publicações. Além da pesquisa quantitativa, que forneceu dados numéricos que determinaram os autores que mais publicaram sobre tradução automática no Brasil, seus respectivos coautores, as instituições às quais pertencem tais autores e os termos correlatos mais recorrentes nos resumos dessas publicações, a presente dissertação também levou em consideração o conteúdo tratado nesse *corpus* com relação à tradução automática na perspectiva brasileira.

Palavras-chave: Estudos da Tradução. Tradução Automática. Bibliometria. Cienciométrica.

ABSTRACT

The first research in machine translation was inaugurated in the late 1940s, with the theory of linguistic code breaking by the mathematician Warren Weaver, and from this possibility, the topic has been investigated to the present day. However, these research movements were not continuous and suffered discouragement, especially with the ALPAC (Automatic Language Processing Advisory Committee) report, published in 1966, which concluded that machine translation was imprecise, expensive, and moved at a slow pace. The report slowed down research for almost a decade, until the topic was investigated again in the mid-1970s, with research divided into two main axes: theoretical studies regarding linguistics in the context of machine translation systems and studies dealing with computational tests, with empirical approaches, aimed at building these systems (HUTCHINS, 2007, 2015; STUPIELLO, 2010). With the technological evolution and popularization of the Internet, this subarea of Computational Linguistics has gained worldwide notoriety and has become a relevant research point for Translation Studies. Considering the importance of the scientific production on machine translation for Translation Studies (GILE, 2015), this paper investigated, through scientometric indicators (VANTI, 2002; ESQUEDA, 2020), combined with a qualitative bias, the scientific production on machine translation in the Brazilian scenario. Thus, it demonstrates which discussions have been held in the country over the past 30 years - i.e., from 1990 to 2020, and the impact of such discussions on the training of Brazilian translators. We collected all articles from Brazilian Translation Studies journals that dealt with the topic: “machine translation”. Thus, this search in the titles, abstracts or keywords of the publications resulted in 17 articles throughout all the editions of 14 verified journals. Due to the reduced number of publications, the search scenario was expanded and the same keywords were submitted to the Google Academic search platform, composing a research *corpus* with 87 publications. In addition to the quantitative research, which provided numerical data that determined the authors who have most published on machine translation in Brazil, their respective coauthors, the institutions to which these authors belong to, and the most recurrent correlated terms in the abstracts of these publications, this dissertation also took into consideration the content dealt with in this *corpus* with regard to machine translation from the Brazilian perspective.

Keywords: Translation Studies. Machine Translation. Bibliometrics. Scientometrics.

LISTA DE FIGURAS

Figura 1- Representação gráfica das palavras-chave com maior densidade dentro do <i>corpus</i> “Tradução Automática”	32
Figura 2 - Mapa de densidade de produção no Brasil	41
Figura 3 - Mapa de visualização de densidade de produção dentro do <i>corpus</i> geral	48
Figura 4 - <i>Cluster</i> de coautoria	49
Figura 5 - Visualização de ligações dos autores	53
Figura 6 - Visualização por períodos de produção	57

LISTA DE GRÁFICOS

Gráfico 1 - Fonte de extração dos textos para compilação do <i>corpus</i> de pesquisa	30
Gráfico 2 - Número de publicações anuais no período de 30 anos	32
Gráfico 3 - Quantidade total de documentos por instituição de ensino.....	39
Gráfico 4 - Agências de financiamento que investiram em pesquisas sobre “Tradução Automática”, no Brasil, de 1990 a 2020	58
Gráfico 5 - Representação das categorias de discussões sobre a TA	68

LISTA DE TABELAS

Tabela 1 - Palavra-chave responsável pela extração do documento	30
Tabela 2 - Tipos de produção textual presente no <i>corpus</i> geral e suas quantidades	33
Tabela 3 - Tipos e quantidade de documentos de cada instituição de ensino	36
Tabela 4 - Número de documentos por área de conhecimento dentro do corpus geral	41
Tabela 5 - Tipos de documento por área de conhecimento e instituição de ensino	41
Tabela 6 - Autores que discutiram a respeito de TA no Brasil de 1990 a 2020	44
Tabela 7 - Relação de autores com área de pesquisa, instituição e coautoria	50
Tabela 8 - Número de coautorias por pesquisador	51
Tabela 9 - Número de orientações de teses e dissertações	54
Tabela 10 - Quantidade de financiamentos mencionados em cada tipo de documento por agência de financiamento	58
Tabela 11 - Quantidade de artigos por categoria de discussão	62
Tabela 12 – Produção das áreas de pesquisa para cada categoria de produção	66

LISTA DE QUADROS

Quadro 1 - Palavras-chave submetidas aos campos de busca dos periódicos	22
Quadro 2 - Diagrama metodológico da pesquisa.....	27
Quadro 3 - Duplicatas.....	34
Quadro 4 - Relação dos artigos categorizados como “TA e os Estudos da Tradução” e respectivos anos de publicação	69

LISTA DE ABREVIACES

FIVR/UNISEPE: Faculdades Integradas do Vale do Ribeira

ET: Estudos da Traduço

NILC: Ncleo Interinstitucional de Lingustica Computacional

PoP: Publish or Perish

PUC – RIO: Pontifcia Universidade Catlica do Rio de Janeiro

PUC – RS: Pontifcia Universidade Catlica do Rio Grande do Sul

PUC – SP: Pontifcia Universidade Catlica de So Paulo

SMT: Sistemas de Memria de Traduço

TA: Traduço Automtica

UECE: Universidade Estadual do Cear

UFC: Universidade Federal do Cear

UFJF: Universidade Federal de Juiz de Fora

UFPB: Universidade Federal da Paraba

UFRGS: Universidade Federal do Rio Grande de Sul

UFSC: Universidade Federal de Santa Catarina

UFSCAR: Universidade Federal de So Carlos

UNESP: Universidade Estadual Paulista “Jlio de Mesquita Filho”

UNIBERO: Centro Universitrio Ibero-Americano

UNIOESTE: Universidade Estadual do Oeste do Paran

USP - So Carlos: Universidade de So Paulo / Campus So Carlos

USP: Universidade de So Paulo

UTFPR: Universidade Tecnolgica Federal do Paran

SUMÁRIO

INTRODUÇÃO	09
CAPÍTULO 1: TRADUÇÃO AUTOMÁTICA - BREVES REGISTROS HISTÓRICOS, EVOLUÇÃO E BASES LINGUÍSTICAS DAS FERRAMENTAS	13
1.1 Entendendo as bases de construção de um sistema de TA	14
1.2 Além do Google Tradutor	16
CAPÍTULO 2: ASPECTOS METODOLÓGICOS	17
CAPÍTULO 3: ANÁLISE BIBLIOMÉTRICA E MAPAS CIENCIOMÉTRICOS SOBRE TRADUÇÃO AUTOMÁTICA EM CENÁRIO BRASILEIRO: resultados quantitativos	28
CAPÍTULO 4: RESULTADOS QUALITATIVOS: O QUE TRATAM AS PESQUISAS EM CENÁRIO BRASILEIRO SOBRE A TRADUÇÃO AUTOMÁTICA?	60
4.1 TA e o desenvolvimento, a arquitetura e/ou o funcionamento das ferramentas	62
4.2 TA e avaliação de performance e qualidade das ferramentas através de comparações linguísticas	63
4.3 TA e a análise, apresentação ou comparação de uma ou mais ferramentas de TA	64
4.4 TA e o uso e a recepção dos usuários	65
4.5 TA e o ensino de línguas estrangeiras	65
4.6 TA e os Estudos da Tradução	68
CONSIDERAÇÕES FINAIS	74
REFERÊNCIAS DOS PERIÓDICOS DO <i>CORPUS</i>	76
REFERÊNCIAS	78
APÊNDICE 1	84
APÊNDICE 2	85

INTRODUÇÃO

Meu interesse pela temática “Tradução Automática” surgiu quando ainda era aluna do curso de Bacharelado em Tradução da Universidade Federal de Uberlândia. Tive, durante a graduação, a oportunidade de participar de grupos de pesquisa nos quais discutíamos estudos baseados em Linguística de *Corpus*, que têm conexão direta com as tecnologias da tradução, tais como os sistemas de memória de tradução e tecnologias de tradução automática (doravante TA). Tais tecnologias são basicamente compostas por *corpora* paralelos, isto é, textos-fonte e textos-alvo de um determinado par linguístico.

Em minha monografia de final de curso de graduação, intitulada “Sistemas de memória de tradução e *Corpora*: integração que cria novas possibilidades ao tradutor” (ALVES, 2017), discuti a relação entre tais sistemas, a tradução automática e os *corpora*. Percebi a imediata relação entre os *corpora* paralelos alinhados, a construção de memórias de tradução e a arquitetura das tecnologias de tradução automática. Essa relação pôde, na monografia, ser confirmada a partir do que colocam Qun e Xiaojun (2015), de que a tradução automática é uma tecnologia que fornece ao seu usuário uma tradução para o idioma de chegada, fazendo uso concomitante de dicionários, cálculos algoritmos, bancos de dados de memórias de tradução e *corpora* paralelos.

Assim, após tal estudo, decidi aprofundar, em nível de mestrado, minhas investigações sobre tradução automática, principalmente por pensar que ela não é “tão automática assim”. Segundo Esqueda (2021), essa tecnologia utiliza avançadas análises linguísticas computacionais para processar documentos automaticamente e criar seus correspondentes em outras línguas sem a aparente intervenção humana, pois os *corpora* utilizados por empresas que a ela se dedicam são frutos de muito trabalho prévio e de atualização de tradutores humanos, embora também de vários outros profissionais, tais como engenheiros da computação, analistas de sistemas, linguísticas, apenas para citar alguns.

Ao produzir tanto minha monografia quanto meus estudos em nível de mestrado também percebi que a história da tradução automática surgiu juntamente com a evolução e há registros de seu surgimento que datam do ano de 1940, nos Estados Unidos, principalmente no pós-guerra. Se, nesta data, a abordagem principal utilizada para que essa tecnologia funcionasse bem era a abordagem baseada em regras gramaticais, que, como o nome já prevê, utilizava regras gramaticais para traduzir os textos a elas submetidos, as abordagens subsequentes foram trazendo sofisticação aos sistemas, com a implementação da abordagem estatística e, atualmente, a neural, as quais explicitaremos brevemente no Capítulo 1.

Além disso, se, a partir dos anos 1970, as tecnologias de tradução automática funcionavam através de *corpora* linguísticos gerais, com a evolução tecnológica dos anos de 1980 as pesquisas aumentaram e os pesquisadores perceberam que ao reduzirem o grau de automação e a complexidade do uso da língua, submetendo aos sistemas textos com afinidades temáticas, as tecnologias de TA seriam mais assertivas (MARTINS; VOLPE, 2005).

Hoje em dia, podemos constatar o surgimento de empresas que desenvolvem tais tecnologias direcionadas a um domínio específico (direito, turismo, medicina etc.) envolvendo determinados pares linguísticos. Esse é o caso, no exemplo citado por Esqueda (2021), do projeto TAUM (*Traduction Automatique de l'Université de Montréal – Machine Translation Project*) criado pela Universidade de Montreal, província de Quebec, no Canadá, em 1976 e que é utilizado até hoje. METEO é uma tecnologia de tradução automática desse projeto canadense que traduz a previsão do tempo no par linguístico inglês-francês. Essa tecnologia foi especificamente criada para informar os cidadãos anglófonos e francófonos a previsão do tempo e contém, em seu banco de dados, um vocabulário específico e limitado à sintaxe dessa temática. As tecnologias de TA, assim, podem ser gratuitas, gerais e abertas ao público geral, ou corporativas, direcionadas a um domínio específico e exclusivas de grandes empresas ou de governos.

As bases de dados que alimentam essas tecnologias já são consideradas produtos de grande potencial lucrativo no mercado de tradução e movimentam milhões de dólares anualmente. A TAUS - *Translation Automation User Society*, empresa que oferece dados e recursos para as indústrias da língua e da tradução, atua como mediadora dos compradores e das empresas provedoras de serviços e tecnologias linguísticas, e, além de oferecer uma gama abrangente de serviços, oferece também ferramentas de avaliação da qualidade da TA. A TAUS possui, hoje em dia, uma biblioteca de compartilhamento de dados contendo 21 tipos de *corpora* específicos (de domínio), incluindo bancos de dados relacionados à Covid-19 em mais de 30 pares linguísticos¹. As indústrias da língua e da tradução não veem a TA como uma ameaça e sim como uma oportunidade real de crescimento que atende as demandas da sociedade e integram os serviços dos tradutores aos produtos.

De um lado, a sociedade em geral utiliza a TA em contextos variados e seus produtos podem ter certa aceitação, como é o caso do Google Tradutor, da empresa norte americana Google. De outro, podemos verificar um interesse crescente em traduções automáticas para domínios específicos. No caso da área da saúde, por exemplo, a TA facilita a comunicação entre os agentes

¹ <https://www.taus.net/data-services/translation-data>

dessa área que falam línguas diferentes. Mas dificilmente em situações específicas uma tecnologia de TA com *corpora* gerais, como o Google Tradutor, resolveria de forma satisfatória os problemas e barreiras linguísticas. Segundo Roland (2020):

No domínio da saúde, onde a proteção de dados privados é extremamente importante, é fundamental que seja usado um sistema de TA seguro, e não uma ferramenta de tradução gratuita disponível na Internet, pois as traduções que utilizam esta última são lançadas na rede e podem ser acessadas por qualquer pessoa. (ROLAND, 2020, p.30 – tradução nossa)²

Em se tratando de ferramentas gratuitas, necessitamos entender que, de fato, não existem serviços gratuitos e que de algum modo estaremos oferecendo algum tipo de “pagamento” em troca desses serviços. Nós, tradutores, não podemos negar o fato de que o uso das ferramentas de tradução automática gratuita para serviços especializados de tradução compromete a privacidade de dados do cliente e do autor do texto.

Sobre tal comprometimento, em agosto de 2020, um tribunal polonês estabeleceu uma série de regras referentes ao uso de ferramentas gratuitas de tradução automática no contexto de direitos de propriedade intelectual, bem como à avaliação de sua qualidade³. Tudo isso porque um tradutor e a agência para a qual ele trabalhava submeteram a tradução de um livro inteiro a uma ferramenta de tradução automática, o Google Tradutor, e infringiram regras de direitos autorais, além de finalizarem o trabalho de pós-edição de forma incompleta e negligente.

Outro exemplo de uso inadequado das tecnologias de TA aconteceu no contexto legal. Nos Estados Unidos, por exemplo, o simples uso da TA para questões de comunicação durante uma abordagem policial pode desqualificar a ação policial e inocentar um suspeito, pois considera-se, neste caso, o produto da tradução ambíguo e confuso⁴.

Com os avanços das abordagens da tecnologia da tradução automática e das análises implementadas pela Linguística Computacional, podemos dizer que não apenas sua utilização cresce, para suprir as mais diversas necessidades das sociedades, mas também a necessidade de investigá-la. Compreender a TA e fazer uma avaliação apropriada de seus resultados e das pesquisas que a envolvem é essencial para uma integração de sucesso dessa tecnologia à sociedade.

² *In the healthcare domain, where protection of private data is extremely important, it is critical that a secure MT system be used, and not a free internet translation tool, as translations using the latter are indexed on the internet and searchable by anyone.*

³ <https://www.linkedin.com/pulse/lsp-responsibility-process-translation-rules-using-wojciech/?trackingId=8RqLU%2FPeSv6YTQ3bigF%2Fbg%3D%3D>

⁴ https://ecf.ksd.uscourts.gov/cgi-bin/show_public_doc?2017cr40100-24

Diante do exposto, indagamos:

- Como podemos definir a TA nos dias de hoje e o que se tem discutido sobre essa tecnologia, especialmente no cenário brasileiro? Como podemos avaliar sua integração à sociedade?

Como explicaremos adiante, trata-se de uma pesquisa descritiva que, por meio de técnicas bibliométricas e cienciométricas, buscará aferir o volume de trabalhos publicados em cenário brasileiro sobre o tema Tradução Automática, mapeando os polos de maior discussão, os autores que mais produziram e as áreas de pesquisa mais dedicadas à Tradução Automática. Nossa dissertação também buscará analisar o conteúdo desses trabalhos, com o propósito de chegarmos ao estado da questão. Para Therrien e Therrien,

A finalidade do “estado da questão” é de levar o pesquisador a registrar, a partir de um rigoroso levantamento bibliográfico, como se encontra o tema ou o objeto de sua investigação no estado atual da ciência ao seu alcance. Trata-se do momento por excelência que resulta na definição do objeto específico da investigação, dos objetivos da pesquisa, em suma, da delimitação do problema específico de pesquisa. (THERRIEN; THERRIEN, 2004, p. 7)

Podemos afirmar que esse tema de pesquisa nos parece relevante, pois como tradutores necessitamos compreender a TA e seu uso. Encontramos, com frequência, várias discussões acadêmicas dentro das bases de dados tais como o Google Acadêmico, *Web of Science*, *Scopus*, *Microsoft Academic*, que versam sobre a capacidade das tecnologias de TA, dissociadas, porém, daqueles que são seus maiores interessados: os tradutores. Em nossa opinião, essa pesquisa trará não apenas informações relevantes para mapearmos o desenvolvimento das discussões sobre a TA no Brasil, ou seja, o que os pesquisadores brasileiros discutiram e pesquisaram ao longo dos últimos 30 anos, mas também para averiguarmos que áreas foram responsáveis por tais investigações e, principalmente, para quem e com quem dialogam.

CAPÍTULO 1: TRADUÇÃO AUTOMÁTICA – BREVES REGISTROS HISTÓRICOS, EVOLUÇÃO E BASES LINGUÍSTICAS DAS FERRAMENTAS

Para falarmos de TA, precisamos, a nosso ver, primeiramente discutir seu significado. Segundo Hutchins e Somers (1992) os sistemas, motores ou ferramentas de tradução automática, que ficaram conhecidos pelos acrônimos MT, isto é, *Machine Translation*, em inglês, e TA ou Tradução Automática, em português, são programas computacionais que traduzem textos automaticamente sem a aparente necessidade de um agente humano.

O britânico John Hutchins (1939-2021), na qualidade de um dos pesquisadores que mais pesquisaram e publicaram sobre a TA no mundo, escreveu, em 2015, o artigo “*Machine Translation History of Research and Applications*”, publicado na *The Routledge Encyclopedia of Translation Technology*, no qual estabeleceu uma retrospectiva detalhada sobre a história da TA. Esse artigo, bem como outros artigos escritos pelo autor, constituíram a base para a escrita deste capítulo de nossa dissertação.

Apesar de a TA ser popularmente conhecida, principalmente em razão do Google Tradutor, essa tecnologia foi descoberta recentemente e as primeiras pesquisas sobre TA datam quase da mesma época do surgimento dos computadores.

Segundo Hutchins (2005; 2007; 2015), a ideia de se utilizar computadores para a tradução surgiu em 1947, assim que os computadores comerciais (e não mais para fins militares) começaram a ser desenvolvidos. A TA se espalhou rapidamente e foi a primeira função não numérica dos computadores.

Em 1954, nos Estados Unidos da América, dois anos após as primeiras descobertas, houve a primeira demonstração do funcionamento da TA. Embora ainda com um vocabulário e gramática muito restritos, a demonstração foi convincente o suficiente para impulsionar e especular os estudos sobre o tema. Já naquela época, surgiu algo que continua rondando as discussões sobre a TA – o medo de as máquinas substituírem os homens, o eterno paradoxo entre ajuda e ameaça. Muitos pesquisadores concluíram que a TA logo se desenvolveria e seria uma ameaça aos tradutores humanos e que todo texto poderia ser traduzido através do apertar de um botão.

No final da década de 1950, impulsionados pela Guerra Fria, o estudo da TA começa ser explorado pelo IBM em parceria com o setor de Tecnologia Estrangeira das Forças Armadas Americanas e com a Universidade de Georgetown, que instalou ferramentas de TA na Agência de Energia Atômica dos Estados Unidos. Com o avanço das pesquisas sobre TA voltadas para área de Inteligência Militar, o governo dos Estados Unidos cria o ALPAC – *Automatic Language*

Processing Advisory Committee, um comitê encarregado de avaliar o progresso da Linguística Computacional e, principalmente, da TA.

Em 1966, o ALPAC, após analisar os projetos desenvolvidos e executados até aquele momento nos EUA, publica um relatório que muda o cenário de expansão TA. No relatório, o grupo de pesquisadores se mostra cético perante a ferramenta e classifica os resultados como insatisfatórios, questiona a necessidade real desse sistema, além de enfatizar a falta de conhecimento sobre o que mais tarde viria a ser chamado de Linguística Computacional. O Governo Americano esperava uma tecnologia que acelerasse o processo tradutório e reduzisse os custos, entretanto, o relatório afirmava que os textos submetidos à ferramenta careciam de edição prévia e posterior. Isso justificou a redução drástica da maioria dos recursos governamentais ligados às pesquisas a respeito de TA e, mais que isso, desincentivou as pesquisas internacionais. As pesquisas que estavam a todo vapor naquelas últimas duas décadas foram estagnadas e, até 1975, nenhum projeto foi financiado pelo governo americano.

Apesar de o ALPAC ter interrompido as pesquisas por mais de uma década, países como Canadá, França e Alemanha seguiram com suas pesquisas, principalmente baseados na tradução de textos curtos e específicos. Nesse cenário, fora dos Estados Unidos, muitas das áreas do conhecimento contribuíram para a construção de sistemas de TA, tais como a Linguística Estrutural, Formal, Semiótica, Matemática, Linguística Quantitativa e a Linguística Computacional. Assim, ao longo da década de 1970, as pesquisas teóricas e práticas mostraram-se relevantes para a computação e o aprendizado das máquinas.

Com o avanço da informática e o surgimento de outras tecnologias nos anos de 1980, ressurge o interesse pela TA, também impulsionado pelo avanço das teorias linguísticas, como por exemplo, os estudos sobre gerativismo de Chomsky e, também, pelo surgimento da União Europeia. Na década de 1990, a revolução digital e acesso à *internet* transformaram as formas de comunicação em um mundo globalizado. Com esse panorama, a demanda por tradução cresceu exponencialmente e (re)alavancou as pesquisas à respeito de TA (HUTCHINS, 2015). Nesse período, várias ferramentas de TA foram construídas, algumas foram disponibilizadas *online* e gratuitamente e se popularizaram. Mas, nessa nova fase, como a TA foi arquitetada?

1.1 Entendendo as bases de construção de um sistema de TA

As tecnologias de TA podem compreender estruturas bilíngues (tradução entre duas línguas) ou multilíngues (mais que duas), que podem ser processadas das seguintes formas: por

meio de regras gramaticais (*ruled-based approach*), de sistemas baseados em *corpora* (TA estatística e TA baseada em exemplos) e por meio de redes neurais.

Como explica Hutchins (2015, p. 120) as primeiras pesquisas baseavam-se nas abordagens voltadas para o paradigma linguístico, guiadas pelas regras das línguas fonte e alvo. Essa abordagem é chamada de *tradução direta*, que tem como metodologia principal o sistema lexicográfico, ou seja a tradução do léxico fonte para o léxico alvo. A abordagem baseia-se no uso de dicionários, e é desenvolvida para fins de tradução de uma língua específica para outra, com pouca ou nenhuma reorganização sintática.

Outra abordagem que ganhou notoriedade após o relatório do ALPAC, principalmente com a evolução das pesquisas envolvendo a Linguística Computacional, Estatística e Ciência da Computação foi a *tradução interlíngua* (tradução para uma “língua artificial”), que utilizou o princípio matemático da Linguística Universal, princípio esse composto de códigos complexos com representações abstratas da linguagem. Assim, a tradução automática era feita em dois estágios: da língua-fonte para uma interlíngua e depois para a língua-alvo. O interesse por essa abordagem, vindo de países europeus, durou até meados dos anos de 1980. Os estudos e investimentos nessa abordagem perderam força após problemas como a rigidez dos níveis de análise, perda de informação entre os dois processos e a quantidade excessiva de análises parciais de textos. Todavia, podemos afirmar que a *tradução interlíngua* alavancou as pesquisas de TA estatística que surgiram posteriormente.

No início da década de 1990, com o avanço da Linguística de *Corpus*, o paradigma empírico ganhou destaque nas pesquisas sobre TA. Surge a abordagem de *tradução por transferência*, na qual a tradução é feita através de regras de transferência sintática. Ou seja, a ferramenta “aprendeu” a produzir uma tradução semelhante a algo presente em um determinado *corpus* da língua-alvo, através de padrões recorrentes presentes no *corpus* de treinamento. A *tradução por transferência* utiliza um sistema baseado em exemplos, que vai além da tradução de palavras isoladas e o resultado dessas traduções são um *corpus* paralelo bilíngue.

Outras abordagens que também seguem o paradigma empírico, e que são baseadas em *corpus*, são as *traduções baseadas em estatística* e, a mais recente delas, a *neural*. Os sistemas de TA baseados em estatística trabalham com um *corpora* bilíngue gigantesco procurando na língua-alvo ocorrências com a maior probabilidade possível. Essa abordagem difere daquela baseada em exemplos, pois não buscam por sentenças traduzidas previamente que estão presentes no *corpus*, mas avalia a probabilidade de tradução de palavras ou frases com base nesse *corpora* de treinamento gigantesco (CASELI, 2017). O sistema de TA que utiliza a abordagem baseada em

estatística normalmente é multilíngue, pois precisa de *corpora* paralelos multilíngues e não de regras linguísticas de todas elas.

Ainda segundo Caseli (2017), a abordagem estatística utilizada no Google Tradutor foi substituída, no ano de 2016, pela abordagem neural. Com relação às redes neurais, Caseli explica:

Redes neurais artificiais são técnicas computacionais usadas na área de Aprendizado de Máquina, as quais estão baseadas no modelo neural de organismos inteligentes, ou seja, em uma rede neural artificial uma unidade de processamento (um neurônio) recebe entradas, as processa e propaga a saída para outras unidades de processamento (outros neurônios) organizadas em camadas. Após os processamentos nas diversas camadas, um resultado é produzido na camada final e dado como resposta. (CASELI, 2017, p. 1785)

A então chamada TA neural é a abordagem mais recente. Diferentemente da TA estatística, a TA neural usa redes neurais com informações sobre a morfologia, frequência e ocorrências para gerar a tradução mais apropriada na língua-alvo. A TA neural otimiza toda a rede, gerando um produto mais robusto.

Toral e Way (2018) acreditam que a significativa melhora na qualidade das ferramentas de TA nos dias atuais, principalmente para textos mais complexos como os literários, se dá por dois fatores principais: o alto crescimento do mercado de livros eletrônicos – que disponibiliza uma variedade enorme de livros originais e suas traduções, que são grandes *corpora* bi- ou multilíngues, que então treinam as ferramentas baseadas em estatística; e o crescimento das pesquisas sobre TA neural.

Não obstante a todos esses fatores, não podemos nos esquecer também das deficiências dessas abordagens. Enquanto a *abordagem baseada em regras* depende de muito tempo e dedicação de pesquisadores para gerar as estruturas de aprendizagem, as *abordagens estatística e neural* dependem de *corpora* paralelos abundantes para que o sistema de TA “aprenda a traduzir”. Para Caseli (2017) e Toral e Way (2018), a *TA estatística* era considerada o estado-da-arte até o surgimento da *TA neural*, que tenta solucionar os obstáculos presentes na *TA estatística*, tais como, ambiguidade, incapacidade de generalizar, dificuldade com léxicos ou sintaxe complexos. Só a história dirá se tais obstáculos serão superados.

1.2 Além do Google Tradutor

Como antes mencionamos, as tecnologias de TA podem ser gratuitas, gerais e abertas ao público geral, ou corporativas, direcionadas a um domínio específico e exclusivas de grandes empresas ou de governos. Existem várias ferramentas de TA além do Google Tradutor, sendo

muitas delas criadas exclusivamente para atender necessidades específicas de um mercado e são extremamente evoluídas tecnologicamente e não são divulgadas para o público geral. Outras foram criadas ao longo da evolução das pesquisas sobre TA, possuem diversas abordagens diferentes e existem há bastante tempo.

O Systran é uma das ferramentas mencionadas por Hutchins (2015) que sobreviveu à queda das pesquisas após o relatório do ALPAC. Tendo surgindo na década de 1980, a ferramenta traduz para vários pares linguísticos. Mais recentemente, O Systran forneceu tecnologia para o Yahoo! Babel Fish e para a tradução automática presente no macOS.

Outra ferramenta mencionada por Hutchins no ressurgimento dos interesses mundiais pela TA no final da década de 1980 foi o METAL, um sistema de TA desenvolvido pela Siemens Nixdorf. Trata-se de uma ferramenta especializada em traduções técnicas baseada em regras, que traduz a partir do par linguístico inglês-alemão. Em nosso *corpus*⁵ foram recuperadas discussões interessantes sobre o METAL, datadas de 1996, sobre como alimentar a ferramenta com termos técnicos.

O Apertium é um exemplo de ferramenta ou sistema de TA baseada em regras, popularizado nos anos 2000. Foi desenvolvido pela Universidade de Alicante, na Espanha, e tem código-fonte e linguístico abertos, isto é, qualquer pesquisador ou usuário pode acessar a ferramenta. O Apertium é amplamente estudado por pesquisadores brasileiros do grupo NILC – Núcleo Interinstitucional de Linguística Computacional da USP.

O Unitran é uma ferramenta de tradução automática tipo *interlíngua*, que traduz dos e para os idiomas inglês, espanhol e alemão. Em um dos artigos de nosso *corpus*, a autora Mirna de Oliveira, em 2012, estudou parâmetros para incluir a língua portuguesa nessa ferramenta.

O DeepL, lançado em 2017, é uma das ferramentas mais modernas da atualidade e, segundo a página da ferramenta, tem atingido grandes resultados com sua ferramenta de tradução automática por rede neural. Segundo a empresa: “as inovações permitiram alcançar uma qualidade de tradução automática que ultrapassa todas as grandes empresas tecnológicas em nível mundial”. Para a empresa, a ferramenta é capaz de lidar até com jargões técnicos ou terminologia técnica, uma incapacidade de outras ferramentas que é criticada e enfatizada pelos pesquisadores.

O DeepL, assim como a tecnologia da TAUS, estão disponíveis no mercado e afirmam estar voltadas para o tradutor profissional e o mercado tradutório. O DeepL Pro e os *corpora* da TAUS podem ser acoplados a sistemas de memória de tradução. O DeepL Pro possui a opção de

⁵ Nesse trabalho não estamos usando *corpus* no sentido da Linguística de Corpus, pois não fazemos as análises linguísticas típicas da LC. Aqui chamamos de *corpus* um conjunto de documentos.

traduzir um documento inteiro mantendo os estilos do texto-fonte, imagens e outros elementos de formatação dos documentos originais. Além disso, ele possibilita a edição do documento após a tradução, aprimorando e capacitando ainda mais o sistema. Características como essas são interessantes para o tradutor profissional, que lida diariamente com vários tipos de documentos, desafios com a formatação e tradução de imagens, além da dificuldade de salvar uma memória de tradução (MT) após uma pós-edição.

As empresas que oferecem os sistemas de TA como produto especializado para tradução garantem conformidade com as leis de proteção de dados americana e da União Europeia e possuem certificação ISO 27001⁶, que certifica a proteção de dados submetidos a essas ferramentas.

Como anunciamos no título deste capítulo, buscamos traçar um breve histórico sobre o surgimento da TA, suas bases de construção e a menção a alguns dos mais relevantes sistemas de TA disponíveis, para além do popular Google Tradutor.

Passemos, na seção seguinte, aos aspectos metodológicos que nortearam a pesquisa.

⁶ Certificação de proteção de dados ISO 27001 - <https://www.iso.org/isoiec-27001-information-security.html>

CAPÍTULO 2: ASPECTOS METODOLÓGICOS

Antes de detalharmos a metodologia de pesquisa utilizada em nossa dissertação, pensamos ser importante aludirmos às principais contribuições dos estudos bibliométricos e cienciométricos.

A bibliometria e a cienciométrica tornaram-se importantes recursos de avaliação e análise da produtividade de pesquisadores, da cooperação entre universidades e instituições de pesquisa, dos impactos do financiamento estatal para a ciência e outros indicadores relacionados à divulgação científica.

Para Araújo e Alvarenga (2011), a bibliometria, quando aplicada com o objetivo de avaliar um campo científico, utiliza também métodos da cienciométrica. Utilizando métodos como análises em bases de dados gerais ou específicas, que reúnem produções científicas de um ou vários países, instituições ou temas, podemos mensurar a produção dos pesquisadores, de grupos de pesquisa, instituições e produzir um mapeamento das produções que auxiliam na averiguação das contribuições de determinada área e seus impactos na sociedade ao longo dos anos.

Segundo Vanti (2002), o termo cienciométrica alcançou notoriedade com o início da publicação da revista *Scientometrics*, em 1977, editada originalmente na Hungria e atualmente na Holanda. Para a autora, a cienciométrica baseia-se na aplicação de métodos quantitativos para estudar a história da ciência e do progresso científico e tecnológico, a saber:

- analisar as produções de uma área e delinear o crescimento de uma determinado campo de pesquisa;
- cartografar a ciência, apontando o avanço e as fronteiras de determinadas áreas da ciência;
- reunir pensamentos e olhares científicos de maneira imparcial; e
- demonstrar cientificamente as lacunas de determinada área de pesquisa.

As técnicas cienciométricas podem ser utilizadas para identificar as tendências e o crescimento do conhecimento de uma área; identificar as revistas científicas de maior impacto; mensurar o impacto das revistas secundárias; prever as tendências de publicação; estudar a dispersão e a obsolescência da literatura científica; prever a produtividade de autores individuais, organizações e países; medir o grau e padrões de colaboração entre autores; analisar os processos de citação e cocitação; determinar o desempenho dos sistemas de recuperação da informação;

avaliar os aspectos estatísticos da linguagem, das palavras e das frases; avaliar a circulação e uso de documentos em um centro de documentação e, principalmente, medir o crescimento de determinadas áreas e o surgimento de novos temas.

Usar as técnicas bibliométricas e cienciométricas para investigar, em nosso caso, a TA, possibilita delinear a evolução de suas pesquisas, identificarmos seus avanços e, talvez, retrocessos.

Para que uma pesquisa bibliométrica ou cienciométrica seja feita, o ponto de partida é estabelecer a base de dados que oferece boa abrangência e catalogação de estudos em uma determinada área, para que dela possam ser extraídas as informações desejadas pelo pesquisador.

Inicialmente, pensamos em selecionar apenas os periódicos acadêmicos dos Estudos da Tradução e investigar, em seus próprios portais, as publicações referentes à TA. Porém, uma busca preliminar feita nesses portais nos remeteu a um número reduzido de artigos sobre o tema, sendo necessária a inclusão da base de dados do Google Acadêmico para a expansão do *corpus*. Como pensamos que no Google Acadêmico poderíamos encontrar os mesmos artigos científicos sobre TA, já que vários periódicos dos Estudos da Tradução estão nele indexados, tornando, mais uma vez, o *corpus* reduzido, delimitamos que dissertações de mestrado e teses de doutorado também fariam parte de nosso estudo, ou seja, tanto os artigos encontrados nos periódicos quanto os outros documentos⁷ encontrados no Google Acadêmico foram contabilizados.

Assim, o objetivo geral desta dissertação é compreender o estado da arte (por meio de técnicas bibliométricas e cienciométricas) e o estado da questão (por meio de análise de conteúdo) das pesquisas dedicadas à TA em cenário brasileiro. Para operacionalizar tal objetivo geral, traçamos os seguintes objetivos específicos:

1. Mapear, nos periódicos brasileiros dos Estudos da Tradução, todos artigos científicos dedicados ao tema da TA, bem como na base de dados do Google Acadêmico, em suas páginas em português do Brasil, artigos, teses de doutorado e dissertações de mestrado; tais bases de dados foram escolhidas por buscarmos traçar o percurso da pesquisa sobre TA especificamente em nosso país;
2. Analisar os resultados que esses trabalhos alcançaram;
3. Descrever e problematizar os temas de pesquisa mais recorrentes no interior das discussões implementadas por esses trabalhos.

⁷ “Documentos” e “trabalhos” serão utilizados nesta dissertação de forma intercambiável.

Trata-se, portanto, de uma dissertação que lança mão de uma metodologia descritiva, que, a partir de técnicas bibliométricas e cienciométricas, compreende e avalia o impacto acadêmico dos trabalhos sobre TA para os Estudos da Tradução em cenário brasileiro.

Vale ressaltar que os periódicos científicos são bases de dados importantes para as pesquisas bibliométricas e cienciométricas, pois desempenham um papel fundamental na disseminação da informação, principalmente na divulgação das atividades científicas, de acordo com a área de pesquisa e abrangência determinadas por eles. Como afirmam Tenopir e King (2001, p. 8) “a informação contida nos periódicos se presta a muitas finalidades (pesquisa, ensino, serviços de alerta, leitura básica etc.) para os cientistas”, contribuindo, assim, para a propagação das descobertas acadêmicas e para o desenvolvimento de determinado campo científico. Sendo assim, o primeiro passo foi averiguar as publicações sobre TA em periódicos.

Na pesquisa, selecionamos periódicos reconhecidos de discussão de pesquisas na área dos Estudos da Tradução, geralmente vinculados a programas de graduação e pós-graduação. O objetivo das publicações presentes neles é o de veicular os resultados das pesquisas no campo dos Estudos da Tradução e da Interpretação no Brasil e no exterior, promover os debates mais relevantes da área, respeitando critérios específicos e temáticos para publicação dos trabalhos, tais como originalidade, idoneidade da pesquisa e do pesquisador e relevância para a área.

Foram selecionados, com base em Esqueda e Jesus (2015) e Esqueda (2020), 14 periódicos brasileiros de domínio público dos Estudos da Tradução, a partir dos quais efetuamos nossa análise:

1. Rónai: Revista de Estudos Clássicos e Tradutórios (UFJF)
2. Belas Infiéis (UNB)
3. Cadernos de Literatura em Tradução (USP)
4. Cadernos de Tradução (UFRGS)
5. Cadernos de Tradução (UFSC)
6. Cultura e Tradução (UFPB)
7. In-Traduções Revista do Programa de Pós-Graduação em Estudos da Tradução da UFSC
8. TRADTERM (USP)
9. Tradução e Comunicação: Revista Brasileira de Tradutores (UNIBERO/ANHANGUERA)
10. Tradução em Revista (PUC-RIO)

11. Traduzires (UNB)
12. Translatio (UFRGS)
13. Transversal – Revista em Tradução (UFC)
14. Scientia Traductionis (UFSC)

Os periódicos listados mantêm disponíveis, em seus portais, edições anteriores de suas publicações *online*.

Em todos os portais dos periódicos selecionados existem campos de busca que têm por objetivo facilitar a localização de artigos presentes em toda coleção e que podem ser encontrados por alguma expressão de busca ou palavra-chave. As palavras-chave submetidas aos campos de busca⁸ em cada um dos periódicos brasileiros foram pensadas para abranger todos os possíveis descritores similares para o termo “tradução automática”, e foram inseridos no campo de busca dos portais de cada um dos periódicos utilizando o indicador booleano⁹ “OR”: Tradução Automática; Tradução Eletrônica; Tradução de Máquinas e Tradução Mecânica. Todas as quatro palavras-chave foram testadas com suas variantes - feminino, masculino e plural, o que resultou num total de 16 variantes submetidas à busca e separadas pelo operador booleano “OR”, como exemplifica o Quadro 1.

Quadro 1 - Palavras-chave submetidas aos campos de busca dos periódicos

<ul style="list-style-type: none"> – Tradução Automática OR Tradutor Automático OR Traduções Automáticas OR Tradutores Automáticos – Tradução Eletrônica OR Tradutor Eletrônico OR Traduções Eletrônicas OR Tradutores Eletrônicos – Tradução de Máquina OR Tradutor de Máquina OR Traduções de Máquinas OR Tradutores de Máquina – Tradução Mecânica OR Tradutor Mecânico OR Traduções Mecânicas OR Tradutores Mecânicos

Fonte: a autora.

Nem sempre os campos de busca funcionaram como o esperado, ora o operador booleano não funcionou e as palavras-chave tiveram que ser inseridas uma a uma, em outros casos o

⁸ O campo de busca das revistas, quando acionados pelas expressões de busca ou palavras-chave, verificam a incidência do termo pesquisado percorrendo os campos: título, resumo e palavras-chave.

⁹ Os Operadores Booleanos agem como códigos que informam ao sistema de busca como combinar os termos de sua pesquisa. Eles são: AND, OR e NOT e é importante que estejam em letras maiúsculas para serem identificados na hora da busca. O operador booleano AND funciona como a palavra “E”, capturando apenas artigos que contenham todas as palavras-chave digitadas, restringindo a varredura da pesquisa. O OR é utilizado para encontrar artigos que contenham um dos termos listados e NOT exclui documentos que contenham as palavras listadas na busca.

buscador “escondia” muitos resultados ou mostrava resultados não relacionados. Esses foram os primeiros desafios da pesquisa, pois apesar de existir a técnica bibliométrica para executar a pesquisa, na prática, muitos buscadores não funcionam como deveriam, além do uso equivocado das palavras-chave por parte de alguns pesquisadores.

No periódico *Tradução em Revista*, por exemplo, foi necessário verificar manualmente as 20 edições disponíveis *online*, pois o campo de busca não era confiável¹⁰. No periódico *Cadernos de Literatura e Tradução*, o buscador identificou quatro artigos, mas nenhum deles discutia ou tinha presente as palavras-chave mencionadas. No periódico *Cadernos de Tradução (UFRGS)*, o buscador não identificou quaisquer dados e as 12 edições foram verificadas manualmente. Já no periódico *Cadernos de Tradução (UFSC)*, o buscador revelou resultados não relacionados ao tema, que precisaram ser verificados manualmente. Na pesquisa do portal do periódico *TradTerm*, verificamos 36 edições, cujos dados pareciam promissores, pois identificamos 69 resultados relacionados ao tema, porém apenas dois dos resultados foram relevantes, após serem analisados caso a caso.

Encontramos, por fim, a soma de 17 artigos publicados, cujo primeiro artigo é datado de 1996 e o último 2020, ou seja, o primeiro artigo publicado no Brasil em periódicos especializados foi há 26 anos, mesmo as discussões terem começado internacionalmente entre as décadas de 1940 e 1950.

Como o número de trabalhos extraídos dos periódicos foi bastante reduzido para um estudo bibliométrico e cienciométrico, expandimos a busca para o banco de dados do Google Acadêmico, como antes mencionamos.

O Google Acadêmico é uma plataforma *online* e gratuita¹¹ de busca que reúne, em um banco de dados, documentos acadêmicos indexados a partir de páginas com conteúdo acadêmico, tais como periódicos, cadernos, revistas e repositórios, e os fornece através de lista simples em sua página da internet, porém não disponibiliza uma visão geral do número de documentos recuperados, sua fonte de publicação, o link para *download*, data de publicação, dentre outras informações. Utilizando simplesmente o Google Acadêmico, sem uma ferramenta de pesquisa que facilita o acesso a esses dados, seria necessário abrir e fazer o *download* de cada documento selecionado individualmente.

¹⁰ Para testar o campo de busca, inserimos o termo “tradução”, um termo genérico que logicamente estaria presente em algum documento presente em um periódico especializado em Estudos da Tradução. Se pelos menos um resultado fosse apresentado, consideramos que o buscador estava em funcionamento.

¹¹ A *Web of Science* e o *Scopus* também são exemplos de base de dados robustas e reconhecidas, porém, não são gratuitas.

Há um número considerável de ferramentas de pesquisa que pode ser atrelado ao Google Acadêmico para facilitar a busca e o acesso aos dados bibliográficos. Porém, para fins de análise bibliométrica e cienciométrica e elaboração de mapas cienciométricos, utilizamos, nesta dissertação, o software Publish or Perish (PoP).

O PoP é uma ferramenta de acesso livre e usabilidade simples, que permite o *download* de dados bibliográficos, funcionando como uma interface de busca do Google Acadêmico. A pesquisadora Anne-Wil Harzing (2010), desenvolvedora do PoP, descreve o programa como “ideal para mesurar o impacto das publicações acadêmicas com os dados do Google Acadêmico”¹² [tradução nossa]. Ao selecionarmos a base de dados Google Acadêmico e ao submetermos uma palavra-chave no campo de busca do PoP, a ferramenta recupera uma seleção de referências a ela relacionadas presentes nessa base, e fornece estatísticas referentes ao impacto acadêmico das publicações, produtividade dos autores, o ano de publicação, veículo da publicação e ainda fornece acesso ao documento. Além disso, o PoP pode analisar e incluir os registros de citações, bem como disponibiliza o documento para *download* na própria interface da ferramenta.

Portanto, fizemos a busca utilizando o PoP, que facilitou a varredura das publicações sobre TA. Entretanto, foi difícil determinar as publicações específicas do cenário brasileiro, pois o programa encontrou diversos tipos de documentos escritos em português, tanto do Brasil quanto de Portugal e de outros países de fala portuguesa. Assim, 140 documentos foram examinados manualmente, incluindo: apresentações em .pptx, artigos, resumos, dissertações, dossiês, ensaios, monografias, relatórios, reportagens e teses.

Dentre eles, 12 artigos que já haviam sido selecionados nos periódicos brasileiros dos Estudos da Tradução se repetiram nessa busca. Essas duplicatas foram então excluídas. Dezoito (18) documentos que discutiam o tema da Tradução Automática para Língua de Sinais foram excluídos, pois pensamos que esse tema poderá ser objeto de uma investigação relacionada especificamente a essa língua. Vinte (20) documentos de publicações portuguesas e 2 publicações¹³ de outros países, uma da Universidade de Macau, na China, e outra da Universidade de Santiago de Compostela, na Espanha, também foram excluídas do *corpus* de pesquisa, pois nosso foco principal de investigação é o cenário brasileiro.

¹² No original: “*Publish or Perish is ideally suited to measure the impact of academics and journals with Google Scholar data.*” (HARZING, 2010)

¹³ Ressaltamos que três artigos presentes no *corpus* dos periódicos brasileiros foram escritos por autores da Bélgica, Portugal e Alemanha. Esses artigos foram mantidos no *corpus*, já que não identificamos se seus autores estavam ou não trabalhando no Brasil.

Além dos documentos excluídos, havia três outros artigos não relacionados com nosso tema, por exemplo, uma dissertação e um artigo sobre tradução automática de problemas de escalonamento dentro da área de Engenharia Elétrica, outro de tradução automática de voz, que pensamos fazer parte de estudos da área de Interpretação, e outro sobre estudos relacionados à acessibilidade e tecnologias assistivas.

Ou seja, fez-se necessária não somente uma avaliação dos títulos, mas também uma leitura dos resumos desses documentos, para obtermos a certeza de que as produções estivessem, de fato, de acordo com o escopo da pesquisa, a saber: documentos que contivessem as palavras-chave definidas, que tivessem sido publicados em território brasileiro, fossem produções científicas (excluindo resumos, questionários, resumos de evento e slides de apresentações em .pptx (totalizando 11¹⁴ exclusões de documentos em formato incompatível) e principalmente, tratassem de Tradução Automática de textos. Mugnaini e Strehl (2008) afirmaram que resultados divergentes em pesquisas no Google Acadêmico podem ser encontrados, pois são gerados por meio de diferentes metodologias. Noruzi (2005) também ressalta o problema quando se trata de acessar um banco de dados tão abrangente:

O *Google Acadêmico*, às vezes, tem suas desvantagens. Ele busca resultados que nem sempre podem ser considerados materiais acadêmicos pelo ponto de vista tradicional, tais como, notas administrativas, visitas guiadas de bibliotecas, apostilas etc. Infelizmente, o algoritmo do *Google Acadêmico* não consegue distinguir esses materiais. Além de muitos periódicos e revistas acadêmicas não estarem indexadas à plataforma. (NORUZI, 2005, p. 174, tradução nossa)¹⁵

Depois de submeter esses 140 resultados a uma análise manual e individual, com critérios rigorosos de seleção, 70 documentos foram encontrados e adicionados ao *corpus* geral de pesquisa (junto aos outros 17 artigos encontrados nos periódicos). Percebemos um espaço temporal de 30 anos, isto é, de 1990 (primeiro documento apresentado pelas bases de dados) a 2020.

Assim, somando-se a coleta de documentos com o tema “Tradução Automática” em 14 periódicos brasileiros dos Estudos da Tradução, que resultou em 17 documentos, com a expansão de busca à base de dados do Google Acadêmico através do software PoP, que resultou em 70

¹⁴ Total de exclusões: 12 duplicatas, quatro duplicatas dentro do PoP, 18 documentos sobre TA de Sinais, 22 documentos de publicação fora do Brasil, três documentos de outras áreas de conhecimento, 11 documentos de formato incompatível – 70 documentos removidos dos documentos iniciais recuperados pelo PoP na base de dados *Google Acadêmico*.

¹⁵ No original: “*Google Scholar is, however, not without its disadvantages, sometimes, Google Scholar includes administrative notes, library tours, student handbooks, etc., which are not exactly scholarly material from the point of view of the traditional definition of scholarly information. Unfortunately, Google Scholar’s algorithm cannot distinguish between articles, editorial notes or library guides. Many scholarly periodicals and magazines are not indexed.*”

documentos (artigos, monografias, dissertações de mestrado e teses de doutorado), compilamos um *corpus* com 87 documentos, que foram salvos e seus metadados transferidos para uma tabela do Excel. A tabela e os 87 documentos podem ser acessados através do endereço eletrônico disponibilizado no apêndice 1. Pela quantidade de arquivos, a relação completa dos 87 artigos está no apêndice 2 e não nas referências desse trabalho.

Cada um dos 87 documentos foram analisados e receberam tratamento bibliométrico, reportados e representados através de tabelas e gráficos com os seguintes campos:

- Nome do documento no *corpus*;
- Área de conhecimento;
- Base de dados em que o documento foi encontrado;
- Link do documento;
- Tipo de Publicação;
- Título da Publicação;
- Palavra-chave de busca;
- Nome do autor(a);
- País de origem do autor(a);
- Nome do(s) coautor(es);
- Nome do orientador(a) (caso houvesse);
- Instituição de Ensino geradora do trabalho – Polo educacional;
- Ano de escrita;
- Ano de publicação;
- Idioma da publicação; e
- Agência de Financiamento.

Esses dados foram obtidos um a um pela análise manual dos documentos, transferidos para uma tabela do Excel, que gerou os gráficos e tabelas que serviram de base para nosso estudo bibliométrico e para nosso mapeamento cienciométrico, que serão discutidos nas próximas seções.

Já os dados cienciométricos foram reorganizados através do organizador de referências Mendeley. Essa ferramenta é um gerenciador de referências que possibilita o armazenamento, a organização e o compartilhamento de referências e dados de pesquisa (MENDELEY, 2022). Ele gera automaticamente as referências bibliográficas, uma lista de referências que podem ser exportadas para ferramentas de construção e visualização de ligações bibliométricas, no caso, o software VOSviewer. Portanto, o Mendeley organiza as informações bibliométricas e o VOSviewer as representa por mapas de relações.

O software VOSviewer é baseado no procedimento de visualização de similaridades que cria rede de ligações entre os documentos submetidos à análise. O objetivo principal dessa ferramenta é fornecer uma visualização na qual objetos são localizados de modo que a distância entre os pares reflita a semelhança entre eles com a maior precisão possível. Essas redes de

semelhança podem, por exemplo, incluir periódicos, pesquisadores ou publicações individuais, e podem ser construídas com base em relações de citação, acoplamento bibliográfico, cocitação ou coautoria (VOSVIEWER, 2022). O procedimento de criação das relações entre autores dessa pesquisa foi realizado através do acoplamento bibliográfico gerado em formato RIS, que quando submetidos ao VOSviewer fornecem as relações cienciométricas presentes no *corpus*.

A mineração, que é utilizada para construir e visualizar redes de coocorrência de termos importantes extraídos de um *corpus* também é feita pelo VOSviewer, e com ele podemos determinar os termos mais recorrentes dentro de uma área de conhecimento. Assim, o diagrama metodológico a seguir mostra visualmente o fluxo de todos os processos metodológicos que ocorreram na pesquisa.

Quadro 2 - Diagrama metodológico da pesquisa



Fonte: a autora

CAPÍTULO 3: ANÁLISE BIBLIOMÉTRICA E MAPAS CIENCIOMÉTRICOS SOBRE TRADUÇÃO AUTOMÁTICA EM CENÁRIO BRASILEIRO: resultados quantitativos

Após a coleta dos dados e a inserção dos dados em planilhas, desenvolvemos um processo de análise bibliométrica, relativo às informações relacionadas à publicação, isto é, seus metadados: título, resumo, palavras-chave, autor ou autores, ano de publicação, título da publicação etc. Essas informações bibliométricas, aliadas ao impacto da publicação, à instituição onde foi desenvolvida, o fomento que obteve, a rede de autores que a citam, dentre outros, expandem a análise bibliométrica, o que Vanti (2002) chama de análise cienciométrica, pois mensura os impactos da ciência, para além daquela determinada publicação.

Vários indicadores bibliométricos e cienciométricos podem ser extraídos dos dados obtidos: o número de publicações dentro do banco de dados, áreas de pesquisa que mais publicam sobre o tema, tipo de publicação mais comum, os autores que publicam mais, as ligações de coautoria, os orientadores que mais incentivam pesquisas naquela área, os polos/ instituições que mais publicam sobre aquele tema, picos de publicação, espaço temporal de pesquisa daquele tema, o quanto os centros e agências de fomento incentivam as pesquisas sobre o tema, a nacionalidade dos pesquisadores, os idiomas de publicação mais recorrentes, o número de citações, a densidade de produção de um autor em relação ao período de publicação, as palavras-chaves mais recorrentes dentro do *corpus*, entre outras observações que podem variar de acordo com o escopo da análise. Essas informações são usadas para delimitar o impacto causado pelas pesquisas dentro da área de conhecimento.

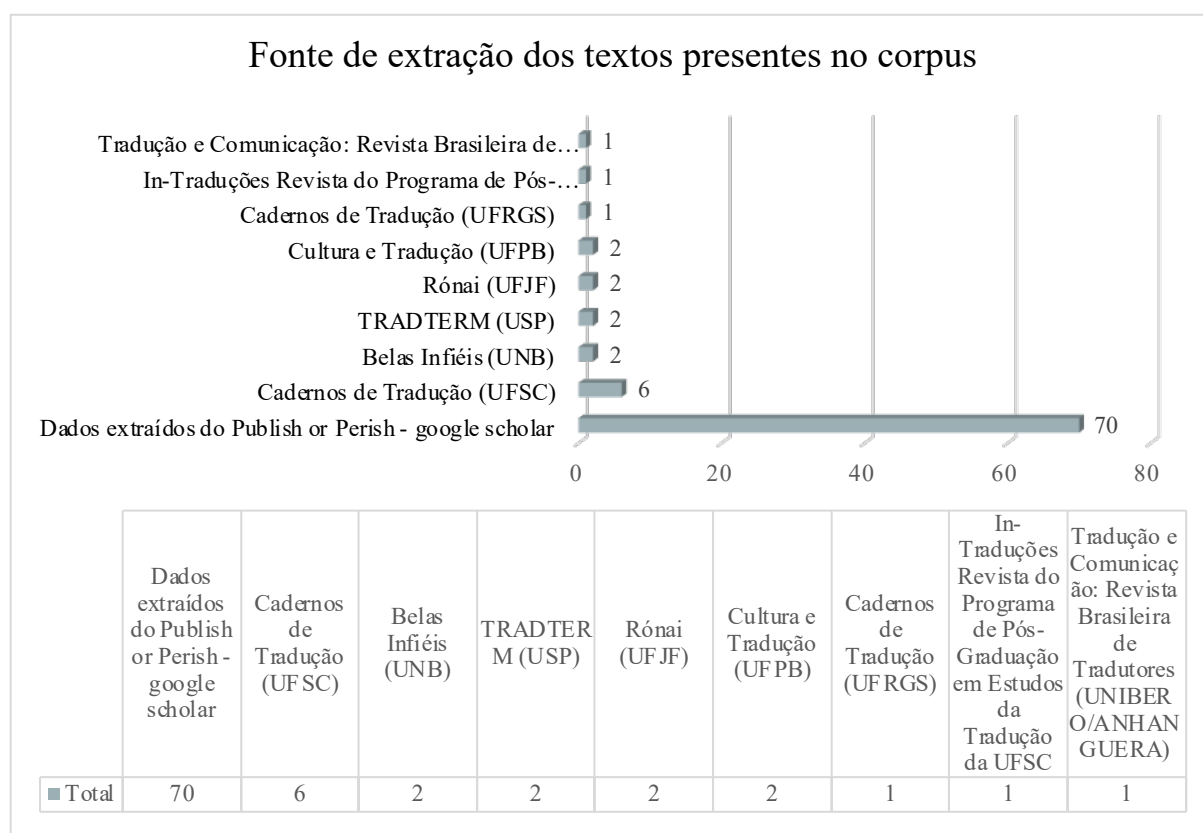
O enriquecimento e expansão da análise bibliométrica são feitos através de mapas cienciométricos, que representam a relação entre diferentes autores, instituições, ano de publicação, densidade de publicação etc. Um mapa cienciométrico é uma representação espacial e visual de uma série de ligações/relações bibliométricas (ECK, 2011), ou seja, de que maneira as áreas do conhecimento, as especialidades das áreas e os autores se relacionam, utilizando dados concretos para representar graficamente as estruturas sociais da ciência.

Buscando mapear o crescimento e o impacto das publicações a respeito da TA em cenário brasileiro, nos propomos a averiguar uma série de indicadores bibliométricos e cienciométricos a partir do *corpus* geral com 87 documentos.

Partindo desse objetivo, o primeiro grande desafio da pesquisa surgiu quando foi encontrado um número bastante reduzido de artigos tratando de TA nos periódicos brasileiros dos Estudos da Tradução. Havia uma pressuposição nossa de que os números não seriam muito

animadores, pois as pesquisas em TA são, em sua maioria, feitas em outras áreas do conhecimento, como Ciência da Computação. Assim, a quantidade reduzida de artigos encontrados nos periódicos brasileiros dos Estudos da Tradução confirmou essa pressuposição e ainda indicaram que as pesquisas em TA não atingem um de seus mais interessados grupos de profissionais: os tradutores. O gráfico 1 mostra as bases de dados de onde cada documento presente no *corpus* foi retirado.

Gráfico 1 - Fonte de Extração dos textos para compilação do *corpus* de pesquisa



Fonte: a autora.

À exceção do periódico “Cadernos de Tradução”, da Universidade Federal de Santa Catarina, que contribuiu com seis artigos sobre TA, em suas 57 edições, desde 1996, a investigação conduzida nos outros 13 periódicos evidenciou a exígua presença de artigos publicados sobre TA. Apenas 11 artigos foram encontrados em todos os outros periódicos analisados, reflexo do número de pesquisas acadêmicas ainda bastante reduzido sobre o tema.

A base de dados do Google Acadêmico, se mostrou, por sua vez, a opção mais abrangente para identificar as principais publicações sobre TA que respeitassem o escopo da pesquisa: palavras-chave em língua portuguesa e artigos que tivessem sido publicados no Brasil. Ainda que

excessivamente abrangente, ou seja, recuperando documentos que não são relevantes para a pesquisa, tais como apresentações em .pptx, relatórios de bolsa, pesquisas relacionadas com outras áreas de abrangência e resumo de apresentações, a base de dados expandiu nosso *corpus* de pesquisa, na qual obtivemos 70 documentos, dentre eles artigos, relatórios, monografias, dissertações de mestrado e teses de doutorado, cujas fontes e informações detalharemos a seguir.

Para obter dados confiáveis e garantir que o máximo de documentos possíveis relacionados à TA fossem detectados, quatro palavras-chave e todas as suas variantes foram submetidas aos buscadores, tanto nos periódicos quanto no banco de dados do Google Acadêmico, com o propósito de abranger a maior variação terminológica possível dentro dos títulos, resumos e palavras-chave.

A Tabela 1 mostra a quantidade de arquivos encontrados a partir de cada palavra-chave ou sua variante dentro dos periódicos ou banco de dados do Google Acadêmico. Setenta e um (71) documentos foram recuperados através da palavra-chave “tradução automática”. Sendo assim, se as variantes não tivessem sido submetidas aos buscadores, haveria uma perda de 16 documentos (18,40% do *corpus*).

Tabela 1 - Palavra-chave responsável pela recuperação do documento

Palavra-chave	N. de Documentos recuperados
Tradução Automática	71
Tradutor Automático	5
Tradutores Automáticos	4
Traduções Automáticas	3
Tradutor Eletrônico	2
Tradução Eletrônica	2
Total Geral	87

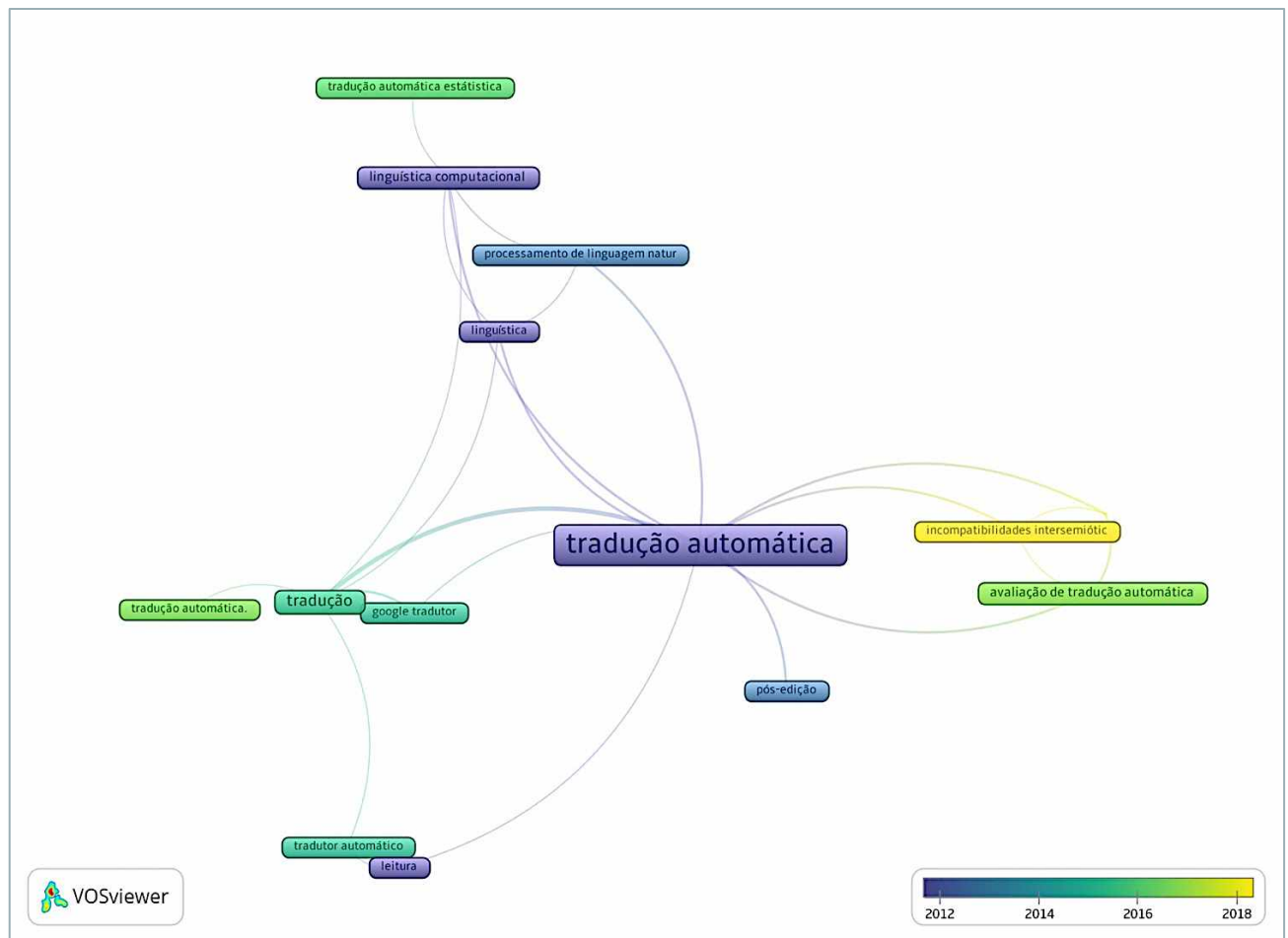
Fonte: a autora.

Uma das técnicas principais das análises bibliométricas e cienciométricas é determinar as palavras-chave da área de conhecimento, para que o mapeamento seja confiável e represente aquela área sem perda de pesquisas científicas no momento das buscas.

Para verificar as palavras com maior número de ocorrência dentro dos textos, um mapa de densidade foi criado no VOSviewer, usando os 87 documentos. A ferramenta procura similaridades entre os títulos, resumos e palavras-chave, delineando ligações entre termos principais. Na análise bibliométrica e ciencimétrica, essas ligações são determinadas pela “Lei de Zipf”, também conhecida como “Lei do Mínimo Esforço”, que consiste em medir a frequência do aparecimento das palavras em vários textos, gerando uma lista ordenada de termos de uma

determinada disciplina ou assunto” (VANTI, 2002, p.153). Desse modo, é possível perceber a relação e evolução das subáreas pesquisadas dentro daquela grande área através das palavras-chave, conforme exhibe a Figura 1..

Figura 1 - Representação gráfica das palavras-chave com maior ocorrência dentro do *corpus* “Tradução Automática”



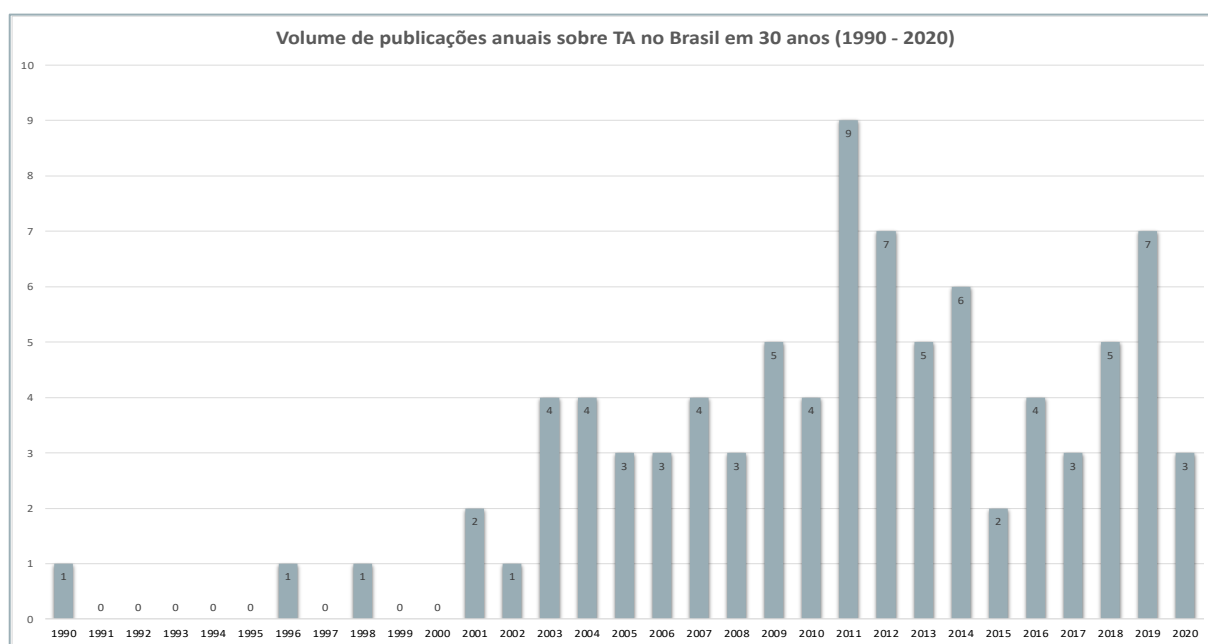
Fonte: a autora.

A Figura 1 mostra a ocorrência das palavras-chave no *corpus* e sua temporalidade, ou seja, em que momento aqueles tópicos relacionados à palavra-chave foram mais discutidos e que relação eles têm com o tema principal. A palavra mais frequente no *corpus*, geralmente, tem maior relação com o documento e tem destaque na imagem. É possível observar que todas as outras palavras de destaque têm relação com “tradução automática”. A partir desse ponto, pode-se ver ramificações de subáreas, como por exemplo, TA e pós-edição – que estava sendo discutido, principalmente em 2013/2014, ou TA relacionada com Linguística – Linguística Computacional e partindo daí, Processamento de Linguagem Natural, mais discutidos no início da década de 2010.

Quando a TA está relacionada à ferramenta em si, logo aparece o Google Tradutor nas pesquisas – entre 2015/2016.

Na pesquisa, encontramos um espaço temporal de 30 anos, ou seja, de 1990 (data primeira publicação recuperada pelas bases) a 2020, conforme exhibe o Gráfico 2. A pesquisa não abordou resultados após julho de 2020, por entender que os efeitos da pandemia podem ter causado diminuição ou interrupção das publicações e pela dificuldade da base de dados em indexar documentos que acabaram de ser publicados. Como nos coloca Vanti (2002), publicar pesquisas leva tempo e, qualquer interrupção passa a ter impacto na produção dos pesquisadores. Outros estudos precisaram ser feitos para verificar se de fato houve um pico ou uma queda no número de trabalhos discutindo TA após o período de pandemia.

Gráfico 2 - Número de publicações anuais no período de 30 anos



Fonte: a autora.

No ano de 2011, nove trabalhos foram publicados, um número significativo dentro do universo analisado, representando um pico de publicação da TA, e a quantidade média de publicações se manteve maior que os anos anteriores a 2011. Em 20 anos, de 1990 – data do primeiro artigo recuperado pelas bases de dados, até 2010, 36 documentos foram encontrados (41,4% do total do *corpus*), uma média de 1,8 documentos produzidos por ano, considerando que na década de 90 quase não houve discussões sobre TA no Brasil. Após 2010, as pesquisas aumentaram, juntamente com as discussões mundiais a respeito da TA, embora ainda oscilante, o número de produções cresceu consideravelmente em relação as duas décadas anteriores.

Nos últimos 10 anos, de 2011 a 2020, 51 documentos foram recuperados (58,6% do *corpus* total), totalizando uma média de 5,1 documentos produzidos por ano. Esse crescimento nas publicações justifica-se pela popularização dos tradutores automáticos, das pesquisas sobre pós-edição, do surgimento de novos cursos de Tradução, de programas de pós-graduação voltados aos Estudos da Tradução e, principalmente, com o crescimento da interdisciplinaridade entre programas de pós-graduação nas universidades brasileiras.

Ao realizarmos as análises bibliométricas e cienciométricas foi necessário definirmos a metodologia da forma mais abrangente possível e respeitarmos o escopo estipulado para análise. Cada detalhe presente no *corpus* pode revelar muito a respeito das tendências de pesquisa da área estudada. Um exemplo disso são os tipos de documentos presentes no *corpus*. Um número elevado de artigos dentro *corpus* demonstra uma tendência da área de pesquisa em contribuir e disseminar conhecimento, criando um diálogo entre os pesquisadores da área. Já as monografias, dissertações e teses mostram a inserção e evolução de novos pesquisadores para aquela área de conhecimento, ilustrando um crescimento do campo de pesquisa daquele tema e um cenário promissor no futuro, além de apontar a influência dos pesquisadores orientadores, que através de seus orientandos, propagam e estimulam o crescimento naquela área. Relatórios e dossiês caracterizam as pesquisas que estão em desenvolvimento e são campos férteis para uma análise qualitativa dos desafios práticos que os pesquisadores vêm enfrentando. Nesse sentido, a Tabela 2 a seguir mostra os tipos de documentos presentes no *corpus*.

Tabela 2 - Tipos de produção textual presente no *corpus* geral e suas quantidades

Tipos de produção textual presente no <i>corpus</i> geral	N. de documentos do tipo
Artigo científico	55
Dissertação de mestrado	12
Tese de doutorado	7
Relatório de pesquisa	6
Monografias de final de curso	4
Dossiê científico	1
Ensaio crítico	1
Reportagem	1
Total Geral	87

Fonte: a autora.

O artigo científico é um tipo de produção textual científico-acadêmica que tem como característica principal discutir inovações, métodos, processos, análises, técnicas e resultados, com a finalidade de disseminar conhecimento e promover discussões dentro de uma determinada área do conhecimento. 63,2% dos documentos que compõem nosso *corpus* geral são artigos, o que

demonstra o interesse dos pesquisadores em comunicar seus resultados e dialogar cientificamente uns com os outros. Dentre os 55 artigos, oito foram publicados por pesquisadores da USP-São Carlos e sete da UFSC e quatro pelo grupo NILC – Núcleo Interinstitucional de Linguística Computacional, formado pela Universidade de São Paulo em São Carlos, pela Universidade Federal de São Carlos (UFScar) e pela Universidade Estadual Paulista (UNESP), entre outros.

Vale lembrar que ocorreram, como já anunciamos, duplicatas, isto é, 12 artigos encontrados nos periódicos também apareceram na base de dados do Google Acadêmico, através do PoP e foram retirados do *corpus*. O Quadro 3 mostra os artigos duplicados.

Quadro 3 - Duplicatas: documentos presentes tanto nos periódicos quanto na base da dados

Nome da Revista	Título do Artigo	Nome da Periódico apareceu no PoP?
Belas Infieis (UnB)	<i>Uma Breve Reflexão Historiográfica Sobre A Tradução Automática</i>	Não, aparece como repositório da UnB
Cadernos de Tradução (UFRGS)	<i>Um Estudo Comparativo Da Desambiguação Lexical Realizada Por Tradutor Eletrônico e Tradutores Humanos</i>	Sim, foi retirado do periódico
Cadernos de Tradução (UFSC)	<i>El general de brigada es un tipo de caramelo – tradução automática e aprendizagem cultural</i>	Sim, foi retirado do periódico
Cadernos de Tradução (UFSC)	<i>Tradução Automática: Uma Ferramenta de Auxílio ao Tradutor</i>	Sim, foi retirado do periódico
Cadernos de Tradução (UFSC)	<i>Um estudo de expressões cristalizadas do tipo V+SN e sua inclusão em um tradutor automático bilíngüe (português/inglês)</i>	Sim, foi retirado do periódico
Cultura e Tradução (UFPB)	<i>Por Uma Estética Na Tradução Automática: Um Estudo Baseado Em Corpora Eletrônicos</i>	Não, apareceu dentro de repositório
Cultura e Tradução (UFPB)	<i>Coleta, etiquetagem e anotação de incompatibilidades intersemióticas geradas por tradução automática</i>	Não, o artigo apareceu no PoP dentro de repositório
In-Traduções Revista do Programa de Pós-Graduação em Estudos da Tradução (UFSC)	<i>Um panorama do fluxo de recepção da tradução automática no cenário nacional</i>	Não, o artigo apareceu no PoP dentro de repositório
A Rónai (UFJF)	<i>As imperfeições da tradução automática on-line</i>	Sim, foi retirado do periódico
A Rónai (UFJF)	<i>Tradução automática e competência tradutória: repensando interseções</i>	Sim, foi retirado do periódico
TradTerm (USP)	<i>Terminografia para a tradução automática: alimentar metal com termos</i>	Sim, foi retirado do periódico

Tradução e Comunicação: Revista Brasileira de Tradutores	<i>Um tradutor automático de expressões nominais com idéia de posse</i>	Sim, foi retirado do periódico
--	---	--------------------------------

Fonte: a autora.

Ao saber de onde as duplicatas foram recuperadas dentro da base de dados é possível concluirmos quais são os periódicos indexados no Google Acadêmico e aqueles sem indexação à plataforma. Os periódicos Belas Infiéis (UnB), Cultura e Tradução (UFPB), In-Traduções Revista do Programa de Pós-Graduação em Estudos da Tradução (UFSC) possuem indexação à base de dados do Google Acadêmico, mas o título desses periódicos não aparecem na base de dados. Além disso, nem todos os periódicos indexados na base de dados tem todos os seus artigos representados, por exemplo: TradTerm (USP) com dois artigos no *corpus* e Cadernos de Tradução (UFSC), com seis, não têm todos seus artigos presentes na base de dados. Sem mencionar que a TradTerm tinha apenas um artigo recuperado e Cadernos de Tradução três. Tal inconsistência deixa claro que nem todos os trabalhos acadêmicos estão na base de dados do Google Acadêmico, pois essa base deveria ter encontrado os outros cinco artigos que não apareceram nas duplicatas e o motivo não fica claro, pois não é somente a falta de indexação da instituição ou dos periódicos que podem ocultar o artigo na base de dados do Google Acadêmico.

Para a devida indexação, segundo as explicações dadas pelo próprio Google Acadêmico, as publicações precisam:

- a) ser hospedadas em uma página de conteúdo acadêmico que não contenha erros ou *links* quebrados, do contrário o acesso a elas se perde;
- b) ter até 5 MB e formato .PDF;
- c) ter sido publicadas de forma gratuita e estar disponível para aqueles que as acessam pela ferramenta;
- c) ocultar quaisquer tipos de publicidade ou propaganda;
- d) conter informações bibliográficas coerentes e completas, do contrário não serão identificadas;
- e) evidenciar apenas um local de publicação; se disponíveis em várias edições ou publicadas em mais de um local elas poderão ser desconsiderados; e
- f) ter fórum de discussão; embora não haja priorização ou hierarquia de um ou outro tema, a falta de discussões em uma área pode afetar a indexação.

Além desses fatores, o tempo de publicação pode influenciar a indexação. Tal indexação dos documentos pode levar semanas ou até meses, fazendo com que os documentos muito próximos do período da realização de uma pesquisa não apareçam na base de dados.

Portanto, consideramos esses fatores limitações dos estudos bibliométricos e cienciométricos. Por maior que seja o cuidado com a abrangência da pesquisa, é difícil encontrar uma base de dados livre de duplicatas, falhas de busca, documentos não relacionados com o tema ou incompatíveis com textos acadêmicos mais convencionais.

Ainda sobre os tipos de documentos no *corpus* geral, 33,3% dos documentos estão relacionados com trabalhos de conclusão de curso de graduação (quatro monografias), mestrado (12 são dissertações), doutorado (sete teses) ou relatórios de grupos de pesquisa (seis relatórios) de Tradução. Ter um número considerável de trabalhos acadêmicos de conclusão de graduação e pós-graduação no *corpus* vai além de mostrar a iniciação e evolução de novos pesquisadores, pois aponta também influência das pesquisas dos professores orientadores. Ademais, é pela quantidade desses tipos de documentos presentes no *corpus* que podemos visualizar os polos de ensino em nível de graduação e pós-graduação dos Estudos da Tradução e quais são os seus agentes mais atuantes quando o tema é discutir a TA. A Tabela 3 mostra a produção de cada instituição de acordo com o tipo de documento produzido.

Tabela 3 - Tipos e quantidade de documentos de cada Instituição de Ensino

Tipos de documento X Instituição de Ensino	N. de documentos por Instituição
Artigo científico	55
USP - São Carlos	8
Universidade Federal de Santa Catarina	7
USP – São Carlos; Universidade Federal de São Carlos - UFSCar;	
Universidade Estadual Paulista - UNESP	4
Instituição não mencionada	2
PUC-RIO	2
Universidade de Brasília	2
Universidade de Santa Cruz do Sul	2
Universidade Estadual do Ceará – UECE	2
Universidade Federal de Campina Grande	2
Universidade Federal de São Carlos (UFSCar)	2
Universidade Federal do Rio de Janeiro	2
Universidade Estadual Paulista (Unesp)	1
Centro Federal de Educação Tecnológica de Minas Gerais	1
Faculdade de Letras da Universidade Federal de Juiz de Fora	1
Faculdade Integrada Brasil Amazônia	1
FIVR/UNISEPE	1
PUC - SP	1
PUC-RS	1
UNIOESTE/Foz do Iguaçu	1
Universidad de las Américas Puebla	1
Universidade Católica de Louvain	1

Universidade de São Paulo (USP)	1
Universidade de Trás-os-Montes e Alto Douro	1
Universidade Estadual de Feira de Santana	1
Universidade Estadual de Goiás	1
Universidade Estadual de Maringá	1
Universidade Estadual do Sudoeste da Bahia	1
Universidade Federal de Ouro Preto	1
Universidade Federal do Paraná	1
Universidade Sagrado Coração de Bauru-SP	1
Universidade Tecnológica Federal do Paraná (UTFPR)	1
Dissertação de mestrado	12
Universidade Estadual Paulista (Unesp)	3
PUC-RIO	2
Universidade Federal de Santa Catarina	2
Fundação Getúlio Vargas	1
Universidade de São Paulo (USP)	1
Universidade Estadual de Campinas	1
Universidade Federal do Rio Grande do Sul	1
Universidade Federal de São Carlos (UFSCar)	1
Dossiê científico	1
Universidade Federal do Paraná	1
Ensaio crítico	1
Universidade Federal da Bahia	1
Monografia de final de curso	4
Universidade de Brasília	2
Escola de Aperfeiçoamento de Oficiais, Rio de Janeiro	1
Universidade Federal da Bahia	1
Relatório de pesquisa	6
USP – São Carlos; Universidade Federal de São Carlos - UFSCar;	
Universidade Estadual Paulista - UNESP	5
Universidade de São Paulo (USP)	1
Reportagem	1
Instituição não mencionada	1
Tese de doutorado	7
USP - São Carlos	2
Universidade de Brasília	1
Universidade Estadual Paulista (Unesp)	1
Universidade Federal de Juiz de Fora	1
Universidade Federal de Santa Catarina	1
Universidade Federal do Ceará	1
Total Geral	87

Fonte: a autora.

Com as informações presentes na Tabela 3, é possível medir a produção acadêmica de cada instituição de ensino e, não somente isso, mas também averiguar o tipo de documento que cada polo de ensino produz com maior frequência.

As instituições que produzem um número alto de dissertações e teses, logicamente, têm programas de pós-graduação que se dedicam às pesquisas sobre TA. E se as mesmas instituições se repetem nos artigos, por exemplo, pode ser um indício de que as pesquisas de pós-graduação estão gerando discussões que são publicadas em artigos.

Já na Tabela 2 foi possível perceber que os artigos são o tipo de documento mais comum presente no *corpus* e somando-se isso às informações da Tabela 3 percebe-se quais instituições publicaram a maior quantidade de artigos: USP – São Carlos, com oito publicações, seguida pela UFSC, com sete publicações. Essas duas instituições se destacam consideravelmente quando comparadas às outras instituições brasileiras. Não podemos deixar de mencionar as produções do NILC e que USP-São Carlos, UNESP e UFScar fazem parte desse núcleo e produzem em parceria, aumentando o número de publicações para essas instituições.

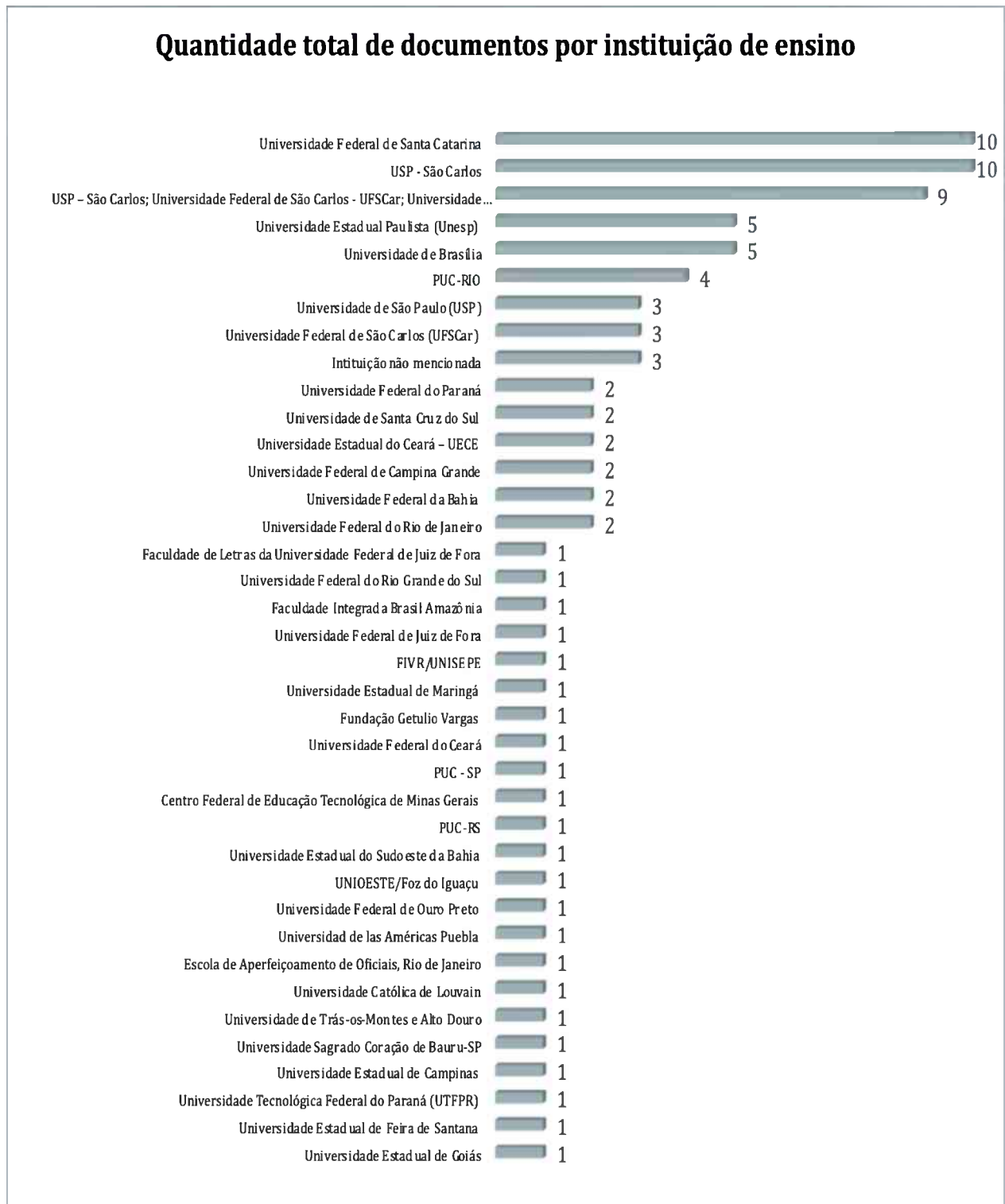
Quanto à procedência geográfica, fica evidente que o sul e sudeste ainda dominam o cenário de discussão sobre TA no Brasil. Podemos dizer que é bastante típico dos resultados dos estudos bibliométricos e cienciométricos brasileiros terem maior concentração no sul e sudeste do país, regiões nas quais estão concentradas as indústrias e o comércio de produtos e serviços. Alguns pesquisadores se preocupam com esses resultados (VANTI, 2002), principalmente por parecer indicar que o ensino e a pesquisa estão em favor da indústria, do comércio e das instituições financeiras localizadas nos grandes centros urbanos e não das instituições de ensino e pesquisa como um todo, independentemente da região à qual pertencem.

As instituições do Estado de São Paulo, do Distrito Federal e de Santa Catarina são responsáveis por 49 produções (55,2%), mais da metade de tudo o que foi encontrado sobre TA em 30 anos de discussões, com destaque para a USP - São Carlos, que além de abrigar a maior quantidade de discussões, trabalha em parceria com outras instituições de ensino do estado de São Paulo (NILC), que, em conjunto, são responsáveis por mais nove publicações dentro do *corpus*. Mas vale lembrar que apenas na segunda parte dessa pesquisa, quanto implementaremos a análise de conteúdo do que vem sendo discutido nesses documentos, é que poderemos perceber a relevância desses trabalhos e instituições para a TA e em consonância, ou não, com a formação de tradutores.

O Gráfico 3 mostra a produção total das instituições às quais os pesquisadores pertencem e que se propõem a discutir TA.

O número de documentos publicados por cada uma delas foi adicionado para facilitar a percepção das informações.

Gráfico 3 - Quantidade total de documentos por Instituição de Ensino



Fonte: a autora.

Para facilitar a visualização, um mapa ilustrativo de densidade mostra as regiões do Brasil com maior produção de trabalhos sobre TA nesses 30 anos, conforme evidencia a Figura 2.

Figura 2- Mapa de densidade de produção no Brasil



Fonte: a autora.

As instituições de ensino são os polos educacionais que abrigam os pesquisadores, sendo eles graduandos, pós-graduandos ou professores. Se um pesquisador tem grande número de publicações discutindo um assunto em uma área de conhecimento específico, consequentemente, a instituição à qual ele pertence vai se destacar nas discussões a respeito desse assunto.

Na Tabela 4, a seguir, apresentamos as grandes áreas e as áreas específicas do conhecimento dos documentos investigados que tratam da TA.

A grande área dos Estudos da Tradução detém 22 trabalhos, totalizando 25,3% dos textos que compõem o *corpus*. As áreas da Linguística e área de interface entre a Linguística e o Ensino 21 trabalhos, representando 24,13% do total analisado. A área de Letras, com 16%, tem 14 trabalhos presentes no *corpus*.

Contudo, cinco trabalhos que informavam ser da área de Letras e quatro da área da Linguística foram publicados nos periódicos brasileiros dos Estudos da Tradução. Isso demonstra

que podem haver, nos textos do *corpus*, discussões a respeito do ensino de Línguas (Letras) e de análises sintáticas (Linguística) utilizando TA.

Tabela 4 - Número de documentos por área de conhecimento dentro do *corpus* geral

Área do conhecimento que executou e publicou o estudo sobre TA	N.º. de documentos por área do conhecimento
Estudos da Tradução	22
Linguística	20
Ciência da Computação	15
Letras	14
Linguística Computacional	10
Não informado	2
Linguística e Ensino	1
Matemática Aplicada	1
Ciências Militares	1
Ciência da Informação	1
Total Geral	87

Fonte: a autora.

O que chama atenção são os 26 trabalhos (29,9%) envolvendo a grande área das Ciências Exatas, com a Ciência da Computação detendo 15 trabalhos (17,2%), a Linguística Computacional com 10 (11,5%) e Matemática Aplicada com um trabalho (1,15%) discutindo TA, o que prevê certas discussões dentro do *corpus* a respeito da arquitetura e desenvolvimento das ferramentas de TA, como já havíamos presumido.

Para entendermos melhor o que justifica a grande quantidade de trabalhos que envolvem Ciência da Computação e Linguística Computacional e de onde essa tendência é oriunda, foi necessário acrescentar a Tabela 5, que mostra quais foram os tipos de documentos publicados em cada área de conhecimento e quais as instituições responsáveis por esses estudos.

Tabela 5 - Tipos de documento por área de conhecimento e instituição de ensino

Áreas de conhecimento por: Tipos de documento x instituição de ensino	Contagem de Área de conhecimento
Ciência da computação	15
Artigo científico	10
USP - São Carlos	5
Centro Federal de Educação Tecnológica de Minas Gerais	1
Universidade de Trás-os-Montes e Alto Douro	1
Universidade Estadual de Feira de Santana	1
Universidade Estadual de Goiás	1
Universidade Federal de São Carlos (UFSCar)	1

Dissertação de mestrado	2
Universidade Federal do Rio Grande do Sul	1
Universidade Federal de São Carlos (UFSCar)	1
Relatório científico	1
USP – São Carlos; Universidade Federal de São Carlos - UFSCar; Universidade Estadual Paulista - UNESP	1
Tese	2
USP - São Carlos	2
Ciência da Informação	1
Tese de doutorado	1
Universidade de Brasília	1
Ciências Militares	1
Monografia de final de curso	1
Escola de Aperfeiçoamento de Oficiais, Rio de Janeiro	1
Estudos da Tradução	21
Artigo científico	16
Universidade Federal de Santa Catarina	5
Instituição não mencionada	2
Faculdade Integrada Brasil Amazônia	1
FIVR/UNISEPE	1
PUC – SP	1
Universidade de Brasília	1
Universidade de Santa Cruz do Sul	1
Universidade Estadual do Ceará – UECE	1
Universidade Estadual do Sudoeste da Bahia	1
Universidade Estadual Paulista (Unesp)	1
Universidade Federal de São Carlos (UFSCar)	1
Dissertação de mestrado	2
Universidade Federal de Santa Catarina	2
Monografia de final de curso	1
Universidade de Brasília	1
Reportagem	1
Instituição não mencionada	1
Tese de doutorado	1
Universidade Federal de Santa Catarina	1
Letras	14
Artigo científico	7
Faculdade de Letras da Universidade Federal de Juiz de Fora	1
PUC-RIO	1
Universidade Federal de Campina Grande	1
Universidade Federal de Ouro Preto	1
Universidade Federal do Paraná	1
Universidade Federal do Rio de Janeiro	1
Universidade Sagrado Coração de Bauru-SP (atual Unisagrado)	1
Dissertação de mestrado	2
PUC-RIO	2
Dossiê científico	1
Universidade Federal do Paraná	1
Ensaio crítico	1
Universidade Federal da Bahia	1
Monografia de final de curso	2
Universidade de Brasília	1
Universidade Federal da Bahia	1
Tese de doutorado	1
Universidade Federal do Ceará	1
Línguas estrangeiras e Tradução	1
Artigo científico	1
Universidade de Brasília	1

Linguística	20
Artigo científico	12
PUC-RIO	1
PUC-RS	1
UNIOESTE/Foz do Iguaçu	1
Universidad de las Américas Puebla	1
Universidade de Santa Cruz do Sul	1
Universidade Estadual de Maringá	1
Universidade Estadual do Ceará – UECE	1
USP – São Carlos; Universidade Federal de São Carlos - UFSCar;	1
Universidade Estadual Paulista - UNESP	1
Universidade Federal de Campina Grande	1
Universidade Federal de Santa Catarina	1
Universidade Federal do Rio de Janeiro	1
Universidade Tecnológica Federal do Paraná (UTFPR)	1
Dissertação de mestrado	5
Universidade Estadual Paulista (Unesp)	3
Universidade de São Paulo (USP)	1
Universidade Estadual de Campinas	1
Relatório científico	1
Universidade de São Paulo (USP)	1
Tese de doutorado	2
Universidade Estadual Paulista (Unesp)	1
Universidade Federal de Juiz de Fora	1
Linguística computacional	10
Artigo científico	6
USP - São Carlos	3
USP – São Carlos; Universidade Federal de São Carlos; UFSCar;	3
UNESP	1
Relatório científico	4
USP – São Carlos; Universidade Federal de São Carlos; UFSCar;	4
UNESP	1
Linguística e Ensino	1
Artigo científico	1
Universidade Federal de Santa Catarina	1
Matemática Aplicada	1
Dissertação de mestrado	1
Fundação Getulio Vargas	1
Não informado	2
Artigo científico	2
Universidade Católica de Louvain	1
Universidade de São Paulo (USP)	1
Total Geral	87

Fonte: a autora.

Ao observar o campo de Ciência da Computação, sete instituições se dividem nas produções sobre TA, contudo a USP-São Carlos se destaca com sete dos 15 documentos presentes, sendo fonte de cinco artigos e duas teses. Em se tratando da área de Linguística Computacional, os 10 documentos presentes no *corpus* têm participação da USP- São Carlos, sendo que três artigos e quatro relatórios que foram produzidos em parceria com outras instituições do Estado de São Paulo (USP, UFScar e Unesp). Ou seja, os 19 documentos que fazem a USP-São Carlos, juntamente com os parceiros, liderar a produção das discussões sobre TA em cenário brasileiro,

tratam de assuntos relacionados às Ciências Exatas. Já a USP, Unesp e UFScar também produziram, pelo menos, um documento dentro do campo Letras, Linguística e Estudos da Tradução.

Na área dos Estudos da Tradução, a UFSC se destaca com oito dos 21 documentos. Sendo fonte de cinco artigos, duas dissertações de mestrado e uma tese de doutorado. A UFSC também relaciona os estudos da TA com a área da Linguística e Linguística e Ensino, com mais dois artigos publicados. Dentro da Linguística, a Unesp se destaca com três dissertações, uma tese e um artigo a respeito de TA, totalizando 5 documentos.

Araújo e Alvarenga (2011) apontam que uma ampla distribuição nos trabalhos por área de conhecimento significa uma considerável difusão do tema, neste caso a TA, e indica um possível enriquecimento das discussões, sejam elas temáticas ou metodológicas, apontando uma interdisciplinaridade entre as principais áreas de conhecimento que tratam do tema. Por isso, ao perceber uma grande discussão da Ciência da Computação e Linguística Computacional a respeito de TA, e as parcerias entre instituições, começam a se formar as redes de ligação e as tendências interdisciplinares.

A Tabela 6 mostra a diversidade de autoria nas discussões de TA em cenário brasileiro, 68 autores para um total de 87 documentos do *corpus*, uma média de 1,28 documentos por autor. Essa quantidade de autores mostra uma área mais dispersa em relação ao número de autores. Aqueles que produziram 2 trabalhos ou mais, já se destacam perante os outros.

Tabela 6 – Autores que discutiram a respeito da TA no Brasil de 1990 a 2020

Autores que publicaram sobre TA no Brasil de 1989 a 2020	Quantidade de textos publicados pelo autor
Cleydstone Chaves dos Santos	4
Felipe Tassario Gomes	4
Helena de Medeiros Caseli	4
Mirna Fernanda de Oliveira	2
Lucia Specia	2
Diego R. Amancio	2
Augusto Velloso dos Santos Espindola	2
Maria José Damiani Costa	2
Thiago Blanch Pires	2
Ronaldo T. Martins	2
Wilker Ferreira Aziz	2
Lucia Helena Rozario da Silva	2
Tiago Martins da Cunha	2

Markus J. Weininger	1
Tatiane Silva Tavares	1
Marcello Modesto dos Santos	1
Débora Borsatti	1
Marcos de Campos Carneiro	1
Diana Santos	1
Norma Barbosa de Lima Fonseca	1
Aline Graciela Lermen dos Santos	1
Sara Cerqueira	1
Eliane Martins de Aguiar.	1
Luciano Rodrigues Lima	1
Érika Nogueira de Andrade Stupiello	1
Aline Naoe	1
Bento Carlos Dias-da-Silva	1
Milena Uzeda Garrão	1
Fernando da Silva	1
Paulo Henrique Barchi	1
Francisco Ramos de Melo	1
Roberlei Alves Bertucci	1
Gabriela Castelo Branco Ribeiro	1
Sílvia Gusmão Sales	1
Gislaine Caprioli Costa	1
Daniel Emílio Beck	1
Brita Banitz	1
Marcela Franco Fossey	1
Janailton Mick Vitor da Silva	1
Maria Alice O. C. Leal	1
Jorge Galarce	1
Maria Paula Fiorim Piruzelli	1
José Paulo de Araújo	1
Michel de Sousa Teixeira	1
Juliana Ramos do Nascimento	1
Cristiane de Souza	1
Juliano Desiderato Antonio	1
Nylcea Thereza de Siqueira Pedra	1
Lara Cristina Santos Talhaferro	1
Priscilla Pellegrino de Oliveira	1
Thiago Lima Vieira	1
Rafael Reis Costa	1
Vera Wannmacher Pereira	1
Cristiano Gonçalves Araújo	1
Carolina Alfaro	1

Sheila de Souza de Melo	1
Luana Moro	1
Silvia Malena Modesto Monteiro	1
Carolina Bisson de Souza	1
Thais Miranda e Borges	1
Caroline de Schaetzen	1
Débora Beatriz de Jesus Martins	1
Luciana Alves Graziuso	1
Leila Cristina de Mello Darin	1
Vinícius Silva Nogueira	1
Leonel Morgado	1
Adriana Riess Karnal	1
Liliane Simões dos Santos	1
Total Geral	87

Fonte: a autora.

Apenas três autores possuem quatro produções durante o período de 30 anos. Cleydstone Chaves dos Santos, com três artigos e uma tese; Felipe Tassario Gomes, com quatro artigos e Helena de Medeiros Caseli, com dois artigos, um relatório científico e uma tese de doutorado.

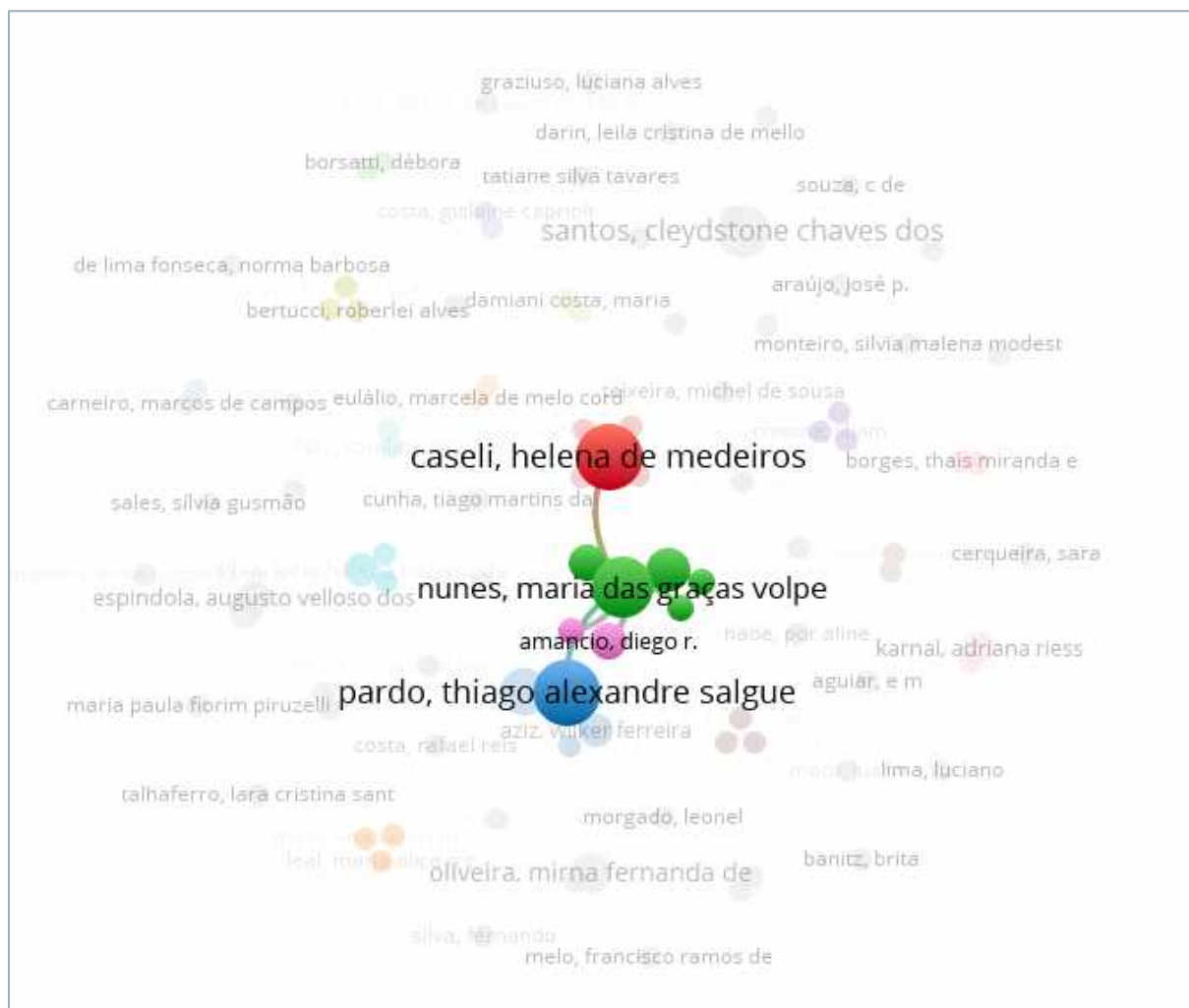
Dez outros autores possuem dois documentos cada. Ou seja, 13 autores detêm 32 documentos em um total de 87 encontrados, representando 36,8% do total de produção dessas discussões.

A Lei de Lotka, ou Lei do Quadrado Inverso, é utilizada para medir a produtividade dos autores e segue um modelo de distribuição do tamanho do *corpus versus* a frequência de todos os autores naquele conjunto de documentos (VANTI, 2002).

A Figura 3 destaca os autores com maior produtividade. Cada ponto no mapa destacado em amarelo representa a soma das produções e coproduções dos autores dentro do *corpus*. Quanto maior a produção e representatividade do autor dentro do mapa, maior o tamanho do ponto de destaque e mais amarelo e notório ficam esses pontos, com o nome do autor em destaque. E quanto mais próximos os *links* de coautoria, mais próximos os autores ficaram um do outro no mapa.

Do mesmo modo, autores com uma produção menor apareceram em tons de amarelo claro e com menor destaque dentro do mapa.

Figura 3 - Mapa de visualização de densidade de produção dentro do *corpus* geral



Fonte: a autora.

É preciso relacionar as produções e as instituições aos autores e coautores, a fim de entender a dispersão física entre eles, definir as redes de ligação do *corpus*, determinar quem são os autores de maior impacto e se os principais autores pertencem às instituições com maior impacto de produção, com quem eles se relacionam (coautorias e parcerias) e a qual área de conhecimento pertencem. Essas informações, juntamente com os picos de publicação, definem o mapeamento cienciométrico que nossa dissertação se propôs a fazer.

Os autores com mais de um documento constam na Tabela 7, que nos possibilita observar quais eram suas relações de coautoria, de que instituição e área do conhecimento eles pertenciam. Apenas os autores com dois ou mais documentos foram considerados, por questões de espaço e para evitar o excesso de informação.

Tabela 7 - Relação de autores com área de pesquisa, instituição e coautoria

Autor	Área de Pesquisa	Instituição	Coautoria	N. de docs por autor
Cleydstone Chaves dos Santos				4
	Estudos da Tradução			4
		Universidade Federal de Santa Catarina		4
		Maria Carmelita P. Dias		1
Felipe Tassario Gomes				4
	Ciência da Computação			2
		USP - São Carlos		2
		Thiago Alexandre Salgueiro Pardo		2
	Linguística computacional			2
		USP - São Carlos		2
		Thiago Alexandre Salgueiro Pardo		2
Helena de Medeiros Caseli				4
	Ciência da Computação			1
		USP - São Carlos		1
	Estudos da Tradução			1
		Universidade Federal de São Carlos (UFSCar)		1
	Linguística computacional			2
		USP – São Carlos; Universidade Federal de São Carlos - UFSCar; Universidade Estadual Paulista - UNESP		2
		Maria das Graças V. Nunes		1
		Israel Aono Nunes		1
Mirna Fernanda de Oliveira				2
	Linguística			2
		UNIOESTE/Foz do Iguaçu		1
		Universidade Estadual Paulista (Unesp)		1
Lucia Specia				2
	Ciência da Computação			2
		USP - São Carlos		2
		Maria das Graças V. Nunes		1
Diego R. Amancio				2
	Linguística computacional			2
		USP - São Carlos		1
		Maria das Graças V. Nunes; Thiago Alexandre Salgueiro Pardo; Osvaldo N.O. Jr		1
		USP – São Carlos; Universidade Federal de São Carlos - UFSCar; Universidade Estadual Paulista - UNESP		1
		Maria das Graças V. Nunes		1
Augusto Velloso dos Santos Espindola				2
	Letras			1
		Universidade de Brasília		1

Línguas estrangeiras e Tradução	1
Universidade de Brasília	1
Thiago Blanch Pires	1
Maria José Damiani Costa	2
Linguística	1
Universidade Federal de Santa Catarina	1
Ana Paula de Carvalho Demétrio	1
Linguística e ensino	1
Universidade Federal de Santa Catarina	1
Meta Elisabeth Zipser; Silvana Ayub Polchlopek	1
Thiago Blanch Pires	2
Ciência da Informação	1
Universidade de Brasília	1
Estudos da Tradução	1
Universidade de Brasília	1
Ronaldo T. Martins	2
Linguística computacional	2
USP – São Carlos; Universidade Federal de São Carlos - UFSCar; Universidade Estadual Paulista - UNESP	2
Maria das Graças V. Nunes	1
Wilker Ferreira Aziz	2
Ciência da Computação	2
USP - São Carlos	2
Ivandré Paraboni; Thiago Alexandre Salgueiro Pardo	2
Lucia Helena Rozario da Silva	2
Linguística	1
Universidade de São Paulo (USP)	1
Não informado	1
Universidade de São Paulo (USP)	1
Marcello Modesto; Alon Lavie	1
Tiago Martins da Cunha	2
Letras	1
Universidade Federal do Ceará	1
Linguística	1
Universidade Estadual do Ceará – UECE	1
Total	32

Fonte: a autora.

A Tabela 7 descreve as redes de coautoria entre os pesquisadores fornecidos pelo VOSviewer, apresentados nas Figuras 2 e 3. Ou seja, ela mostra todas as informações que a ferramenta usa para criar os *clusters*, porém de forma detalhada, permitindo identificar essas relações. O que mais chama atenção nessa análise é a quantidade de trabalhos vindos das mesmas

instituições e as relações de coautoria entre os autores dessas instituições. Ao observar os dois pesquisadores da UnB, por exemplo, existe uma relação entre eles e Thiago Blanch Pires, autor de dois textos, que também trabalha em conjunto com Augusto Velloso dos Santos Espindola. As áreas de conhecimento desses pesquisadores são diferentes, sendo que o primeiro se dedica à Ciência da Informação e aos Estudos da Tradução e o outro, à área de Letras, Línguas Estrangeiras e Tradução. Isso mostra a interdisciplinaridade entre essas conexões, e comprova a afirmação Araújo e Alvarenga (2011), de que essas relações interdisciplinares acabam enriquecendo as discussões e expandindo o tema.

Essa mesma relação de interdisciplinaridade acontece com os autores da USP-São Carlos, UFScar e Unesp, pois formam um Núcleo Interinstitucional de Linguística Computacional - NILC. Os autores estão ligados por *links* de coautoria e enriquecem as discussões vindas de áreas e instituições diferentes. Existe uma relação entre Felipe Tassario Gomes, que escreve em parceria com Thiago Alexandre Salgueiro Pardo, que participa de coautorias com Diego R. Amancio, Wilker Ferreira Aziz, Osvaldo N.O. J, Ivandré Paraboni e Maria das Graças V. Nunes – que é a pesquisadora que se conecta aos pesquisadores Helena de Medeiros Caseli e Israel Nunes, Ronaldo T. Martins e Lucia Specia. Todos esses pesquisadores trabalham de forma interdisciplinar, aliando Ciência da Computação, Linguística Computacional e Estudos da Tradução, embora os trabalhos vindos da USP-São Carlos e seus parceiros sejam, em sua grande maioria, relacionados com as duas primeiras áreas. Vale lembrar que foram analisadas as relações apenas dos autores que mais se destacam no *corpus*, porém, outras coautorias também podem ser destacadas com autores que possuem apenas um documento publicado.

Os números do Tabela 8 mostram quantas vezes, no *corpus* geral com 87 documentos, quais pesquisadores são citados como coautores.

Tabela 8 - Número de coautorias por pesquisador

Coautores	Quantidade de coautorias
Maria das Graças V. Nunes	4
Thiago Alexandre Salgueiro Pardo	4
Helena de Medeiros Caseli	3
Thiago Alexandre Salgueiro Pardo; Ivandré Paraboni	2
Maria Carmelita P. Dias	2
Fátima de Gênova Daniel	1
Janine Maria Mendonça Pimentel	1
Aléssio Miranda Júnior	1
Rosângela Gabriel	1

Hellen Carmo de Oliveira Matos; Emilie Rebecca Bastos Dias	1
Mirna Fernanda de Oliveira	1
Ana Paula de Carvalho Demétrio	1
Meta Elisabeth Zipser; Silvana Ayub Polchlopek	1
Líliam Cristina Marins; Giovanna Martinez Ursulino	1
Maria Lígia Freire Guilherme; Bárbara Branco Puppi	1
Israel Aono Nunes	1
Maria das Graças V. Nunes; Thiago Alexandre Salgueiro Pardo; Osvaldo N.O. Jr	1
Tatiana Pedrolongo; Ronaldo Teixeira Martins; Maria das Graças V. Nunes	1
Thiago Blanch Pires	1
Cleydstone Chaves dos Santos	1
Adriana Riess Karnal	1
Ana Lúcia L. M. Maia; Thiago D’Martin Maia	1
Marcela de Melo Cordeiro Eulálio	1
Ruth Bohunovsky	1
Lincoln Paulo Fernandes	1
Marcello Modesto; Alon Lavie	1
(vazio)	
Total Geral	36

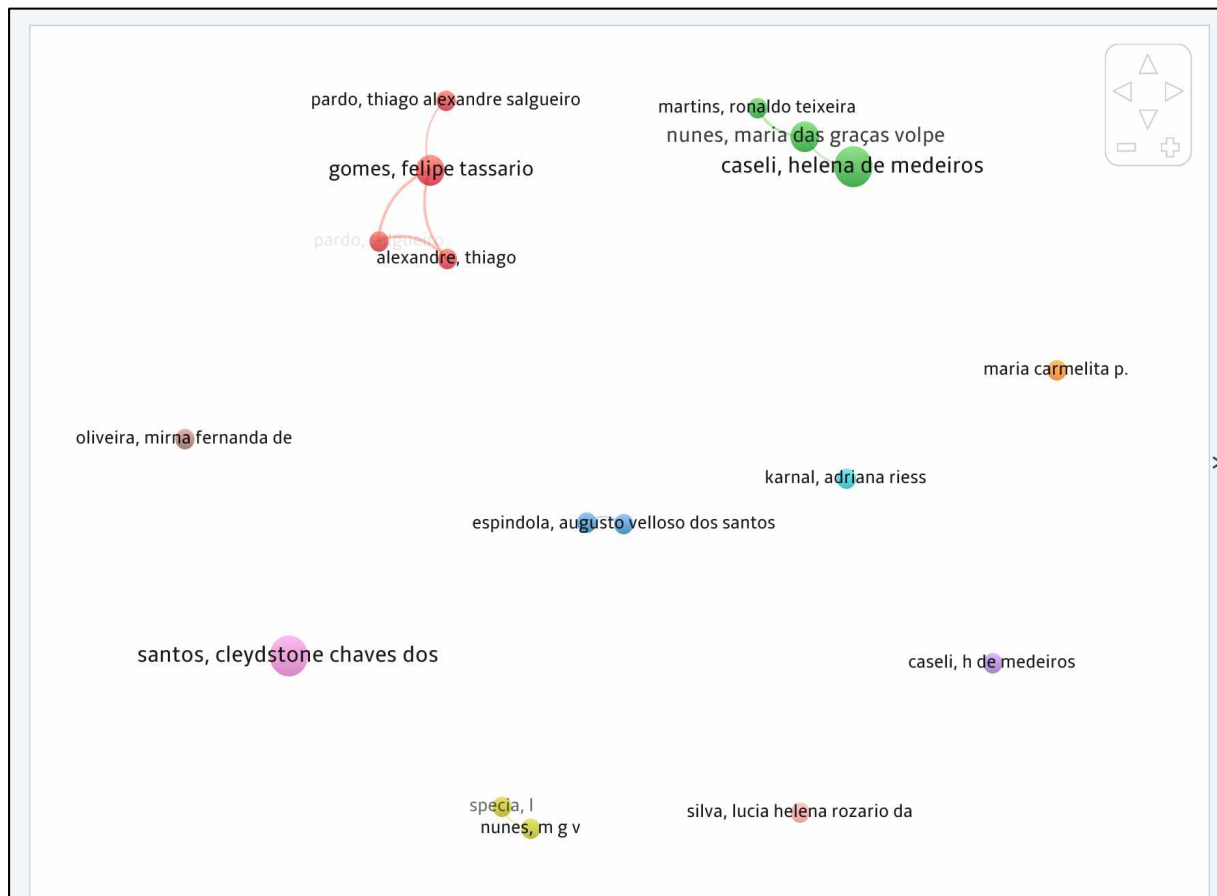
Fonte: a autora.

A Tabela 8 mostra que Thiago Alexandre Salgueiro Pardo é o autor que possui mais trabalhos em coautoria: além de publicar quatro trabalhos sozinho, estabeleceu duas coautorias com Ivandré Paraboni e uma com Osvaldo N. O. Jr e Maria das Graças V. Nunes.

Maria das Graças V. Nunes é a segunda autora com maior ocorrência de coautorias, cinco no total. Helena de Medeiros Caseli também tem destaque, aparecendo em três trabalhos como coautora, e Cleydstone Chaves dos Santos, com duas coautorias.

A Figura 5 mostra o mapa de visualização de ligações de coautoria, em que podemos perceber as relações ora mencionadas nas redes de coautoria no mapa.

Figura 5- Visualização de ligações dos autores



Fonte: a autora.

Quanto mais próximos dois autores estiverem representados no mapa de coautoria, maior é a rede de coautoria entre eles. Ou seja, pesquisam em conjunto ou citam uns aos outros. A esfera representa a dimensão do elemento em relação à dimensão do *corpus* analisado. Portanto, os autores de maior relevância na área de conhecimento em questão são representados por esferas maiores e a representação da relação com outros autores, chamadas de *clusters*, expressa as similaridades de pesquisa dos envolvidos, por exemplo, autores que discutem o mesmo tema de pesquisa tendem a formar um *cluster*.

No mapa de redes de visualização podemos perceber que autores das mesmas instituições aparecem em um *cluster* em vermelho – autores da USP-São Carlos e do NILC que mostram uma forte ligação de produção dentro do *corpus* e assim sucessivamente.

A seguir, 24 documentos dizem respeito a monografias, dissertações, teses e relatórios que receberam orientação de professores (Tabela 9). Embora não sejam o foco de nossa dissertação na análise qualitativa, foram aqui elencadas para fins de visualizarmos os dados cientométricos.

Tabela 9 - Número de orientações de teses e dissertações

Orientador	Tipo de texto	Contagem por Orientador
Prof. Bento Carlos Dias da Silva	Tese e Dissertação	2
Profa. Dra. Érika Nogueira de Andrade Stupiello	Dissertações	2
Profa. Dra. Maria das Graças Volpe Nunes	Teses	2
Cap QCO Carlos Adreson da Silva	Monografia	1
Profa. Cláudia Freitas	Dissertação	1
Prof. Me. Francisco C. Sampaio de Menezes	Monografia	1
Prof. Dr. Cláudio Gottschalg Duque	Tese	1
Profa. Dra. Helena de Medeiros Caseli	Dissertação	1
Prof. Dr. Leonel Figueiredo Alencar	Tese	1
Prof. Dr. Lincoln P. Fernandes	Tese	1
Prof. Dr. Marcello Modesto	Dissertação	1
Prof. Renato Rocha Souza	Dissertação	1
Prof. Dr. Markus J. Weininger	Dissertação	1
Prof. Dr. Karin Becker	Dissertação	1
Profa. Dra. Ina Emmel	Dissertação	1
Prof. Cleystone Chaves dos Santos	Artigo	1
Profa. Manoela C. C. Carvalho da Silva	Monografia	1
Prof. Dr. Rodrigo Esteves de Lima Lopes	Dissertação	1
Profa. Maria Carmelita Padua Dias	Dissertação	1
Prof. Dr. Thiago Blanch Pires	Monografia	1
Prof. Dr. Tiago Timponi Torrent	Tese	1
Total Geral		24

Fonte: a autora.

Nas orientações, se destacam a Profa. Dra. Érika Nogueira de Andrade Stupiello e o Prof. Bento Carlos Dias da Silva, ambos da Unesp, com duas orientações cada e a Profa. Dra. Maria das Graças Volpe Nunes, também com duas orientações pela USP-São Carlos. Todos eles orientando teses e dissertações.

Um dado interessante a ser mencionado é que alguns alunos que receberam orientações mais tarde se tornam orientadores, como por exemplo o Prof. Dr. Cleystone Chaves dos Santos e Profa. Dra. Helena de Medeiros Caseli.

No caso de Cleystone Chaves do Santos, um dos autores com mais documentos no *corpus* (quatro documentos como autor), podemos perceber sua evolução como pesquisador. Em 2011 e 2012 ele escreve dois artigos referentes à sua pesquisa de doutorado; em 2014 defende a tese de doutorado, orientada pelo Prof. Dr. Lincoln P. Fernandes – todos os trabalhos feitos enquanto

aluno da UFSC. Ainda em 2012 ele orienta um artigo e, em 2013, é coautor de outro, já como professor da UFCG – Universidade Federal de Campina Grande. Isso mostra que um pesquisador pode orientar e ser orientado e migrar por instituições diversas, proliferando discussões ao longo de sua trajetória acadêmica, pois, até então, a UFCG ainda não tinha publicado nada sobre TA.

A Profa. Dra. Helena de Medeiros Caseli é uma das autoras que mais se destaca no *corpus*, e é responsável por quatro documentos: artigo em 2005, tese de doutorado em 2007, um relatório em 2009 e um último artigo em 2017, foi coautora em três documentos: dois relatórios (2010/2013) e um artigo em 2011, e ainda orientou uma dissertação em 2012. A professora Caseli começou a produzir em 2005 enquanto aluna de doutorado pelo NILC, sendo orientada pela Profa. Dra. Maria das Graças Volpe Nunes e seguiu com projetos de pesquisa produzindo relatórios - NILC e orientando alunos após a conclusão do doutorado. Caseli tem produções pela USP-São Carlos em 2005 e foi orientada pela Profa. Dra. Maria das Graças Volpe Nunes em 2007, seguiu produzindo relatórios em 2009. A partir desse período, as produções mudam de instituição, Caseli é coautora de um artigo pela UFSCar em 2011, orienta uma dissertação pela mesma instituição em 2012 e produz um artigo em 2017. Ou seja, ela começa sua carreira de pesquisadora na USP-São Carlos e migra para a UFSCar por volta de 2010/2011, mas ainda mantém suas relações de parceria com o NILC, podendo ser um dos elos para o diálogo entre as instituições paulistas.

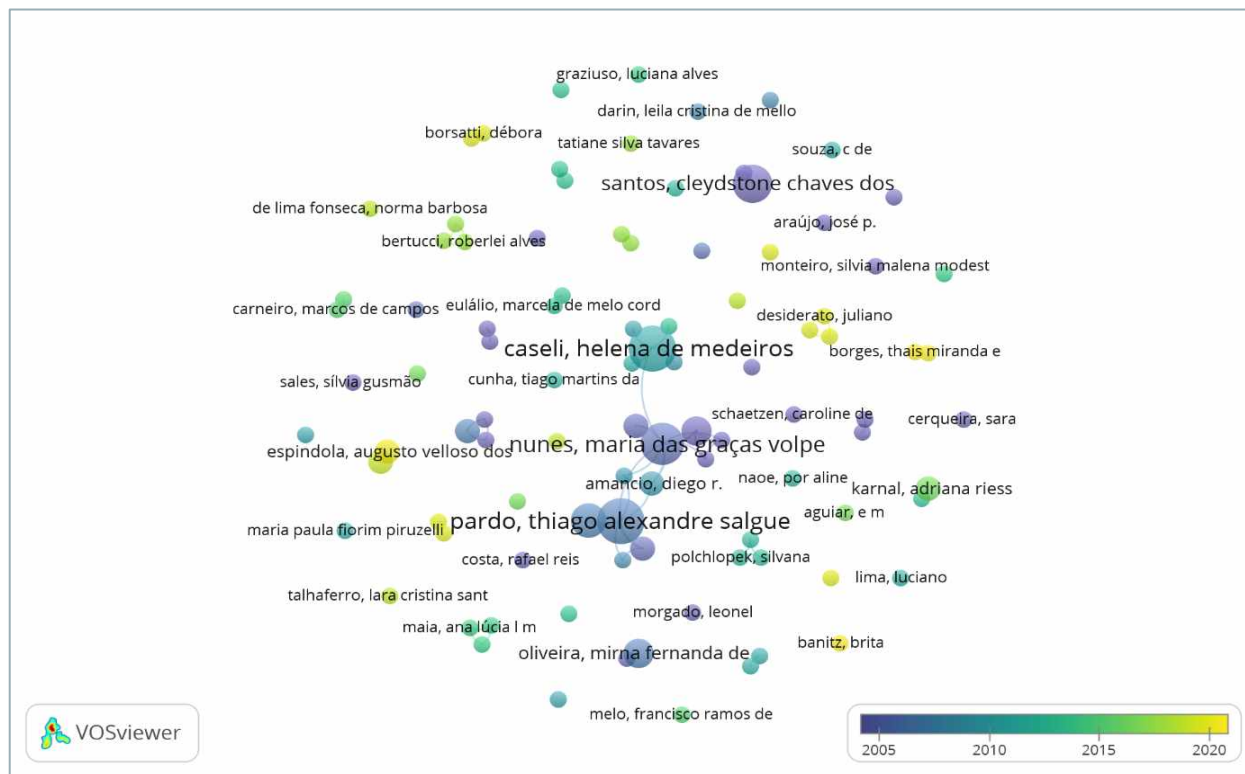
A já mencionada Profa. Dra. Maria das Graças Volpe Nunes¹⁶, que possui dois orientações e seis coautorias, é uma pesquisadora/orientadora-chave nas redes de produção que discute TA no Brasil. Ela se conecta a pesquisadores importantes dentro do *corpus*, como Caseli, antes mencionada, Thiago Alexandre Salgueiro Pardo – com sete coautorias, Ronaldo Teixeira Martins – com dois documentos e Lucia Specia com uma tese dentro do *corpus*. Todos pesquisadores da USP-São Carlos com um pico de produção entre 2003 a 2009. Após essa data, Volpe Nunes não participa mais de nenhuma produção, o que pode indicar uma aposentadoria e uma sucessão por sua orientanda, Helena Caseli.

A ferramenta VOSviewer também fornece dados para a análise de período de produção, na qual os ícones, com o nome do autor do documento, aparecem com cores referentes ao período de produção do documento. A escala temporal começa com o primeiro pico de produção (≈ 2005) com uma coloração arroxeadada e progride no tempo mudando para tons gradientes de azul (≈ 2010),

¹⁶ Ao concluir o doutorado e voltar a São Carlos, em 1991, Maria das Graças Volpe Nunes encontrou dois outros professores que estavam atuando nesse campo: Sandra Aluísio, no ICMC, e Osvaldo Novais de Oliveira Júnior, no IFSC. A união de forças levou o grupo a aceitar o desafio de criar o primeiro revisor automático para o português, o ReGra, projeto que Volpe Nunes coordenou de 1993 a 2008.

verde (≈ 2015) e amarelo (≈ 2020). A Figura 6 mostra essa visualização por períodos de produção dos documentos presentes no *corpus* geral.

Figura 6 - Visualização por períodos de produção



Fonte: a autora.

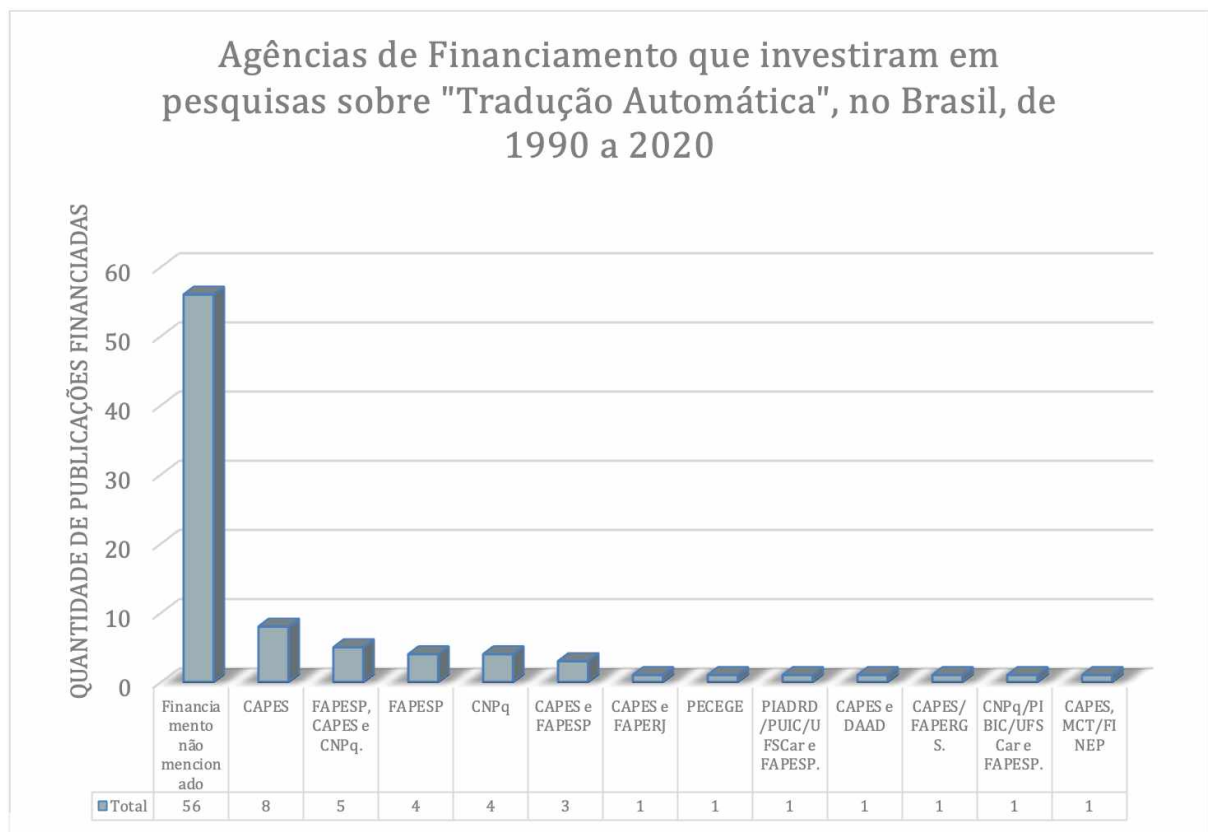
Com a figura de visualização por período, podemos perceber as épocas de produção dos *clusters* principais e quais são os autores com discussões mais atuais. O *cluster* de pesquisadores do NILC teve o pico de publicação entre 2005 e 2010, com cores roxas e azuis. Como discutido antes, esses dados atestam a interdisciplinaridade entre a Ciência da Computação, Linguística Computacional e Estudos da Tradução. Essas duas informações levam a crer que esses autores tratam da construção e desenvolvimento das ferramentas, que tiveram grandes investigações nos anos 2000. Podemos assim inferir que as pesquisas nessa área demoram entre cinco a dez anos para se concretizarem.

Já os ícones mais claros, verde claro e amarelo são discussões mais atuais e se destacam os autores da UnB, como Augusto Velloso dos Santos Espindola, da Unesp, Carollina Bisson de Souza e da Universidade de Santa Cruz do Sul, Adriana Riess Karnal e Débora Borsatti. E, ao recorrermos novamente à Figura 1, que é a representação gráfica das palavras-chave com maior densidade dentro do *corpus*, é possível prever quais discussões estão sendo levantadas em relação

ao tempo de produção. Por exemplo, os autores mencionados em amarelo estão relacionados com a palavra-chave “Incompatibilidade Intersemiótica”, os ícones em verde, “Estatística” e discussões sobre o “Google Tradutor”; os mais azulados tratam de “Tradução Automática”, “Linguística” e “Linguística Computacional”.

As agências de financiamento e fomento de pesquisas em território nacional nos ajudam a perceber a quantidade de pesquisadores que dependem do auxílio governamental para dar prosseguimento às suas pesquisas, o quanto o governo – federal e estadual – estão auxiliando nas pesquisas sobre Tradução Automática, incentivando a expansão dessa tecnologia em solo nacional e, além disso, é um fator de peso para publicações no Brasil, pois se o financiamento for do governo, as publicações devem ser de livre acesso (Gráfico 4).

Gráfico 4 -Agências de financiamento que investiram em pesquisas sobre TA, no Brasil, de 1990 a 2020



Fonte: a autora.

O Gráfico 4 mostra que, das 87 pesquisas em TA em cenário brasileiro, 31 pesquisas (35,6% do total) mencionaram que receberam auxílio e incentivo à pesquisa, sendo que o principal financiador foi a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES, responsável único pelo financiamento de oito pesquisas, e financiando outras 12, em parceria com

outras agências de fomento. Ou seja, dentro das 31 pesquisas que mencionam algum tipo de financiamento, 20 (64,5%) tiveram apoio da CAPES. A CAPES se mostra, nesse cenário, como principal apoiadora das pesquisas em TA no Brasil. A Tabela 10 mostra o número exato de financiamentos para cada tipo de documento presente no *corpus* geral.

Tabela 10 - Quantidade de financiamentos mencionados em cada tipo de documento por agência de financiamento

Tipos de documentos	Contagem de Financiamento
Artigo científico	55
CAPES	2
CAPES/FAPERGS.	1
CNPq	1
FAPESP	2
FAPESP, CAPES e CNPq.	5
Financiamento não mencionado	44
Dissertação de mestrado	12
CAPES	3
CAPES e FAPERJ	1
CAPES e FAPESP	2
CNPq	2
FAPESP	1
Financiamento não mencionado	2
PECEGE	1
Dossiê científico	1
Financiamento não mencionado	1
Ensaio crítico	1
Financiamento não mencionado	1
Monografia de final de curso	4
Financiamento não mencionado	4
Relatório científico	6
CNPq	1
CNPq/PIBIC/UFSCar e FAPESP.	1
FAPESP	1
Financiamento não mencionado	2
PIADRD/PUIC/UFSCar e FAPESP.	1
Reportagem	1
Financiamento não mencionado	1
Tese de doutorado	7
CAPES	3
CAPES e DAAD	1
CAPES e FAPESP	1
CAPES, MCT/FINEP	1

Financiamento não mencionado	1
Total de documentos	87

Fonte: a autora.

A CAPES, fundada pelo Ministério da Educação (MEC), se propõe a expandir e consolidar a pós-graduação *stricto sensu* (mestrado e doutorado) em território nacional (CAPES, 2022), o que parece se justificar pela quantidade de documentos de pós-graduação presentes no *corpus* geral. Na Tabela 10 é possível perceber que a CAPES está presente em todos os financiamentos de tese de doutorado (sete teses) e em metade das dissertações de mestrado (seis dissertações).

Outro dado que se pode inferir é que os artigos científicos muitas vezes são oriundos de pesquisas de dissertações de mestrado e teses de doutorado. Embora a maioria dos artigos não mencionem financiamento (44 artigos dos 87 gerais), dos 12 que mencionam, em oito deles a CAPES era, pelo menos, uma das financiadoras. Ou seja, existia um trabalho de pós-graduação *stricto sensu* por trás do que foi discutido naquele artigo.

A FAPESP - A Fundação de Amparo à Pesquisa do Estado de São Paulo, também se destaca no financiamento das pesquisas, aparecendo em 10 das 31 pesquisas financiadas. A fundação “é uma das principais agências de fomento à pesquisa científica e tecnológica do país e está ligada à Secretaria de Desenvolvimento Econômico. Trabalha com o orçamento de 1% do total da receita tributária do Estado, e financia a investigação, o intercâmbio e a divulgação da ciência e da tecnologia produzida em São Paulo” (FAPESP, 2022). Isso mostra que muitas pesquisas no *corpus* são desenvolvidas no estado de São Paulo e que, em sua grande maioria, são financiadas.

Alguns grupos de pesquisa também são financiados por alguma agência de fomento, principalmente a FAPESP. As monografias não receberam financiamentos, talvez por serem pesquisas de curta duração e requisitos para conclusão de cursos de graduação. E, geralmente, os artigos não mencionam financiamento.

Na soma dos 95 autores e coautores, 93 são de nacionalidade brasileira, um português, uma autora alemã e uma belga. Sendo que os três estrangeiros estavam presentes nas publicações das revistas. A autora belga Caroline de Schetzen publicou o artigo em duas versões, português e francês no periódico TradTerm (USP), em 1996. E, outros dois artigos do *corpus* estavam em inglês – o artigo da alemã Brita Banitz, publicado no periódico Cadernos de Tradução em 2020 e o do brasileiro Thiago Blanch Pires, de 2015. Os dois artigos possuíam título, resumo e palavras-chave em português. Com exceção dos trabalhos mencionados acima, o *corpus* foi inteiramente composto por textos em língua portuguesa.

CAPÍTULO 4: RESULTADOS QUALITATIVOS: O QUE TRATAM AS PESQUISAS EM CENÁRIO BRASILEIRO SOBRE A TRADUÇÃO AUTOMÁTICA?

Com a avaliação qualitativa buscamos identificar o que foi discutido pelos pesquisadores através de análises do conteúdo. Ou seja, o viés qualitativo, para nós, é um modelo de investigação que considera aspectos que não podem ser traduzidos em números, mas que permitem que formulemos hipóteses sobre o tema investigado.

Para este capítulo de análise qualitativa da pesquisa, consideramos apenas os artigos científicos presentes no *corpus*, isto é, 55 artigos no total (Tabela 2), sendo 17 deles retirados dos periódicos brasileiros da área de Estudos da Tradução e 38 retirados da base de dados do Google Acadêmico, através do PoP.

Como antes mencionamos, o objetivo inicial de nossa dissertação era identificar e compreender o que vem sendo discutido a respeito da TA nos periódicos brasileiros da área, entretanto, devido ao baixo número de publicações encontradas, expandimos a pesquisa para a base de dados do Google Acadêmico. Com isso, encontramos 87 documentos, de vários tipos textuais.

A fim de mantermos a padronização do tipo de texto e veículos de publicação utilizados, optamos apenas pela análise qualitativa dos artigos científicos, que compõe 63,2% da totalidade do *corpus*. Além disso, os artigos, na maioria das vezes, são oriundos de pesquisas de mestrado e doutorado – as dissertações e teses são outra grande parcela dos documentos presentes no *corpus*, 12 dissertações de mestrado e 7 teses de doutorado – 19 documentos (21,8%). Portanto, mesmo não considerando as dissertações e teses, muitos dos estudos presentes nesses trabalhos podem ter sido representados de forma mais breve nos artigos científicos presentes do *corpus*, principalmente por aqueles autores que têm dois ou mais documentos publicados. Outra justificativa para se contemplar apenas esse tipo de publicação se dá devido aos artigos científicos serem os principais responsáveis pelo diálogo entre os pesquisadores.

Após o mapeamento das pesquisas sobre TA no Brasil, que se configurou em nossa dissertação como o primeiro objetivo específico, a pesquisa qualitativa contempla os dois últimos objetivos específicos, a saber, como antes delimitado: 2. analisar os resultados que esses trabalhos alcançaram; 3. descrever e problematizar os temas de pesquisa mais recorrentes no interior das discussões implementadas por esses trabalhos.

Para atingirmos os objetivos específicos 2 e 3, os 55 artigos passaram por uma análise qualitativa, a qual exigiu uma leitura minuciosa e uma categorização dos assuntos discutidos em

cada um deles, para que pudéssemos apontar quais foram as principais discussões abordadas ao longo desses 30 anos sobre a TA no Brasil.

O processo de categorização das discussões ocorreu da seguinte forma: os trabalhos foram analisados um a um por seus resumos, objetivos, metodologia e resultados, isto é, lidos na íntegra. Buscamos, assim, após a leitura, agrupá-los em categorias de discussão.

Existia a opção de utilizarmos as ferramentas específicas que viabilizam diferentes tipos de análise de dados textuais, tais como o Iramuteq e o Maxqda, por exemplo. Entretanto, já tínhamos o mapa de ligações de palavras-chave feitas pelo VOSviewer e a tabela do Excel com todos os metadados quantitativos referentes a cada um dos documentos presentes no *corpus*, tornando a análise de conteúdo viável e assim evitando o excesso de processos metodológicos. Mas, não descartamos a possibilidade de, em trabalhos futuros, analisarmos esse mesmo *corpus* de forma linguisticamente mais detalhada, utilizando uma dessas tecnologias.

Nos 55 artigos, emergiram, ao nosso ver, seis categorias de discussões principais:

- TA e o desenvolvimento, a arquitetura e/ou o funcionamento das ferramentas.
- TA e os Estudos da Tradução;
- TA e avaliação de performance e qualidade das ferramentas através de comparações linguísticas;
- TA e a análise, apresentação ou comparação de uma ou mais ferramentas de TA;
- TA e o uso e a recepção dos usuários; e
- TA e o ensino de línguas estrangeiras.

A Tabela 11 mostra a quantidade de artigos que correspondem a cada categoria. A maioria dos artigos (15 artigos – 27,3%) publicados no Brasil em 30 anos tratam, portanto, do desenvolvimento ou arquitetura das ferramentas de TA, seguidos por pesquisas relacionadas com os Estudos da Tradução (13 artigos - 23,6%) e avaliação de performance de qualidade das ferramentas com 12 artigos (21,81%). Dentre os 55 artigos, 76,4% dos artigos não demonstram relação direta com os Estudos da Tradução.

Tabela 11 - Quantidade de artigos por categoria de discussão

Categorias de discussão	Qtde de artigos por categoria
TA e o desenvolvimento, a arquitetura e/ou o funcionamento das ferramentas	15
TA e os Estudos da Tradução	13
TA e avaliação de performance e qualidade das ferramentas através de comparações linguísticas	12
TA e a análise, apresentação ou comparação de uma ou mais ferramentas de TA	9
TA e o uso e a recepção dos usuários	3
TA e o ensino de línguas estrangeiras	3
Total Geral	55

Fonte: a autora.

Nas subseções seguintes, analisaremos qualitativamente os conteúdos desses artigos a partir da quantidade decrescente em que aparecem na Tabela 11. Deixaremos, porém, a discussão sobre os artigos que se voltam aos Estudos da Tradução por último (subseção 4.6).

4.1 TA e o desenvolvimento, a arquitetura e/ou o funcionamento das ferramentas

Muitos trabalhos discutem a arquitetura ou explicam o desenvolvimento computacional e linguístico das ferramentas de TA, suas possibilidades de aprimoramento, ampliação dos sistemas ou como elas foram projetadas, dando ênfase ao método de construção. Por exemplo, como ou quais são as características das ferramentas automáticas baseadas em estatística, em regras, ou em rede neural.

Dos 55 artigos, 15, portanto, referem-se à elaboração e averiguação do desenvolvimento de ferramentas de TA, como por exemplo: o estudo e aprimoramento do sistema Apertium; a implementação e teste de um pós-editor intitulado Trapezio; ou a implementação de *corpora* eletrônicos nas ferramentas de TA. Nesses 15 artigos que discutem o desenvolvimento e a engenharia das ferramentas, nove deles foram escritos por pesquisadores da área de Linguística Computacional e Ciência da Computação, dois pela Linguística e dois pelos Estudos da Tradução. O artigo escrito por Helena Caseli (2017), discute três tipos de ferramentas de TA: TA baseada em regras, em estatística e neural.

Esses artigos são altamente técnicos em relação à estrutura das ferramentas e as discussões neles estabelecidas são voltadas aos estudos da Linguística Computacional, sem menção ou preocupação quanto à sua utilização por profissionais e estudantes da área de Tradução. Destacamos, porém, o artigo de Caseli (2017), que tem um viés didático e detalhado sobre os tipos de ferramentas e suas estruturas, que, quiçá, poderia servir à formação de profissionais e estudantes da Tradução em cursos de graduação ou pós-graduação.

4.2 TA e avaliação de performance e qualidade das ferramentas através de comparações linguísticas

Os artigos categorizados na categoria “TA e avaliação de performance e qualidade das ferramentas através de comparações linguísticas” compreendem 21,81% do *corpus* (12 artigos).

Esses artigos, em geral, avaliam os resultados da TA, geralmente comparando os produtos dessas ferramentas com a tradução humana. Eles exploram a capacidade das ferramentas de TA em traduzir estruturas semânticas, lexicais, léxico especializado, expressões idiomáticas e tempos verbais. Nesses artigos aparecem a classificação dos erros das ferramentas e a discussão principalmente sobre a polêmica da “equivalência”. Esses trabalhos são, em sua grande maioria (nove artigos), elaborados por pesquisadores da área da Linguística ou de Letras.

Um dos trabalhos dentro dessa categoria, o de Borges e Pimentel (2019), da área de Letras, compara o DeepL e do Google Tradutor, fazendo uma avaliação humana do resultado da tradução de léxico especializado dessas duas ferramentas. Entretanto, o objetivo das autoras centra-se em avaliar a qualidade da ferramenta quando comparada às capacidades humanas e, como se poderia supor pela evolução das próprias ferramentas, atestaram que ambos os sistemas, DeepL e Google Tradutor, se tornam cada vez mais precisos e satisfatórios. Elas mencionam, apenas na última linha do trabalho, que os tradutores precisam de subsídios para escolher a ferramenta com melhor performance para seu trabalho.

A discussão de resultados, equivalência, imperfeições e incompatibilidades são primordiais para averiguarmos a qualidade das ferramentas e, mais do que isso, mostram a evolução dessas ferramentas ao longo dos anos. Em nossa opinião, porém, precisamos problematizar o fato de sete desses 12 artigos que tratam da avaliação de performance foram produzidos nos últimos 10 anos, três deles produzidos em 2019 e 2020.

Em nossa opinião, esses artigos alimentam a divergência entre a TA e tradutor humano, trazendo à tona a ideia de que os tradutores humanos serão, cedo ou tarde, substituídos pelas máquinas. Assim, a discussão não parece evoluir.

Podemos levantar a problemática de que, somente os linguistas computacionais – desenvolvedores dessas ferramentas, conhecem bem como elas estão respondendo a certos tipos de texto, discussão que poderia chegar aos tradutores se houvesse um diálogo interdisciplinar.

4.3 TA e a análise, apresentação ou comparação de uma ou mais ferramentas de TA

Na categoria “TA e a análise, apresentação ou comparação de uma ou mais ferramentas de TA” os artigos dedicam-se, como o nome já deixa antever, a: apresentar as ferramentas de TA, como, por exemplo, o Portuga, Déjà-vu e PorTal (esse último desenvolvido com o intuito de integrar ferramentas e recursos úteis para TA e o processamento bilíngue) ou recursos que podem ser acoplados a essas ferramentas; fazer comparações de performance entre as ferramentas de TA dependendo do tipo e base de dados (regras, estatística, neural), avaliando e indicando qual responde melhor a certos tipos de texto e alertam sobre suas incompatibilidades. Essa categoria se difere da anterior (categoria 4.3), pois o objetivo dos trabalhos não é fazer avaliações linguísticas, mas comparações relativas às estruturas das ferramentas e os resultados que exibem em relação aos tipos de textos.

Dos nove artigos que se encaixaram nessa categoria de análise de ferramentas, três são da área de Ciência da Computação e um da Linguística Computacional e se dedicam a apresentar as ferramentas, explicar como funcionam estruturalmente e suas funções, capacidades e seus problemas. Entretanto, essa discussão se volta aos desenvolvedores dessas ferramentas, não aos usuários. Dois dos artigos dessa categoria vêm da área da Linguística e trazem um olhar crítico sobre a performance das ferramentas em relação à abordagem estrutural, discutem quais são os melhores tipos de texto para cada tipo de sistema de TA e as dificuldades encontradas, demonstrando a necessidade de melhora em um ou outro aspecto da ferramenta avaliada. Os três artigos restantes são de pesquisadores inseridos na área dos Estudos da Tradução e se voltam para a tradução de textos de linguagem especializada em duas ou mais ferramentas de TA, apresentando essas ferramentas, comparando os resultados e alertando sobre incompatibilidades.

4.4 TA e o uso e a recepção dos usuários

Quanto aos trabalhos que tratam do uso e da recepção dos usuários, dois deles são de pesquisadores da área dos Estudos da Tradução e um da área de Letras. Eles averigam a recepção de textos traduzidos automaticamente por alunos de diversas áreas e o uso da TA por diversos setores da sociedade. No artigo que averigua o uso das ferramentas de TA em cenário nacional, quando comparado ao cenário internacional, Cleydstone dos Santos (2011) levanta os seguintes debates, que inclusive são bastante relevantes para a discussão que fazemos na seção 4.6:

Assim, observando algumas publicações sobre a TA no cenário nacional oriundas dos ETs, geralmente encontram-se comparações entre traduções humanas e traduções automáticas, ressaltando, em sua maioria, apenas as limitações da última em relação à “perfeição” da segunda, não estabelecendo relação direta com as áreas de geração da TA como: ciência da computação, inteligência artificial e linguística computacional. [...] Essa falta de diálogo entre os ETs e outras áreas de confluência com a TA é também criticada por Martins (2008) a partir de um estudo intitulado “Existem Máquinas de Traduzir?”, no qual é discutido o problema da tradução e seu lugar na máquina, levando em conta a importância de se dialogar com as áreas supracitadas. Como resultado das divergências de produção científica na comunidade acadêmica no cenário nacional, ora se observa um olhar crítico desconstrutivo, funcionando como entrave ao dinamismo do fluxo de recepção da TA, ora um olhar crítico construtivo no que se refere à compreensão das limitações e barreiras ainda enfrentadas pela TA com o intuito de serem superadas para o auxílio à tradução humana. (SANTOS, 2011, p. 172).

Como Santos (2011) argumenta, muito se fala sobre as limitações das ferramentas de TA ou sobre como elas respondem a um ou outro tipo de texto, mas, em cenário brasileiro, ainda existe a necessidade de diálogo entre os Estudos da Tradução e as áreas desenvolvedoras dessas ferramentas. Um exemplo prático seria, nas discussões do NILC, incluir também pesquisadores da área dos Estudos da Tradução, para que esses pudessem dialogar de forma a melhor compreender a evolução das abordagens, seus tipos, as bases de dados que utilizam e como os profissionais da tradução podem tirar proveito dessa tecnologia de maneira responsável e ética.

4.5 TA e o ensino de línguas estrangeiras

Outros três artigos tratam do uso da TA no ensino e aprendizagem de língua estrangeira, chamando atenção para a metodologia de dois deles, nos quais os pesquisadores apresentam textos e ferramentas de TA em sala de aula, buscando levantar discussões sobre seu uso. Um exemplo está nas discussões levantadas pela pesquisadora Sílvia Gusmão Sales (2011), autora de um desses três artigos:

Os dados que serão apresentados ao longo deste artigo vem mostrar, que em momento algum da História houve uma necessidade tão grande de superar as barreiras linguísticas que dividem os povos como agora, pois num mundo mais globalizado faz-se necessário que essas barreiras sejam superadas, para que o acesso às informações e transformações que as diversas áreas vêm sofrendo chegue ao conhecimento de todos. Sendo assim, dispor de meios alternativos que poderão proporcionar a otimização do tempo no processo de tradução, para dar conta do ritmo acelerado aos avanços tecnológicos que o homem vem impondo, e a produção de conhecimentos nas suas áreas de atuação, poderá ser um dos meios para tornar possível a quebra dessas barreiras. (SALES, 2011, p.20)

Nos artigos que relacionam a TA à aprendizagem de línguas estrangeiras, a utilização e importância da apresentação dessas ferramentas para os alunos aparecem em dois deles, um da

área de Linguística – que sugere o uso das ferramentas para o ensino de inglês instrumental, e outro dos Estudos da Tradução, o de Sales (2011) já mencionado, que tem como objetivo principal apresentar uma metodologia prática de exercícios de tradução utilizando TA junto a alunos do I e VII semestre de um curso de Letras com habilitação em português – inglês. Vale ressaltar a escassez de artigos que ensinem alunos de Tradução a utilizarem a TA. Como constatamos, os artigos que relacionam a “TA e aos Estudos da Tradução” (seção 4.6) trazem, em sua maioria, um histórico sobre a TA e uma hostilidade entre essa tecnologia e os tradutores humanos.

A Tabela 12 mostra a produção de cada área de pesquisa de acordo com as categorias de discussão.

Tabela 12 - Produção das áreas de pesquisa para cada categoria de discussão

Categoria de discussão por área de pesquisa	Contagem de Categoria de discussão
TA e o desenvolvimento, a arquitetura e/ou o funcionamento das ferramentas	15
Ciência da computação	6
Estudos da Tradução	2
Linguística	2
Linguística computacional	3
Não informado	2
TA e os estudos da tradução	13
Estudos da Tradução	7
Letras	2
Linguística	1
Linguística computacional	2
Linguística e ensino	1
TA e avaliação de performance e qualidade das ferramentas através de comparações linguísticas	12
Ciência da computação	1
Estudos da Tradução	2
Letras	3
Linguística	6
TA e a análise, a apresentação ou a comparação de uma ou mais ferramentas de TA	9
Ciência da computação	3
Estudos da Tradução	2
Línguas estrangeiras e Tradução	1
Linguística	2
Linguística computacional	1
TA e o uso e a recepção dos usuários	3
Estudos da Tradução	2
Letras	1

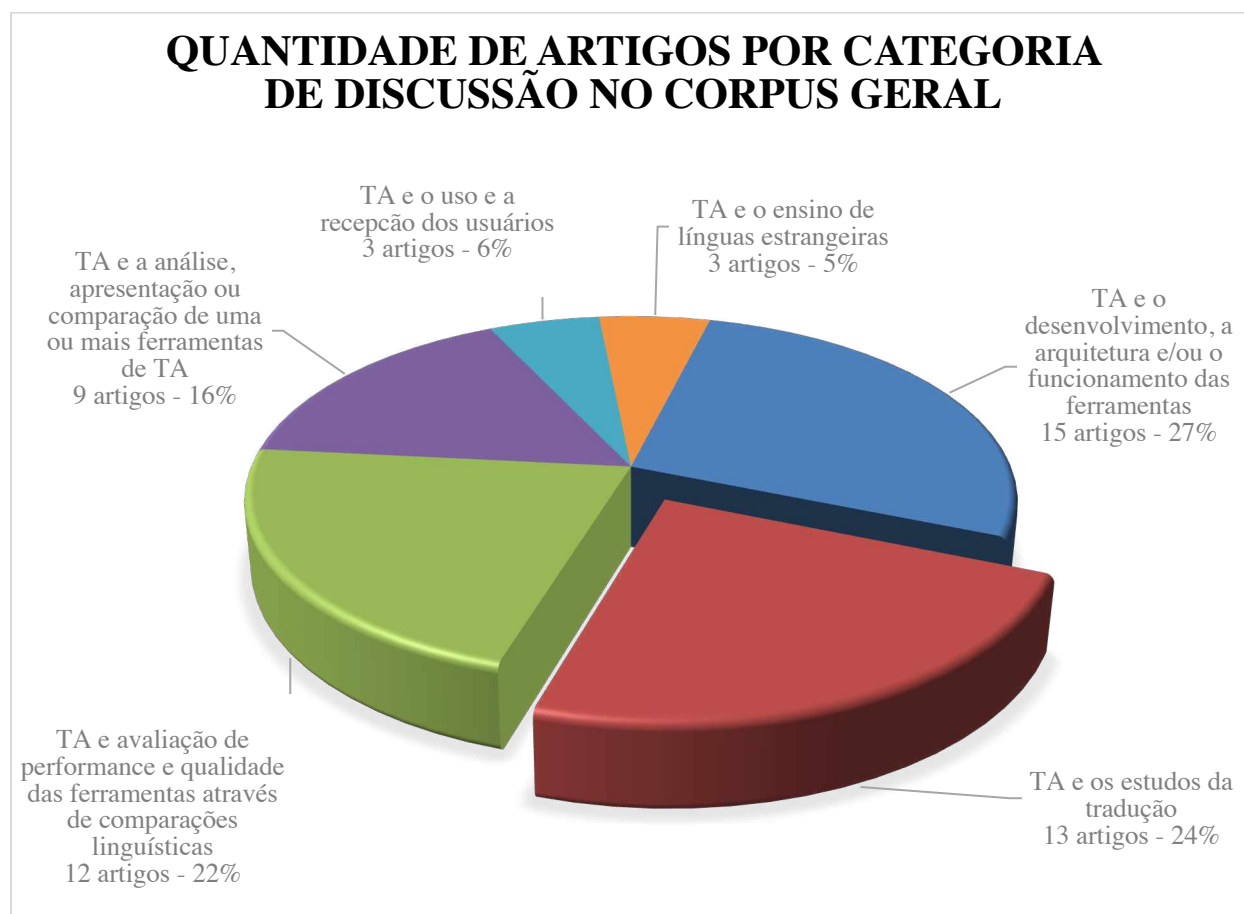
TA e o ensino de línguas estrangeiras	3
Estudos da Tradução	1
Letras	1
Linguística	1
Total Geral	55

Fonte: a autora.

Com relação ao desenvolvimento e à arquitetura das ferramentas de TA, podemos perceber que as pesquisas implementadas nas Ciência da Computação e Linguística Computacional dominam esse eixo. Nas discussões de “TA e Estudos da Tradução”, como era de se esperar, as áreas dos Estudos da Tradução, de Letras e Linguística são predominantes ao trazerem noções gerais de TA, análise teórica das práticas e problematizações do uso das ferramentas e a hostilidade entre TA versus tradutor humano. Os pesquisadores da área de Letras e Linguística se concentraram, em sua grande maioria, na avaliação e averiguação de performance através de comparações linguísticas. A área de Letras também se centra nos usos e na recepção da tecnologia de TA por falantes da língua do texto-alvo.

Assim, os 55 artigos foram analisados e categorizados de acordo com o tipo de discussão presente em seus objetivos, resumos e resultados. Com isso, mapeamos qualitativamente os principais eixos e tendências de discussões sobre Tradução Automática em 30 anos no Brasil. O Gráfico 5 representa visualmente o volume de artigos por categoria de discussões e com ele conseguimos vislumbrar o peso de cada uma delas em cenário brasileiro.

Gráfico 5 - Representação das categorias de discussões sobre TA no Brasil



Fonte: a autora.

Diante do que expomos, na próxima subseção, nos concentraremos naqueles 23,6% (13 artigos) que se encaixaram na categoria “TA e os Estudos da Tradução” para averiguar quais desses estudos se dispõem a relacionar a TA e o tradutor humano, e, mais que isso, se algum desses trabalhos orientam o uso das ferramentas de TA para a formação de tradutores no Brasil.

4.6 TA e os Estudos da Tradução

Podemos perceber que a TA em suas relações com os Estudos da Tradução estão entre os temas mais recorrentemente discutidos em nosso *corpus* de pesquisa. O Quadro 4 mostra o título dos 13 artigos, a área de conhecimento do pesquisador, um breve resumo das discussões e a data de produção de cada um deles. As discussões presentes nos artigos foram colocadas em forma de tabela para facilitar a relação título – discussão e para que a comparação das discussões ficasse visualmente mais fácil de acordo com a área de conhecimento e ordem cronológica.

Quadro 4: Relação dos artigos categorizados como “TA e os Estudos da Tradução” e respectivos anos de publicação

TA E Os Estudos Da Tradução	Área de conhecimento	Discussões abordadas	Ano
ALFARO, Carolina; DIAS, Carmelita: Tradução Automática: Uma Ferramenta De Auxílio Ao Tradutor	Letras	Alfaro e Dias discutem o panorama da TA e das tecnologias da tradução em 1998, contextualizam e apresentam as ferramentas de auxílio ao tradutor, tais como os sistemas de memória de tradução e as ferramentas de TA e discutem a decepção que alguns têm com os resultados das ferramentas, enfatizando a importância de seus produtos para o aperfeiçoamento do trabalho do tradutor, não uma competição entre máquina e homem.	1998
DARIN, Leila Cristina de Mello: O Anseio Da Língua-Mãe Nas Entrelinhas Da Tradução Automática	Estudos da Tradução	Darin relaciona o uso da TA ao anseio de uma língua mãe; uma língua franca através da qual haveria a possibilidade de comunicação universal que a TA proporciona através de um clique de botão; uma promessa de remoção das barreiras linguísticas impostas pela multiplicidade de idiomas. Darin nega os avanços da TA e ignora o uso dessas ferramentas que seria, para a autora, o equivalente a construir uma nova torre de Babel.	2001
MARTINS, Ronaldo T.: Tradução Automática E Estudos Da Tradução: Um Conflito Paradigmático	Linguística Computacional	Martins realiza uma análise teórica das práticas da TA, comparando as duas “áreas”: Tradução automática e os Estudos da Tradução. Discute a tentativa da TA em resumir a língua em códigos, teoria que contraria as teorias dos Estudos da Tradução, que por sua vez busca comprovar que o texto é um fenômeno que ultrapassa a linguagem. Indaga se esse não é o motivo de fracasso da TA até aquele momento, isto é, considerar o texto apenas como código e desconsiderar o fenômeno do sentido.	2003
WEININGER, Markus J.: TM & MT na Tradução Técnica Globalizada – Tendências E Consequências.	Estudos da Tradução	Weininger revisa algumas consequências da aplicação da TA e dos SMT na área de tradução técnica e as tendências de desenvolvimento no campo. Previa que, em 2004, as memórias de tradução seriam incorporadas e utilizadas em maior escala para a área de textos técnicos.	2004
MARTINS, Ronaldo T.; NUNES, Maria das Graças V.: Noções Gerais De Tradução Automática	Linguística Computacional	Martins e Nunes trazem noções gerais da TA. O artigo é composto de uma grande introdução histórica da TA, os parâmetros ou abordagens para a construção das ferramentas e uma explicação técnica do funcionamento das ferramentas. Uma discussão voltada aos desafios e questionamentos daqueles que constroem as ferramentas.	2005
STUPIELLO, Erika Nogueira de: O Texto Adaptado À Máquina: Estratégias De Controle Autoral Para Implementação Da Tradução Automática	Estudos da Tradução	Stupiello apresenta um breve histórico da questão da automação em tradução, introduz o conceito de léxico e ambiguidade e enfoca algumas estratégias de controle lexical. Discute as estratégias de controle de uso do léxico dos textos de origem (frases mais curtas, diretas e sem ambiguidade) como forma de garantir o desempenho da automação, do trabalho de revisão e pós-edição.	2010
CERQUEIRA, Sara: A Nova Torre De Babel – Que Futuro Para A Tradução Automática?	Estudos da Tradução	Cerqueira problematiza o futuro da TA e discute que as críticas à TA são desnecessárias e que só nos faz “virar as costas” para uma realidade existente e capaz de ajudar o tradutor humano. Ver a TA como ameaça só encurta os horizontes do conhecimento e também	2011

		da capacidade do tradutor de se adequar às novas tecnologias. Levanta a discussão de requalificação e reorientação de formação da profissão de tradutor quanto às novas tecnologias, ou seja, uma atualização do tradutor em relação ao uso da TA.	
SOUZA, Cristiane de: A Tradução Automática E Os Seus Limites	Estudos da Tradução	Souza visa demonstrar os limites da TA e o quanto seus produtos estão aquém dos níveis que se esperam de uma tradução de qualidade. A autora estuda: textos abreviados e submetidos aos tradutores automáticos, buscando enfatizar os problemas e equívocos encontrados nos produtos das ferramentas de TA. Mesmo enfatizando a deficiência das ferramentas e a necessidade de intervenção humana no resultado das ferramentas, o artigo defende a necessidade de aliar a TA aos esforços dos tradutores profissionais.	2011
COSTA, Maria José D., ZIPSER, Meta. E., & POLCHLOPEK, Silvana. A.: (Re)Conhecimento Da Tradução Em Sala De Aula: Sobre Uma Experiência Prática Com Tradutor Automático <i>Online</i>	Linguística e Ensino	Costa, Zipser e Polchlopek discutem sobre a prática tradutória usando ferramentas de TA voltada para o ensino de tradução. O estudo evidencia o uso das ferramentas de TA em uma atividade realizada em uma disciplina de prática da tradução com alunos de graduação em Letras da UFSC. As autoras ressaltam que apresentar as ferramentas aos alunos possibilita a demonstração empírica das deficiências, dos tipos de estratégia para conseguir extrair o máximo das ferramentas e da competência tradutória necessária para utilizá-las.	2012
MELO, Sheila de S.: Tradução Automática E Competência Tradutória: Repensando Interseções	Estudos da Tradução	Melo evidencia a dualidade do uso de TA: ameaça próxima ou ferramenta infrutífera e obsoleta. Utilizando a noção de competência tradutória proposta pelo grupo PACTE, o artigo faz uma abordagem reticular entre a emergência da TA, a formação e qualificação de tradutores e necessidades globais e do mercado de tradução.	2013
CARNEIRO, Marcos de C.; FERNANDES, Lincoln P.: Uma Breve Reflexão Historiográfica Sobre A Tradução Automática	Estudos da Tradução	Carneiro e Fernandes constroem uma historiografia da TA e tratam da rivalidade humano <i>versus</i> máquina, explorando o Modelo de Tradução Tripartido de Eugene Nida e a Abordagem Trifásica das ferramentas de TA. Consideram a evolução da TA e propõe uma interpretação das influências e das afinidades observadas entre os processos de tradução humana e TA.	2015
COSTA, Maria José D.; DEMÉTRIO, Ana Paula de C.: O Enfoque Da Sequência Didática Nas Reflexões Sobre Língua E Cultura No Processo Tradutório: O Uso Do Tradutor Automático Como Atividade Diagnóstica	Linguística	Costa e Demétrio mostram o resultado do uso das ferramentas de TA como estratégia para o desenvolvimento do processo tradutório e de produção textual com alunos do curso de Letras na disciplina de Tradução e Ensino de Língua Estrangeira.	2018
FONSECA, Norma B. de L.: Investigando Adjetivos Atributivos Em Traduções Publicadas, Em Um Sistema De Tradução Automática E Na Pós-Edição Monolíngue: Contribuições Para A Pedagogia Da Tradução	Letras	Fonseca investiga o uso de adjetivos atribuídos em traduções publicadas e em tecnologias de TA. Uma parte do estudo se volta a expor para os estudantes de tradução as dificuldades das ferramentas, mostra a necessidade de ensinar estratégias para solucionar essas dificuldades e de familiarizar os alunos com as ferramentas de TA, para que eles se sintam mais seguros ao traduzirem usando-as.	2018

Fonte: a autora.

Com base no Quadro 4, verificamos que as primeiras discussões dos Estudos da Tradução em relação à TA se voltavam para contextualização do que é a Tradução Automática, trazem uma historiografia mais robusta sobre essa e outras tecnologias da tradução, tais como os sistemas de memória de tradução. Como as discussões sobre TA chegaram ao Brasil quando as ferramentas de TA já estavam sendo criadas e utilizadas há mais tempo em outros países, mesmo nos trabalhos iniciais já havia discussões sobre a necessidade de apresentar e incorporar essas ferramentas ao trabalho de tradutor. Porém, sequencialmente, as discussões se voltam à hostilidade entre o tradutor humano e a máquina. Ainda, as discussões mostram os Estudos da Tradução como campo teórico que não aceita as ferramentas de TA, pois a veem como uma ameaça. A pergunta que fica é: os Estudos da Tradução como campo disciplinar deveriam “brigar” com as suas próprias tecnologias? Já os trabalhos mais recentes levantam a discussão da requalificação e reorientação da formação do tradutor quanto às novas tecnologias.

Como antes mencionamos, nas discussões sobre “TA e Estudos da Tradução” as áreas às quais pertencem os pesquisadores são: Estudos da Tradução, Letras e Linguística, porém as áreas se distinguem no teor de suas discussões. Os trabalhos oriundos das áreas de Letras e Linguística se mostram mais receptivos quanto ao uso da TA no ensino de línguas.

A autora Maria José Damiani Costa, em 2012 e em 2018, produz artigos que destacam a necessidade do ensino de línguas e tradução a partir de ferramentas de TA.

Apesar de utilizar as ferramentas de TA para a produção textual com alunos do curso de Letras na disciplina de Tradução e Ensino de Língua Estrangeira, e não especificamente na formação de tradutores, Costa e Demétrio (2018), sob a ótica funcionalista, trazem a TA além do certo ou errado, do equivalente ou não equivalente e localizam o uso atual e inerente dessa ferramenta:

Dentro do ambiente de ensino de línguas, sabe-se que o tradutor automático é um instrumento utilizado pelos alunos como primeiro negociador de sentidos, atuando como suporte para as mais diversas atividades de tradução, produção e compreensão textual. O uso de tal recurso, no entanto, embora aplicado com frequência pelos alunos, muitas vezes é ignorado ou condenado pelo professor e a ferramenta acaba sendo rotulada como algo que deva estar à margem da sala de aula, visto que as traduções realizadas pela máquina nem sempre são aceitáveis e, assim, acabam sendo prejudiciais no processo de produção escrita em LE por parte dos alunos. (COSTA e DEMÉTRIO, 2018, p. 182)

Nos cursos de formação de tradutores acontece algo semelhante, isto é, o uso das ferramentas já se tornou inerente, sobretudo na última década, em que os alunos fazem uso dos tradutores automáticos gratuitos em sala de aula, e o fato desse uso ser ignorado ou, por vezes, condenado pelos professores, coloca a TA à margem do ensino de tradução, contribuindo para sua marginalização. Embora a aceitação e uso da TA para o trabalho do tradutor venham sendo discutidos desde 2000, em sala de aula isso parece não acontecer mediante o que mostram os trabalhos aqui discutidos. No entanto, muitas discussões relevantes sobre a TA podem fazer parte da sala de aula de ensino de tradução, tais como, ética tradutória, proteção de dados, volume de traduções incompatível com a quantidade de tradutores no mercado, tradutores automáticos acoplados a websites, presença das ferramentas de TA em órgãos públicos e privados importantes, importância da TA para a economia e websites de venda internacionais, tipos de texto que necessitam ou não do tradutor humano, entre outros. Precisamos aprender como ensinar o tradutor em formação a utilizar as ferramentas de TA como um auxílio, sobretudo para nos livrarmos do paradoxo “TA “Amável amiga, terrível ameaça” (MELO, 2013).

Podemos perceber esse paradoxo em dois artigos do mesmo ano discutindo “TA e os Estudos da Tradução”, com duas visões totalmente diferentes. Um deles, o de Souza (2011), coloca as tecnologias de TA e os tradutores em lados opostos:

A verdade, porém, é que a chamada tradução automática — também conhecida como “tradução instantânea”, “tradutor automático”, etc. — está muito aquém do nível que se espera de traduções de qualidade, e é uma ingenuidade comparar esse tipo de serviço com os serviços de um tradutor humano. (SOUZA, 2011, p.1)

Ao mesmo tempo, Cerqueira (2011) aponta que:

A proliferação do acesso a tradutores automáticos disponíveis na Internet conduziu, por um lado, à disseminação da obtenção de informação e obriga, por outro, a uma requalificação e re-orientação de formação da profissão de tradutor, que terá, necessariamente de passar por uma abordagem destas novas tecnologias. Este imperativo nem sempre é bem aceito pelos profissionais de tradução. Considerar, porém, a TA utópica, arrogante e insolente só porque os resultados que se obtiveram são nitidamente incipientes e até mesmo anedóticos, não nos deve fazer voltar as costas a uma realidade que não só existe, como tem também provas dadas em vários campos. Ver a TA como uma ameaça para o futuro dos tradutores e a virtualização do trabalho de tradução como um monstro desumanizador só encurta os horizontes de conhecimento e, logo, uma das qualidades imperativas de um bom tradutor seria a assimilação de conteúdos e a adaptação a uma realidade em constante mudança. (CERQUEIRA, 2011, p. 160)

Assim como Cerqueira (2011) aponta, ignorar ou banalizar o uso das tecnologias de tradução é desnecessário e retrógrado. Atualmente, o tradutor deve lidar não somente com os prazos

e pressões editoriais, mas também com questões práticas como formatação, tradução de imagens, formato de arquivos, entre outros. A exposição a uma metodologia prática em relação ao uso das ferramentas de TA, bem como outras tecnologias, beneficiará o tradutor, possibilitando uma automatização do trabalho de tradução, acarretando um crescimento exponencial da qualidade e eficiência de seus serviços.

E quanto as limitações, Caseli (2017), que é uma das pesquisadoras mais produtivas desse *corpus* nos explica que são justamente as limitações da TA que impulsionam a continuação de suas pesquisas:

A tradução da expressão em português “loira gelada”, tradicionalmente empregada para se referir a uma “cerveja gelada”, foi equivocadamente traduzida em inglês para *cold blonde* pelo Google Translator. Esse exemplo demonstra uma das principais limitações dos sistemas automáticos de TA, que é lidar com gírias e coloquialismos. Essas limitações, aliadas à importância cada vez maior dos sistemas de TA em um cenário mundial com cada vez menos barreiras físicas e virtuais entre as pessoas falantes de diversos idiomas, fomentam as pesquisas atuais em TA. (CASELI, 2017, p. 1783)

Entender essas limitações, conhecer o desempenho da TA e compreender que ela traduz e de forma satisfatória o léxico controlado (frases mais curtas, diretas e sem ambiguidade, ou pré-editadas), garantindo assim o bom desempenho da automação, do trabalho de revisão e pós-edição, como ressalta Stupiello (2010), podem ajudar desmistificar a TA e evitar comparações de resultados tradutor humano *versus* máquina.

Por isso, a necessidade de utilizarmos as ferramentas, entendermos seu funcionamento e conhecermos suas limitações e situações de uso parece urgente. Segundo Costa e Demétrio (2018) afirmam:

O tradutor automático pode auxiliá-los (os tradutores) na construção de um texto, porém deve-se tomar essa produção como uma produção inicial – primeiro rascunho – pois a ferramenta apresenta limites de ordem linguística e gramatical. Por outro lado, traz à tona, também, os limites da máquina, sobre as questões que sustentam a situação de comunicação: quem são os interlocutores?; qual o propósito?; os conhecimentos prévios sobre o tema?; há adequação ao gênero textual?; dentre outros elementos extratextuais. Os dois momentos conscientizam o aluno que o tradutor automático é uma boa ferramenta e pode ser utilizado em sala de aula, porém que necessita de sua mediação. O aluno é autor capaz de reelaborar o texto, trazer para sua construção de sentido as pressuposições de seu leitor meta e os seus possíveis movimentos sócio, histórico e culturais. (COSTA; DEMÉTRIO, 2018, p.194)

Também segundo Melo (2018):

Construir reflexões em torno da tradução automática na era das competências tradutórias implica um esforço para inseri-la, aquém das imagens já consolidadas sobre sua fragilidade, no escopo das tecnologias de apoio à tradução. Como toda ferramenta, a TA apresenta limitações, mas o fracasso do empreendimento não está relacionado à sua

natureza ou funcionamento unicamente. Antes, seguindo as considerações do PACTE, o sucesso depende da articulação da ferramenta com outras diversas (ferramentas no aspecto de produtos materiais e intelectuais) por meio de estratégias estabelecidas pelo tradutor no seu projeto de tradução. (MELO, 2018, p. 15)

Apesar de os Estudos da Tradução serem a área de conhecimento que mais discute a “TA e Estudos da Tradução”, compreendendo 7 artigos dos 13, eles se voltam mais a questões históricas, técnicas e de hostilidade envolvendo a TA e os tradutores humanos. Não parece haver uma aceitação dessa ferramenta. Isso pode ser justificado com um dos textos de Ronaldo T. Martins (2003), que discute o conflito paradigmático entre a Tradução Automática e Estudos da Tradução: “*a consideração do atual estado da arte dos Estudos da Tradução introduz uma provocação pelo menos interessante para a Tradução Automática*” (MARTINS, 2003, p. 10), pois a última, segundo o autor, considera a língua somente como códigos, invisibilizando os fenômenos de sentido que são constitutivos da própria linguagem. Ou seja, para Martins (2003), os Estudos da Tradução estão atrelados ao sentido, à arte de traduzir, ao passo que a TA é basicamente o processamento automático dos códigos das línguas naturais.

Para finalizarmos esta seção, pensamos que enquanto esses dois “vetores” (TA e os Estudos da Tradução) não deixarem de ser perpendiculares e passarem a ser paralelos, ou seja, “andarem juntos” em direção ao futuro, um impedirá o outro de alcançar seus objetivos.

CONSIDERAÇÕES FINAIS

Nossa dissertação teve como proposta investigar, por meio de indicadores cienciométricos, aliados a um viés qualitativo, a produção científica sobre tradução automática em cenário brasileiro. Buscamos demonstrar quais discussões vêm sendo realizadas no país, nos últimos 30 anos – *i.e.*, de 1990 a 2020, e o impacto de tais discussões para a formação de tradutores brasileiros.

Coletamos os artigos de periódicos brasileiros dos Estudos da Tradução que tratavam do tema “Tradução Automática”, resultando em 17 artigos ao longo de todas as edições de 14 periódicos verificados. Em virtude do número reduzido de publicações, o cenário de busca foi expandido e as mesmas palavras-chave foram submetidas à plataforma de busca Google Acadêmico, compondo um *corpus* de pesquisa com 87 publicações (estado da arte). Além da pesquisa quantitativa, levamos em consideração o conteúdo tratado nesse *corpus* com relação à TA na perspectiva brasileira (estado da questão).

Pensamos ser importante, antes de concluirmos nossa dissertação, ressaltar algumas sugestões importantes para quem pretende fazer pesquisas baseadas em estudos bibliométricos e

cienciométricos, para que seja possível estabelecer os critérios que necessitam de maior atenção na hora de estabelecer as análises.

O primeiro ponto é a questão temporal: ao delimitar o recorte de tempo, há, conseqüentemente, uma perda da evolução histórica de determinada temática. Embora não tenhamos delimitado o ano de publicação nas bases de dados investigadas, tudo o que foi discutido nesta dissertação deve ser visto dentro daquela realidade temporal. Além disso, ao fazer uma análise usando meios tecnológicos, nós, pesquisadores, deixamos de contemplar as publicações impressas. Um desafio desse tipo, por exemplo, está relacionado ao periódico *Tradução & Comunicação – Revista Brasileira de Tradutores*, que de 1981 a 2006 apenas publicou seus trabalhos na versão impressa.

O veículo de publicação pode também influenciar nas análises bibliométricas e cienciométricas. Questões econômicas, institucionais e ideológicas podem interferir na publicação científica de uma área. No Brasil, não se tem o costume de publicar artigos em periódicos pagos, entretanto, ao fazer uma análise cienciométrica em publicações internacionais, a questão econômica pode ser um fator impactante, já que grande parte dos periódicos mais tradicionais dos Estudos da Tradução são pagos. Que tipo de investimento haveria nesse tipo de publicação?

O idioma da publicação e a variação de terminologia são importantíssimos para os estudos bibliométricos e cienciométricos, pois os termos específicos de determinada área de conhecimento que será pesquisada devem ser utilizados como palavras-chaves no campo de busca nas bases de dados. No caso de uma pesquisa feita utilizando-se as palavras-chave em inglês, a abrangência é infinitamente maior, entretanto, fica inviável a análise por localidade. Se o objetivo da pesquisa é determinar o mapa cienciométrico de um país, como o Brasil, o idioma de publicação é crucial. Assim, determinando um idioma específico as ferramentas buscarão as palavras-chave nele escolhidas. Mas, se o idioma utilizado for compartilhado por mais de um país, os desafios de se analisar as produções de um país só tendem a aumentar, já que as bases de dados podem recuperar, em nosso caso, pesquisas realizadas em português de Portugal, apenas para mencionarmos um único país de fala portuguesa.

Outro ponto de atenção está relacionado ao fato de que as palavras-chave nem sempre se relacionam com os assuntos da pesquisa, pois podem haver erros ou coincidências terminológicas e a ferramenta buscar esses resultados também. Ambos os casos só são possíveis de identificação através de uma pesquisa qualitativa, com manipulação manual de cada documento indicado pela ferramenta e averiguando-se as discussões presentes no documento, o local de publicação e a instituição de ensino provedora.

Depois de estabelecer um *corpus* coerente, respeitando os pontos de atenção ora apontadas, o formato das informações bibliométricas e cienciométricas pode ser desafiador. Nem toda publicação tem o *Digital Object Identifier* (DOI), principalmente aquelas mais antigas. Com base nesse fato, todo o *corpus* precisa ser colocado em uma ferramenta que organize as referências e gere um documento no formato RIS, que possa ser analisado por outra ferramenta que visualize as similaridades e crie os mapas. O formato RIS não possibilita análises tão amplas, e ainda pode existir disparidade de formato entre as ferramentas.

A curva de aprendizagem das ferramentas também deve ser considerada e quando utilizadas devem ser exploradas em todo seu potencial. O pesquisador emprega tempo para encontrar as ferramentas apropriadas para cada etapa da análise, entender os procedimentos de uso, formatos de documentos, todas as funções de análise das ferramentas e lidar com as instabilidades de cada uma delas – incompatibilidade de formato dos documentos entre as ferramentas, incompatibilidade das ferramentas com o processador utilizado pelo pesquisador, informações que precisam ser colocadas manualmente, pois a ferramenta nem sempre consegue ler todas as informações bibliométricas e cienciométricas, entre outros entraves tecnológicos, que cabem ao pesquisador detectar com olhos atentos.

Outro ponto que precisa de atenção é a generalização dos resultados. As ferramentas de análise quantitativa informam resultados gerais a partir dos documentos analisados. Os fatores específicos devem ser estudados detalhadamente pelo pesquisador. Os resultados sobre as obras mais citadas também podem induzir a suposições que não demonstram a realidade. E nem sempre o autor mais citado é o autor mais renomado dentro da área de estudo, pois ele pode estar sendo citado pelas deficiências de sua teoria, ou como foi visto nessa pesquisa, autores do mesmo grupo de pesquisa que se citam dentro de uma área com pouco avanço nas discussões.

Portanto, o que se pode fazer para maximizar a precisão dos estudos cienciométricos é entender as exclusões e os efeitos que esses desafios podem trazer e tentar amenizá-los, delimitando minuciosamente os critérios de análise. Nossa dissertação, por exemplo, verificou apenas publicações de livre acesso, cobriu a maior variação terminológica possível, delimitou um idioma em um âmbito de publicação e, principalmente, buscou analisar especificamente cada resultado geral oferecido pelos gráficos, dados bibliométricos e mapas cienciométricos.

Os dados qualitativos colhidos e analisados nos permitiram desenhar um padrão de discussões por meio das categorias que delimitamos, que varia em termos de áreas de pesquisa e de evolução temporal. Entretanto, devemos ter em mente que, no Brasil, os programas de pós-graduação podem se inserir em institutos de maior abrangência, como, por exemplo, Institutos de

Letras e Linguística, que contém programas voltados para os Estudos da Tradução, o que pode explicar porque a área de Linguística se fez tão presente em todas as categorias de discussões presentes no *corpus*.

Os pesquisadores das áreas de Linguística Computacional e Ciência da Computação, durante a evolução da tecnológica da TA, empenharam-se em investigar a arquitetura, o desempenho, a equivalência e qualidade dos produtos oriundos da TA. Com a popularização de algumas ferramentas, arquiteturas mais avançadas e resultados cada vez mais satisfatórios, outros pesquisadores da Linguística e dos Estudos da Tradução se esforçaram para mostrar as limitações da TA, que as ferramentas não são capazes de traduzir todos os tipos de texto e que não podem substituir o tradutor humano, alertando para os perigos de confiar cegamente em uma tradução de máquina. Com a evolução das tecnologias e a inserção dessas ferramentas no cotidiano de quem trabalha com línguas, houve estudos que aplicavam a TA no ensino e aprendizagem de língua estrangeira, porém o uso da TA no ensino de tradução ainda se mostra tímido.

De nossa parte, buscamos, com esta dissertação, contribuir para as áreas que discutem TA – Ciência da Computação, Estudos da Tradução, Letras, Linguística e Linguística Computacional – fornecendo dados bibliométricos e cienciométricos, seguidos de uma análise qualitativa dos discursos do ano de 1990 a 2020. Com os resultados aqui obtidos conseguimos traçar uma progressão de discussões, mapear os polos de discussão no Brasil, compreender o que discutimos ao longo dos anos e no que devemos nos debruçar para progredirmos cientificamente, fazendo com que a TA seja um produto voltado ao tradutor e não o inverso.

Assim, com estudos mais aprofundados sobre a TA e seus efeitos para a prática profissional em tradução, os tradutores profissionais e também os que estão em formação, poderão não ter tantas dúvidas sobre seu uso, como no exemplo dado na Introdução desta dissertação, em que um tribunal polonês estabeleceu uma série de regras referentes ao uso de ferramentas gratuitas de tradução automática no contexto de direitos de propriedade intelectual, depois que um tradutor e a agência para a qual ele trabalhava submeteram a tradução de um livro inteiro a uma ferramenta de tradução automática, o Google Tradutor, infringindo as regras de direitos autorais.

(Des)automatizando seus discursos, a TA pode se tornar “amável amiga, incrível aliada”!

REFERÊNCIAS DOS PERIÓDICOS DO *CORPUS*

BELAS INFIÉIS: Revista do Programa de Pós-Graduação em Estudos de Tradução da Universidade de Brasília. Brasília, 2012-. Semestral. ISSN 2316-6614. Versão *online*. Disponível em: <https://periodicos.unb.br/index.php/belasinfieis/index>. Acesso em: 14 março 2022.

CADERNOS DE LITERATURA EM TRADUÇÃO. São Paulo. Anual. ISSN: 1981- 2558. Versão *online*. Disponível em: <https://www.revistas.usp.br/clt>. Acesso em: 14 março 2022.

CADERNOS DE TRADUÇÃO. Periódico do Instituto de Letras da Universidade Federal do Rio Grande do Sul. Porto Alegre, 1997- Semestral. Versão *online*. Disponível em: <https://seer.ufrgs.br/cadernosdetraducao>. Acesso em: 14 março 2022.

CADERNOS DE TRADUÇÃO. Revista da Pós-Graduação em Estudos da Tradução da Universidade Federal de Santa Catarina. Florianópolis, 1996- Quadrimestral. ISSN 2175-7968. Versão *online*. Disponível em: <https://periodicos.ufsc.br/index.php/traducao>. Acesso em: 14 março 2022.

CULTURA E TRADUÇÃO. Publicação do Programa de Pós-graduação em Letras da Universidade Federal da Paraíba. João Pessoa. 2011-. Irregular. ISSN: 2238-9059. Versão *online*. Disponível em: <https://periodicos.ufpb.br/index.php/ct/index>. Acesso em: 14 março 2022.

IN-TRADUÇÕES: Revista do Programa de Pós-Graduação em Estudos da Tradução da UFSC. Florianópolis, 2009-2015. Semestral. ISSN 2176-7904. Versão *online*. Disponível em: <http://incubadora.periodicos.ufsc.br/index.php/intraducoes>. Acesso em: 14 março 2022.

RÓNAI – Revista de Estudos Clássicos e Tradutórios da Faculdade de Letras (FALE) da Universidade Federal de Juiz de Fora. Juiz de Fora, 2006-. Semestral. ISSN 2318-3446. Versão *online*. Disponível em: <https://periodicos.ufjf.br/index.php/ronai/index>. Acesso em: 14 março 2022.

TRADTERM. Publicação do Centro Interdepartamental de Tradução e Terminologia. São Paulo, 1994- . Semestral. ISSN: 2317-9511. Versão *online*. Disponível em: <https://www.revistas.usp.br/TradTerm>. Acesso em: 14 março 2022.

TRADUÇÃO E COMUNICAÇÃO: Revista Brasileira de Tradutores da Anhanguera Educacional S.A. Valinhos. 2010- . Anual. ISSN:0101-2789. Versão *online*. Disponível em: <https://diretorio.rcaap.pt/handle/1/278>. Acesso em: 14 março 2022.

TRADUÇÃO EM REVISTA. Periódico eletrônico da área de Estudos da Tradução da Pontifícia Universidade Católica do Rio de Janeiro. Rio de Janeiro, 2004- . Semestral. ISSN 1808-6195. Versão *online*. Disponível em: https://www.maxwell.vrac.puc-rio.br/rev_trad.php?strSecao=sobre&fas=&menufas=1. Acesso em: 14 março 2022.

TRADUZIREs. Revista do Programa de Pós-Graduação em Estudos da Tradução da Universidade de Brasília. Brasília, 2012- . Semestral. ISSN 2238-7749. Versão *online*. Disponível em: <https://periodicos.unb.br/index.php/traduzires/index>. Acesso em: 14 março 2022.

TRANSLATIO - Revista do Núcleo de Estudos de Tradução Olga Fedossejeva Universidade Federal do Rio Grande do Sul. Porto Alegre, 2011- . Semestral. ISSN 2236-4013. Versão *online*. Disponível em: <https://seer.ufrgs.br/translatio>. Acesso em: 14 março 2022.

TRANSVERSAL – Revista em Tradução do Departamento de Letras Estrangeiras da Universidade Federal do Ceará. Fortaleza, 2015- . Anual. ISSN 2446-8959. Versão *online*. Disponível em: <http://periodicos.ufc.br/transversal>. Acesso em: 14 março 2022.

SCIENTIA TRADUCTIONIS. Periódico acadêmico digital vinculado à Pós-Graduação em Estudos da Tradução da Universidade Federal de Santa Catarina. Florianópolis, 2005-2014. ISSN 1980-4237. Semestral. Versão *online*. Disponível em: <https://periodicos.ufsc.br/index.php/scientia>. Acesso em: 14 março 2022.

REFERÊNCIAS

ALFARO, C.; DIAS, M. C. P. Tradução Automática: uma ferramenta de auxílio ao tradutor. **Cadernos de Tradução**, v. 3, n. 1, p. 369-390, 1998. Florianópolis: UFSC/PGET. Disponível em: <https://periodicos.ufsc.br/index.php/traducao/article/view/5392/4936>. Acesso em: 9 set. 2022.

ALVES, D. A. de S.; VASCONCELLOS, M. L. Metodologia de pesquisa em Estudos da Tradução: uma análise bibliométrica de teses e dissertações produzidas no Brasil entre 2006-2010. **D.E.L.T.A.** v. 32, n. 2, p. 375-404, 2016.

ALVES, S. K. **Sistemas de memória de tradução e corpora: integração que cria novas possibilidades ao tradutor**. Monografia (Graduação em Tradução) – Instituto de Letras e Linguística da Universidade Federal de Uberlândia. Uberlândia, 2018.

BORGES, T. M.; PIMENTEL, J. M. M. Avaliação humana da tradução automática de combinações lexicais especializadas: o caso do Google Translate e do DeepL. **Belas Infiéis**, Brasília, Brasil, v. 9, n. 4, p. 21–43, 2020. DOI: 10.26512/belasinfiéis.v9.n4.2020.26755. Disponível em: <https://periodicos.unb.br/index.php/belasinfiéis/article/view/26755>. Acesso em: 9 set. 2022.

CAPES. **História e Missão**. 2022. Disponível em: <https://www.gov.br/capes/pt-br/aceso-a-informacao/institucional/historia-e-missao/>. Acesso em: 20 fev. 2022.

CARNEIRO, M. C., FERNANDES, L. P. Uma breve reflexão historiográfica sobre a tradução automática. **Revista Belas Infiéis**, v. 4, p. 159-168, 2015. Disponível em: <https://periodicos.unb.br/index.php/belasinfiéis/article/view/11318/9955>. Acesso em: 9 set. 2022.

CASELI, H. D. M. Tradução Automática: estratégias e limitações. **Domínios de Linguagem**, v. 11, n. 5, p. 1782-1796, 2017. Disponível em: <https://biblat.unam.mx/es/revista/dominios-de-linguagem/articulo/traducao-automatica-estrategias-e-limitacoes>. Acesso em: 14 março 2022.

- CERQUEIRA, S. A nova Torre de Babel – que futuro para a tradução automática? **POLISSEMA– Revista de Letras do ISCAP**, v.3, p.159-167, 2003. Disponível em: <https://parc.ipp.pt/index.php/Polissema/article/download/3405/1390>. Acesso em: 14 março 2022.
- COSTA, M. J. D., & DEMÉTRIO, A. P. C. O enfoque da sequência didática nas reflexões sobre língua e cultura no processo tradutório: O uso do tradutor automático como atividade diagnóstica. **Revista académica liLETRAd**, v. 4, p. 181-196, 2018. ISSN: 2444-7439.
- COSTA, M. J. D.; ZIPSER, M. E.; POLCHLOPEK, S. A. (Re)Conhecimento da tradução em sala de aula: sobre uma experiência prática com tradutor automático online. **Linguagem & Ensino**, v. 15, p. 365-386, 2012. Disponível em: <https://periodicos.ufpel.edu.br/ojs2/index.php/rle/article/view/15426>. Acesso em: 9 set. 2022.
- DARIN, Leila Cristina de Mello. O anseio da língua-mãe nas entrelinhas da tradução automática. **Galáxia**, n. 2, 2001, p. 149-167. Disponível em: <https://revistas.pucsp.br/index.php/galaxia/article/view/1236/743>. Acesso em: 9 set. 2022.
- DE SOUZA, C. A TRADUÇÃO AUTOMÁTICA E OS SEUS LIMITES. **Revista Científica Linkania Júnior**, v. 1, n. 1, 2011. Disponível em: <https://linkania.org/junior/article/view/13>. Acesso em: 14 março 2022.
- DEEPL. **Por que o DeepL?** 2022. Disponível em: <https://www.deepl.com/pt-BR/whydeepl/>. Acesso em: 24 ago. 2022.
- DOHERTY, S. et al. On Education and Training in Translation Quality Assessment. *In*: MOORKENS, J.; CASTILHO, S; GASPARI, F; DOHERTY, S. (eds.). **Translation Quality Assessment: From Principles to Practice**. Machine Translation Series. Switzerland: Springer International Publishing, 2018, p. 95-106.
- ECK, N. J. V. **Methodological Advances in Bibliometric Mapping of Science**. Rotterdam, Holanda: Erasmos University, 2011.
- ESQUEDA, M. D. & DE JESUS, S.M. Scientific and Technical Translation in Brazilian Journals: A benchmark literature review. **New Voices in Translation Studies**, n. 13, p. 79-101, 2015.

ESQUEDA, M. D. (Org.). **Estudos bibliométricos e cientométricos em Tradução: tendências, métodos e aplicações**. Curitiba: CRV, 2020.

ESQUEDA, M. D. Machine translation: teaching and learning issues. **Trabalhos em Linguística Aplicada**, v. 60 (1), p. 282-299, 2021. DOI: <https://doi.org/10.1590/01031813932001520210212>.

FAPESP. **Sobre a instituição**. 2022. Disponível em: <https://fapesp.br/sobre/>. Acesso em: 20 fev. 2022.

FONSECA, N. B. de L. Investigando adjetivos atributivos em traduções publicadas, em um sistema de tradução automática e na pós-edição monolíngue: contribuições para a pedagogia da tradução. **Texto Livre**, Belo Horizonte-MG, v. 11, n. 1, p. 121-143, 2018. Disponível em: <https://periodicos.ufmg.br/index.php/textolivre/article/view/16784>. Acesso em: 9 set. 2022.

GILE, D. Analyzing Translation studies with scientometric data: from CIRIN to citation analysis. **Perspectives: Studies in Translatology**, v. 23, n. 2, p. 240–248, 2015.

GOOGLE ACADÊMICO. **Sobre o Google Acadêmico**. 2022. Disponível em: <http://scholar.google.com.br/intl/pt-BR/scholar/about.html>. Acesso em: 18 fev. 2022.

HARZING, Anne-Wil. **The publish or perish book**. Melbourne, Australia: Tarma Software Research Pty Limited, 2010. Disponível em: <https://www.harzing.info/download/popbook12.pdf>. Acesso em: 05 Mar. 2020.

HUTCHINS, W. J & SOMERS, H. L. **An introduction to Machine Translation**. Academic Press. San Diego, 1992.

HUTCHINS, W. J. Current commercial machine translation systems and computer-based translation tools: system types and their uses. **International journal of translation**, v.17, n. 1-2, p. 5-38, 2005.

HUTCHINS, W. J. **Machine Translation: a concise history**. 2007. Disponível em: <http://www.hutchinsweb.me.uk/CUHK-2006.pdf>. Acesso em: 10 de nov. 2020.

HUTCHINS, W. J. Machine translation: history of research and applications. *In*: SIN-WAI, C. (ed.). **The Routledge encyclopedia of translation technology**. London/New York: Routledge, p. 120-136, 2015.

MARTINS, R. T. **Tradução Automática e Estudos da Tradução: um conflito paradigmático**, 2003. Disponível em: http://www.nilc.icmc.usp.br/til/til2003/oral/RonaldoMartins_31.pdf. Acesso em: 9 set. 2022.

MARTINS, R. T.; NUNES, M. das G. V. **Noções gerais de tradução automática**. Relatório técnico. São Carlos: NILC-ICMC-USP, 25p., 2005. Disponível em: http://www.nilc.icmc.usp.br/nilc/download/NotasDidaticasICMC_68.pdf. Acesso em: 9 set. 2022.

MELO, S. S. C. de. Tradução automática e competência tradutória: repensando interseções. **Rónai – Revista de Estudos Clássicos e Tradutórios**, v. 1, n. 1, p. 87-104, 2013. Disponível em: <https://periodicos.ufjf.br/index.php/ronai/article/view/23060>. Acesso em: 9 set. 2022.

MENDELEY. **Sobre o Mendeley**. 2022. Disponível em: <https://www.elsevier.com/solutions/mendeley>. Acesso em: 12 fev. 2022.

MUGNAINI, R., STREHL, L. Recuperação e impacto da produção científica na era Google: uma análise comparativa entre o Google Acadêmico e a Web of Science. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação [en linea]**, p. 92-105, 2008. Disponível em: <https://www.redalyc.org/articulo.oa?id=14709808>. Acesso em: 02 Jan. 2022

NÓBREGA-TERRIEN, S. M., & TERRIEN, J. Trabalhos científicos e o estado da questão: reflexões teórico-metodológicas. **Estudos Em Avaliação Educacional**, v. 15, n.30, p.5-16, 2004. DOI: <https://doi.org/10.18222/eae153020042148>

NORUZI, A. Google Acadêmico: the new generation of citation indexes. **Libri**, v. 55, n. 4, p. 170-180, 2005.

O'HAGAN, M. **The Routledge Handbook of Translation and Technology**. New York: Routledge, 2020.

O'BRIEN, S., SPECIA, L., SIMARD, M., CARL, M., BALLING, L. W. **Post-editing of Machine Translation: Processes and Applications**. United Kingdom: Cambridge Scholars Publisher, 2014.

OLIVEIRA, M. F. de. **Enfrentamento do problema das divergências de tradução por um sistema de tradução automática: um exercício exploratório**. 2006.

ROLAND, J. Support patient access to care. Sandport/Idaho: **Multilingual**, p. 26-32, 2020.

SALES, S. G. Tradução automática: os processos da tradução mediada por computador. **Saberes em perspectiva**, v. 1, n. 1, p. 19-37, 2011. Disponível em: <https://philpapers.org/rec/GUSTAO>. Acesso em: 14 março 2022.

SANTOS, C. C. Um panorama do fluxo de recepção da tradução automática no cenário nacional. **In-Traduções**, Santa Catarina, Florianópolis, UFSC, vol. 05, n. 01, pp.167-176, 2011.

STUPIELLO, E. N. A. O texto adaptado à máquina: estratégias de controle autoral para implementação da tradução automática. **Estudos Linguísticos**, p. 696-706, 2010. Disponível em: <<http://hdl.handle.net/11449/122338>>. Acesso em: 9 set. 2022.

TAUS – matching data. 2022. Disponível em: <https://md.taus.net>. Acesso em: 15 março 2022.

TENOPIR, C.; KING, D. W. A importância dos periódicos para o trabalho científico. **Revista de Biblioteconomia de Brasília**, v. 25, n. 1, 2001. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/70824>. Acesso em: 21 fev. 2022.

TORAL, A., WAY, A. What Level of Quality Can Neural Machine Translation Attain on Literary Text? *In*: MOORKENS, J., CASTILHO, S., GASPARI, F., DOHERTY, S. (eds) **Translation Quality Assessment: From Principles to Practice**. Machine Translation Series. Switzerland: Springer International Publishing, 2018, p. 263- 287.

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

VANTI, N. A. P. Da bibliometria à webometria: uma exploração conceitual dos mecanismos utilizados para medir o registro da informação e a difusão do conhecimento. **Ci. Inf. [online]**, v. 31, n. 2, p. 369-379, 2002. DOI: <https://doi.org/10.1590/S0100-19652002000200016>.

VOSViewer. **Apresentação do VOSViewer**. 2022. Disponível em: <https://www.vosviewer.com>. Acesso em: 12 fev. 2022.

WEININGER, M. J. TM & MT na tradução técnica globalizada—tendências e consequências. **Cadernos de tradução**, Florianópolis, v. 2, n. 14, p. 243-263, 2004. Disponível em: <https://periodicos.ufsc.br/index.php/traducao/article/view/6483/5977>. Acesso em: 9 set. 2022.

APÊNDICE 2 – Relação dos documentos e os respectivos autores presentes no *corpus* geral

Base de Dados de onde os documentos foram retirados	Título da Publicação	Nome do Autor
PoP - Google Acadêmico	<i>Indução de léxicos bilíngües e regras para a tradução automática</i>	Helena de Medeiros Caseli
PoP - Google Acadêmico	<i>O anseio da língua-mãe nas entrelinhas da tradução automática</i>	Leila Cristina de Mello Darin
PoP - Google Acadêmico	<i>Noções Gerais de Tradução Automática</i>	Ronaldo T. Martins
PoP - Google Acadêmico	<i>Uma abordagem híbrida relacional para a desambiguação lexical de sentido na tradução automática</i>	Lucia Specia
PoP - Google Acadêmico	<i>Tradução Automática de Abstracts: avaliação do potencial e das limitações de três ferramentas da web</i>	José Paulo de Araújo
PoP - Google Acadêmico	<i>Análise comparativa dos resultados de mecanismos de tradução automática baseados em regras e estatística</i>	Fernando da Silva
PoP - Google Acadêmico	<i>Tradução Automática no Grupo Científico IBM-INESC</i>	Diana Santos
PoP - Google Acadêmico	<i>Tradução Automática: os processos da tradução mediada por computador</i>	Sílvia Gusmão Sales
PoP - Google Acadêmico	<i>Enfrentamento do problema das divergências de tradução por um sistema de tradução automática: um exercício exploratório.</i>	Mirna Fernanda de Oliveira
PoP - Google Acadêmico	<i>Aplicação do Word2vec e do Gradiente descendente estocástico em tradução automática</i>	Eliane Martins de Aguiar.
PoP - Google Acadêmico	<i>Construção de um Córpus Paralelo Alinhado para a Tradução Automática Estatística</i>	Wilker Ferreira Aziz
PoP - Google Acadêmico	<i>Estudo e aprimoramento do sistema Apertium de tradução automática entre português e espanhol</i>	Felipe Tassario Gomes
PoP - Google Acadêmico	<i>Um Modelo para a Desambiguação Lexical de Sentido na Tradução Automática</i>	Lucia Specia
PoP - Google Acadêmico	<i>Por uma linguagem controlada na tradução automática de resumos acadêmicos</i>	Cleydstone Chaves dos Santos

PoP - Google Acadêmico	<i>O texto adaptado à máquina: estratégias de controle autoral para implementação da tradução automática</i>	Érika Nogueira de Andrade Stupiello
PoP - Google Acadêmico	<i>Tradução automática estatística baseada em sintaxe e linguagens de árvores</i>	Daniel Emílio Beck
PoP - Google Acadêmico	<i>Experimentos Comparativos Combinando Aprendizado Supervisionado e Tradução Automática para Mineração de Emoções em Textos Multilíngues</i>	Aline Graciela Lermen dos Santos
PoP - Google Acadêmico	<i>A influência da Tradução Automática (TA) na produtividade do tradutor: um experimento prático com uma ferramenta gratuita online.</i>	Luciana Alves Graziuso
PoP - Google Acadêmico	<i>Ambiguidades linguísticas no inglês e a tradução automática inglês-português: um estudo exploratório</i>	Maria Paula Fiorim Piruzelli
PoP - Google Acadêmico	<i>Tradução Automática: estratégias e limitações</i>	Helena de Medeiros Caseli
PoP - Google Acadêmico	<i>Treinamento linguístico de software na pós-edição de transcrição e tradução automática em cursos de educação a distância</i>	Luana Moro
PoP - Google Acadêmico	<i>O jogo da avaliação : um estudo prático sobre tradução automática</i>	Michel de Sousa Teixeira
PoP - Google Acadêmico	<i>Tradução automática de textos científicos como suporte pedagógico para o desenvolvimento da compreensão leitora em inglês para propósitos acadêmicos</i>	Débora Borsatti
PoP - Google Acadêmico	<i>Investigando adjetivos atributivos em traduções publicadas, em um sistema de tradução automática e na pós-edição monolíngue: contribuições para a pedagogia da tradução</i>	Norma Barbosa de Lima Fonseca
PoP - Google Acadêmico	<i>Tradução automática adaptativa, gerenciamento de terminologia e pós-edição: benefícios e desafios do uso de tecnologia na tradução de finanças</i>	Carolina Bisson de Souza
PoP - Google Acadêmico	<i>Estratégias de Tradução Automática: comparando a estrutura retórica do texto fonte e a estrutura retórica do texto gerado por um tradutor automático.</i>	Juliano Desiderato Antonio
PoP - Google Acadêmico	<i>Incompatibilidades intersemióticas geradas por tradução automática: um estudo exploratório sobre coleta, etiquetagem e anotação em corpus de documentos multimodais estáticos</i>	Augusto Velloso dos Santos Espindola
PoP - Google Acadêmico	<i>Sistemas de memórias de tradução e tecnologias de tradução automática: possíveis efeitos na produção de tradutores em formação</i>	Lara Cristina Santos Talhaferro
PoP - Google Acadêmico	<i>Aprimoramento do sistema Apertium de tradução automática</i>	Felipe Tassario Gomes

PoP - Google Acadêmico	<i>A tradução automática e os seus limites</i>	Cristiane de Souza
PoP - Google Acadêmico	<i>Análise do presente perfeito em contextos de tradução automática</i>	Roberlei Alves Bertucci
PoP - Google Acadêmico	<i>Tecnologias unem linguística e computação para aprimorar tradução automática</i>	Aline Naoe
PoP - Google Acadêmico	<i>Investigação de Modelos Estatísticos de Tradução Automática Baseados em Phrases</i>	Wilker Ferreira Aziz
PoP - Google Acadêmico	<i>Os corpora eletrônicos nos estudos da tradução automática</i>	Cleydstone Chaves dos Santos
PoP - Google Acadêmico	<i>Multimodalidade e avaliação de tradução automática: uma proposta para a investigação de incompatibilidades intersemióticas geradas pelo uso do tradutor automático em documentos multimodais</i>	Thiago Blanch Pires
PoP - Google Acadêmico	<i>Tradução Automática entre línguas com poucos recursos computacionais</i>	Prof. Dr. Marcello Modesto dos Santos
PoP - Google Acadêmico	<i>Desmistificando a tradução automática – um estudo sobre seu uso e sua recepção</i>	Silvia Malena Modesto Monteiro
PoP - Google Acadêmico	<i>Tradução Automática e Estudos da Tradução: um conflito paradigmático</i>	Ronaldo T. Martins
PoP - Google Acadêmico	<i>Ampliando olhares sobre a tradução automática online: um estudo exploratório de categorias de erros de máquina de tradução gerados em documentos multimodais.</i>	Thiago Blanch Pires
PoP - Google Acadêmico	<i>A controvérsia da equivalência na tradução automática de grupos nominais</i>	Juliana Ramos do Nascimento
PoP - Google Acadêmico	<i>A tradução automática de colocação lexical em resumo acadêmico</i>	Janailton Mick Vitor da Silva
PoP - Google Acadêmico	<i>A nova torre de babel – que futuro para a tradução automática?</i>	Sara Cerqueira
PoP - Google Acadêmico	<i>Tradução Automática no contexto militar: um estudo sobre sua aplicação.</i>	Cristiano Gonçalves Araújo
PoP - Google Acadêmico	<i>Alinhamento Sentencial e Lexical de Córpus Paralelos: Recursos para a Tradução Automática</i>	Helena de Medeiros Caseli

PoP - Google Acadêmico	<i>Avaliação de tradução automática no mercado de localização de software: um estudo de caso</i>	Gabriela Castelo Branco Ribeiro
PoP - Google Acadêmico	<i>Um Estudo Comparativo de Três Técnicas de Tradução Automática para Processamento de Textos</i>	Maria Alice O. C. Leal
PoP - Google Acadêmico	<i>Alinhamento de Grafos: Investigação do Alinhamento de ConceptNets para a Tradução Automática</i>	Paulo Henrique Barchi
PoP - Google Acadêmico	<i>Avaliando Tradução Automática e Simplificação Textual com Redes Complexas</i>	Diego R. Amancio
PoP - Google Acadêmico	<i>Anotação manual de erros de tradução automática em textos traduzidos de inglês para português do Brasil</i>	Débora Beatriz de Jesus Martins
PoP - Google Acadêmico	<i>As estruturas léxico-conceituais e a elaboração de bases lexicais para tradução automática: o caso do verbo apunhalar</i>	Bento Carlos Dias-da-Silva
PoP - Google Acadêmico	<i>O tratamento dos papéis temáticos por um sistema de tradução automática do tipo interlíngua: parâmetros para a língua portuguesa.</i>	Mirna Fernanda de Oliveira
PoP - Google Acadêmico	<i>Análise dos softwares gistsummarizer e google tradutor: ferramentas de suporte à sumarização de textos científicos e à tradução automática</i>	Liliane Simões dos Santos
PoP - Google Acadêmico	<i>Tradução Automática Estatística baseada em Frases e Fatorada: Experimentos com os idiomas Português do Brasil e Inglês usando o toolkit Moses</i>	Helena de Medeiros Caseli
PoP - Google Acadêmico	<i>A criação de um sistema híbrido de tradução automática para a conversão de expressões nominais da língua inglesa para a língua portuguesa</i>	Tiago Martins da Cunha
PoP - Google Acadêmico	<i>O tradutor e a tradução automática: uma análise da utilidade da ferramenta google tradutor para a versão de resumos de tccs da área de direito</i>	Rafael Reis Costa
PoP - Google Acadêmico	<i>PorTAL: Recursos e Ferramentas de Tradução Automática para o Português do Brasil</i>	Thiago Lima Vieira
PoP - Google Acadêmico	<i>Trapezio – Translation Post Editor: um ambiente de pós-edição de traduções automáticas</i>	Felipe Tassario Gomes
PoP - Google Acadêmico	<i>Avaliação de qualidade de traduções automáticas usando redes complexas</i>	Diego R. Amancio
PoP - Google Acadêmico	<i>Investigação e Desenvolvimento de uma Ferramenta de Pós-Edição de Traduções Automáticas</i>	Felipe Tassario Gomes

PoP - Google Acadêmico	<i>Aperfeiçoamento de um tradutor automático Português-Inglês: Tempos Verbais</i>	Lucia Helena Rozario da Silva
PoP - Google Acadêmico	<i>O enfoque da sequência didática nas reflexões sobre língua e cultura no processo tradutório: o uso do tradutor automático como atividade diagnóstica</i>	Maria José Damiani Costa
PoP - Google Acadêmico	<i>Desenvolvimento de um tradutor automático Português Brasileiro-Inglês</i>	Lucia Helena Rozario da Silva
PoP - Google Acadêmico	<i>O (re)conhecimento da tradução em sala de aula: sobre uma experiência prática com tradutor automático online</i>	Maria José Damiani Costa
PoP - Google Acadêmico	<i>Análise Comparativa de Tradutores Automáticos Inglês-Português</i>	Marcela Franco Fossey
PoP - Google Acadêmico	<i>Aplicação da métrica bleu para avaliação comparativa dos tradutores automáticos bing tradutor e google tradutor</i>	Francisco Ramos de Melo
PoP - Google Acadêmico	<i>Uma Alternativa Colaborativa para Expansão de Dicionários Morfológicos de Tradutores Automáticos Baseados em Regras</i>	Vinícius Silva Nogueira
PoP - Google Acadêmico	<i>Requisitos para a modelagem de padrões de cunhagem e construções semi-produtivas no constructicon da framenet brasil com foco no fomento ao desenvolvimento de tradutores automáticos</i>	Tatiane Silva Tavares
PoP - Google Acadêmico	<i>Compreensão leitora e o uso do tradutor eletrônico nas aulas de inglês instrumental (esp)</i>	Vera Wannmacher Pereira
PoP - Google Acadêmico	<i>Tradução eletrônica: do riso irônico ao interesse científico</i>	Luciano Rodrigues Lima
PoP - Google Acadêmico	<i>O gênero notícia on line e a tradução eletrônica nas aulas de língua estrangeira moderna inglês na perspectiva da pedagogia crítica da educação</i>	Jorge Galarce
Belas (UNB)	Infieis <i>Uma Breve Reflexão Historiográfica Sobre A Tradução Automática</i>	Marcos de Campos Carneiro
Belas (UNB)	Infieis <i>Avaliação Humana Da Tradução Automática De Combinações Lexicais Especializadas:O Caso Do Google Translate E Do Deepl</i>	Thais Miranda e Borges
Cadernos de Tradução (UFRGS)	de <i>Um Estudo Comparativo Da Desambiguação Lexical Realizada Por Tradutor Eletrônico E Tradutores Humanos</i>	Adriana Riess Karnal
Cadernos de Tradução (UFSC)	de <i>TM & MT na tradução técnica globalizada – tendências e conseqüências.</i>	Markus J. Weininger

Cadernos de Tradução (UFSC)	de	<i>Machine Translation: A Critical Look At The Performance Of Rule-Based And Statistical Machine Translation</i>	Brita Banitz
Cadernos de Tradução (UFSC)	de	<i>El general de brigada es um tipo de caramelo – tradução automática e aprendizagem cultural.</i>	Nylcea Thereza de Siqueira Pedra
Cadernos de Tradução (UFSC)	de	<i>Tradução Automática: Uma Ferramenta de Auxílio ao Tradutor</i>	Carolina Alfaro
Cadernos de Tradução (UFSC)	de	<i>Um estudo de expressões cristalizadas do tipo V+SN e sua inclusão em um tradutor automático bilíngüe (português/inglês)</i>	Milena Uzeda Garrão
Cadernos de Tradução (UFSC)	de	<i>Déjà vu: tradução assistida ao pormenos</i>	Leonel Morgado
Cultura e Tradução (UFPB)	e	<i>Por uma estética na tradução automática: um estudo baseado em corpora eletrônicos</i>	Cleydstone Chaves dos Santos
Cultura e Tradução (UFPB)	e	<i>Coleta, etiquetagem e anotação de incompatibilidades intersemióticas geradas por tradução automática</i>	Augusto Velloso dos Santos Espindola
In-Traduções Revista do Programa de Pós-Graduação em Estudos da Tradução da UFSC		<i>Um panorama do fluxo de recepção da tradução automática no cenário nacional</i>	Cleydstone Chaves dos Santos
Rónai (UFJF)		<i>As imperfeições da tradução automática on-line</i>	Priscilla Pellegrino de Oliveira
Rónai (UFJF)		<i>Tradução automática e competência tradutória: repensando interseções</i>	Sheila de Souza de Melo
TRADTERM (USP)		<i>Google Tradutor: Análise de Utilização e Desempenho da Ferramenta</i>	Gislaine Caprioli Costa
TRADTERM (USP)		<i>Terminografia para a tradução automática: alimentar metal com termos</i>	Caroline de Schaetzen
Tradução e Comunicação: Revista Brasileira de Tradutores (UNIBERO/AN HANGUERA)	e	<i>Um tradutor automático de expressões nominais com idéia de posse</i>	Tiago Martins da Cunha

Fonte: a autora.

