

Os anos da mecanização (1940-1980): uma história da Tradução Automática

The years of mechanization (1940-1980): a history of Machine Translation

Ronaldo de Oliveira Batista

Universidade Presbiteriana Mackenzie (Brasil)

Conselho Nacional de Desenvolvimento Científico e Tecnológico (Brasil)

ronaldo.obatista@gmail.com

Luciana Debonis

Universidade Presbiteriana Mackenzie (Brasil)

luddeb01@yahoo.com

Recebido: 04/06/2024

Aceite: 19/10/2024

Publicado: 23/12/2024

RESUMO: Narra-se uma história da Tradução Automática (TA) dos anos 1940-1980 a partir da Historiografia Linguística e do conceito de programas de investigação. Para tanto, duas perguntas são estabelecidas em função do período histórico selecionado: (1) Quais as características das ferramentas tecnológicas nesses anos?; (2) Quais impactos estão presentes nesses anos em relação à atuação do tradutor humano? O percurso metodológico adotado para apresentação de resultados neste artigo estabeleceu-se em três fases: 1ª. fase heurística: definição da periodização; seleção/tratamento do *corpus*; 2ª. fase hermenêutica: estabelecimento de diretrizes para a análise; 3ª. fase expositiva: apresentação de resultados em uma narrativa historiográfica. O material de análise foi organizado e hierarquizado em dois grandes conjuntos: documentos sobre funcionalidades e ferramentas; documentos sobre a história da TA. Argumenta-se, neste artigo, que em cerca de 40 anos a TA se caracterizou por iniciativas e empreendimentos de um programa de mecanização que começou a alterar o perfil do tradutor e sua inserção profissional.

PALAVRAS-CHAVE: Historiografia Linguística; Tradução Automática; Programas de Investigação; Mecanização.

ABSTRACT: We sketched the history of machine translation (MT) from the 1940s to the 1980s from a linguistic historiography perspective and the concept of research programs. For this purpose, two questions were established according to the selected historical period: (1) What are the characteristics and the use of technological tools? (2) What are the impacts such changes have on social contexts of use in relation to the performance of human translators? The methodological direction adopted to demonstrate the results in this article was established into three phases: 1st. heuristic phase: periodization definition; selection/treatment of corpus; 2nd. hermeneutic phase: guidelines establishment for the

analysis; 3rd. expository phase: results presentation through a historiographical narrative. The analysis material was organized and hierarchized into two large sets: documents regarding functionalities and tools and documents in connection with the history of MT. This article argues that for about 40 years, the MT was characterized by initiatives and operations of a mechanization program that began to change the translator's profile and professional insertion.

KEYWORDS: Linguistic Historiography; Machine Translation; Research Programs; Mechanization.

Introdução

Este artigo revê a história da Tradução Automática (TA) - “modalidade de tradução que consiste na automatização de uma parte ou do todo no processo de tradução de uma língua natural para outra” (Debonis 2021: 32) - entre 1940-1980 a partir da Historiografia Linguística, campo que estuda o desenvolvimento histórico do conhecimento sobre a linguagem, no qual estão também aqueles sobre tradução (cf. Altman, 2019; Batista, 2013, 2020; Koerner, 2014; Swiggers, 2019).

Duas perguntas nos guiam: (1) Quais as características das ferramentas tecnológicas nesses anos?; (2) Quais impactos estão presentes nesses anos em relação à atuação do tradutor humano? A argumentação principal é a de que em 40 anos a TA se caracterizou por empreendimentos de um “programa de mecanização”¹ que alterou o perfil do tradutor e sua inserção profissional.

As décadas de 1940-1980 na TA foram revistas por autores como Hutchins (2001), Poibeau (2017), Pym (1998). No entanto, o ponto de vista aqui adotado é distinto, pois insere a análise em um campo dos estudos linguísticos por meio de caracterização interna (características de ferramentas, métodos e visões teóricas subjacentes a estas) e externa (contextos científico-tecnológico e empresarial de produção, circulação e recepção de ferramentas e métodos) de práticas de tradução em torno do conceito de *programas de investigação*.

O percurso metodológico se deu em três fases e orientou a elaboração deste artigo: 1ª. fase, *heurística*: definição da periodização (1940-1980, os primeiros anos no século XX da história da TA - cf. Arnold *et al.*, 1994; Hutchins, 2001; Poibeau, 2017; Pym, 1998); seleção/tratamento do *corpus* (correspondências, relatórios,

¹ Este artigo retoma, reorganiza e sintetiza argumentações apresentadas em Debonis (2021).

experimentos, revisões históricas - Alpac, 1966; Bar-Hillel, 1951; Dostert, 1957; Holmes, 1972 [2000]; Hutchins, 2001; Masterman & Kay, 1960; Nelson, 1967; Oswald, 1952, 1953; Poibeau, 2017; Pym, 1998; Weaver, 1947-1949; Yngve, 1953); 2ª. fase, *hermenêutica*: com base no conceito teórico de *programas de investigação*, estabelecimento de perguntas (problemas a investigar, já citados) para apreensão de conteúdos presentes, pressupostos ou subentendidos no *corpus* com relação a seus diversos contextos de produção, circulação e recepção; 3ª. fase, *executiva*: apresentação desta narrativa histórica, caracterizada pelo formato expositivo em torno de causalidades e implicações explicativas de fatos diante de questionamentos incidentes sobre instâncias de produção, circulação e recepção de ideias e conteúdos examinados na fase hermenêutica.

Como apontado, a história da TA é observada pelo conceito de *programas de investigação*: modos de tratamento da linguagem definidos por procedimentos adotados e resultados alcançados, derivados de concepções teóricas. O conceito é de Swiggers (1981). Seus programas de investigação (descritivista, de correspondência, sociocultural, de projeção), originalmente definidos para teorias linguísticas, precisam constantemente passar por revisões diante de novas tendências e confluências na linguística e nas práticas a ela relacionadas (como a tradução).

Este artigo, ao rever Swiggers (1981), propõe a existência na história da TA de um “programa de mecanização”. Insere, assim, em outras narrativas históricas sobre métodos de tradução (ampliando o escopo destas), uma perspectiva que articula características técnicas e científicas de um modelo tradutório a elementos sociais que sustentaram práticas profissionais de tradutores.

Para tanto, apresenta uma narrativa histórica em três seções principais (além de uma Introdução na qual o problema central é colocado e de uma Conclusão que retoma os questionamentos iniciais e os responde): a) a primeira, que trata de um clima intelectual e científico nos anos 1940-1970; b) a segunda, que propõe a caracterização de um programa de mecanização na história da TA; c) a terceira, que observa aspectos sociais do impacto desse período de mecanização na história da TA.

1. Como cenário para a TA, um clima intelectual formalista

Estudos linguísticos dos anos 1940-1970 foram em grande parte estruturalistas de feição formalista. Um cenário teórico-metodológico que seria alterado durante os anos 1970 com a presença de teorias interdisciplinares (como a sociolinguística e a psicolinguística), teorias do texto e do discurso e teorias funcionalistas.

Neste artigo, *estruturalismo* é denominação abrangente e se refere a propostas teórico-metodológicas em período que vai da década de 1920 até finais da década de 1970. Dentre estas: o descritivismo nos Estados Unidos (Leonard Bloomfield [1887-1949] e Kenneth Pike [1912-2000]); estruturalismos derivados do Círculo Linguístico de Praga (1920-1930); a glossemática de Louis Hjelmslev (1899-1965); a semântica estrutural de Algirdas Julien Greimas (1917-1992); o gerativismo de Noam Chomsky nas teorias padrão e padrão estendido. A definição *estruturalista* refere-se antes de tudo a um ponto de vista, mais do que a uma corrente teórica única: é uma concepção geral de língua como conjunto organizado e regado de unidades em permanente relação. Já *formalista/formalismo* denotam abordagem imanente de aspectos linguísticos. Para os formalistas, atribuem-se: preexistência, autonomia e primazia analítica da forma; forma anterior ao uso; representações visuais correlacionadas a regras estruturais ou aspectos mentais; explicações internas ao sistema (sem apelo ao social ou à subjetividade dos falantes).

Algumas dessas correntes formalistas influenciaram, por hipótese, procedimentos de tradução, mesmo que não reconhecidas explicitamente fora do universo acadêmico em humanidades. Um paradigma formalista estruturalista que não foi homogêneo em teorias e métodos, pois diferentes propostas foram divulgadas e defendidas. Algumas tiveram seus momentos de destaque e outras de ostracismo. Diante desse pressuposto, diferentes modos de conceber tarefas do linguista eram possíveis.

Nesses anos 1940-1970, teorias formalistas estruturalistas analisavam os sistemas linguísticos como fim em si mesmos. E a forma linguística era definida por relações estruturais e regulares entre as unidades de um sistema. Nessa vertente, teríamos os “linguistas-engenheiros” (Thomas, 2020), que dariam destaque para:

representações formais correlacionadas a regras; explicações internas ao sistema (sem nenhum apelo ao social ou à subjetividade dos falantes); generalizações; observação (com base em método dedutivo) de fatos gramaticais específicos.

Seria essa linguística (em toda sua diversidade) que comporia cenário de fundo para propostas de tradução (especialmente as de natureza técnica) com uso de ferramentas e programas computacionais que mecanicamente fariam transições interlinguísticas.

2. Um programa de mecanização na história da TA

Neste artigo, argumenta-se que os primeiros anos da história da TA foram anos de um “programa de mecanização”, com pesquisas em protótipos de sistemas computacionais motivadas por diferentes conflitos entre povos e nações, diante de demandas de vigilância e compreensão rápida de línguas de outros povos e países.

E por que a caracterização como programa de mecanização? Seguindo diretrizes classificatórias de Swiggers (1981) para outro objeto analítico e contexto científico, destacam-se como elementos para existência desse programa na história da TA:

QUADRO 1 - Caracterização do programa de mecanização na TA

CONCEPÇÃO TEÓRICA: visão de língua como sistema codificado (criptografia) e de gramática como estrutura de relações formais regulares entre unidades.
MÉTODO: técnica de tradução no nível da palavra (incidência no léxico) e modelo de tradução com base em dicionário.

(Fonte: elaborado pelos autores)

Situar mudanças no percurso histórico da TA é também propor paralelo com marcos nos estudos linguísticos em termos de alterações em paradigmas teóricos e metodológicos que delimitaram objetivos e alcances de práticas de pesquisa em torno da descrição e análise linguísticas.

O programa de mecanização está relacionado, até pelo recorte temporal e pelo espaço geográfico no qual inicialmente estava circunscrito (os EUA), ao descritivismo praticado por linguistas formalistas norte-americanos (cf. Altman,

2021) como Leonard Bloomfield e seus seguidores (os pós-bloomfieldianos ou distribucionalistas).

O que permite essa vinculação é em essência o teor mecanicista da linguística de extração bloomfieldiana. Ao privilegiar descrições sincrônicas dos níveis fonológico e morfológico, por meio da seleção de unidades e da explicação dos seus modos de combinação e atuação no sistema, o descritivismo norte-americano (como é reconhecido o paradigma com origem em Bloomfield e Kenneth Pike) associou-se a uma concepção behaviorista de linguagem. Daí, portanto, a aproximação entre o programa de mecanização aqui exposto e o descritivismo de base behaviorista.

O programa de mecanização agrupa pesquisas com objetivo de mecanizar a gramática, ou seja, esforços voltavam-se para criação de sistemas em que a máquina aplicasse regras gramaticais das línguas naturais. Buscava-se programar a máquina para realizar tarefas de tradução como se fosse uma operação matemática. Nesta apenas a aplicação da regra certa poderia transferir um texto de partida para uma língua de chegada (LC), sem considerar o que aconteceria no caminho da transferência. Primeiras ferramentas usavam o modelo de “tradução palavra por palavra” (WMT), no qual eram usados dicionários e glossários bilíngues para transferência linguística.

Pesquisadores tentaram simplificar dicionários bilíngues, fornecendo equivalentes únicos para palavras da língua de partida (LP) que abrangessem uma ampla maioria de sentidos. Assim, os experimentos buscavam, na sua maioria, desenvolver regras de programação para traduzir de forma direta, de uma LP para uma LC específica apenas, com mínima análise e reorganização sintática, evitando análise de contextos e mantendo a ordem das palavras da LP.

Alguns grupos de pesquisa investigaram, ainda, possibilidades de representações “interlínguas”, ou seja, códigos ou símbolos independentes tanto na LP como na LC, em que a tradução ocorreria em duas etapas: 1) a partir da LP para a “interlíngua” e 2) a partir da “interlíngua” para a LC. Porém, essa era uma visão de pesquisa mais a longo prazo e a maioria se dedicou a pesquisas linguísticas básicas, isso porque alguns pesquisadores já acreditavam que apenas

investigações mais específicas sobre o processo de pensamento humano (posteriormente chamado de “inteligência artificial” ou “ciência cognitiva”) poderia melhorar a qualidade da TA (Hutchins, 2001: 11).

A base teórica era o conceito da tradução direta. Primeiros resultados não foram satisfatórios, com baixa qualidade na tradução. Apesar da qualidade ruim no texto da LC, pesquisas e usos da ferramenta continuaram. Isso atendia à agilidade e à redução de custos em comparação com o tradutor humano.

Esse processo de uma tradução direta, ancorada na possibilidade de criação de máquinas, tem em essência um vínculo, mesmo que muitas vezes não intencional, com a visão mecanicista de língua que caracterizou parte do estruturalismo linguístico.

A construção de máquinas para TA com base puramente mecânica estendeu-se até a década de 1960, mas recursos computacionais disponíveis representavam barreira para evolução dos modelos tecnológicos. Nesse período, a máquina só executava a tradução em uma direção, por exemplo, do russo para o inglês, e não conseguia realizar o caminho contrário, para isso era preciso outra máquina.

Primeiras ferramentas e protótipos de TA tiveram menor envolvimento de linguistas e maior concentração de especialistas da área de exatas e da computação (que estava começando). O principal envolvimento de linguistas se dava na interpretação gramatical, para estabelecer regras gramaticais a serem aplicadas na máquina mecanicamente, assim como, na pós-edição do texto traduzido pela máquina, tarefa essa desempenhada por linguistas e tradutores profissionais.

Os grupos de pesquisa atuantes nesse programa estavam em grande parte nos EUA e se formaram, principalmente, dentro de universidades com investimento público e privado. Esses grupos contribuíram para a produção acadêmica nas décadas de 1950-1960, impulsionando avanços computacionais, motivados pela demanda da TA e que serviram, posteriormente, como fundamento para o desenvolvimento de linguagens de programação.

QUADRO 2 - Grupos de pesquisa norte-americanos em 1950-1960

Grupo de Erwin Reiffler, Universidade de Washington (Seattle, EUA): otimizou a abordagem da tradução direta com base em dicionário que envolveu a construção de grandes dicionários bilíngues em que a informação lexical era usada também para resolver problemas gramaticais, sem o uso de análise sintática.
Grupo de Leon Dostert, Universidade de Georgetown (Washington, EUA): conduziu análises de textos apenas quando as informações da gramática tradicional eram inadequadas. Desenvolveu o sistema <i>Georgetown Automatic Translation</i> (GAT), que possuía três níveis de análise: morfológica, sintagmática e sintática. O sistema foi instalado na Itália em 1963 pela Euratom e posteriormente adotado pela Comissão de Energia Atômica Americana em 1964.
Grupo de Anthony Oettinger, Universidade de Harvard (Massachusetts, EUA): concentrou as pesquisas na compilação de um dicionário russo-inglês e em 1959 avançou no desenvolvimento de sistema para identificação de sequências de categorias gramaticais e sua predição probabilística.
Grupo de Victor Yngve, MIT (Massachusetts, EUA): em 1953 concentrava estudos na relação sintática. A LP era analisada como representações de estruturas de frase, em que uma rotina de transferência de estrutura convertia em estruturas de frases equivalentes na LC. O principal problema encontrado era a barreira semântica do texto traduzido. Chomsky fez parte desse grupo por pouco tempo, mas não se encontram evidências de abordagem chomskiana nas pesquisas em TA nesta época.
Grupo de Winfried Lehmann, Universidade do Texas (Austin, EUA): grupo atuando no <i>Linguistic Research Center</i> em 1958, que concentrava estudos na relação sintática. Tentou construir sistema que pudesse fazer tradução bidirecional com base na abordagem da transferência sintática, deixando fundamentos para o posterior desenvolvimento do sistema METAL.
Grupo de Sydney Lamb, Universidade da Califórnia (Berkeley, USA): concentrou-se no desenvolvimento de rotinas de dicionários eficientes e de uma teoria linguística adequada para TA. Este conceito seguia a ideia de relações paralelas com a arquitetura de computadores (nós e redes), em que a tradução era vista como processo de codificação e decodificação.

(Fonte: elaborado pelos autores)

QUADRO 3 - Grupos de pesquisa fora dos EUA em 1950-1960

União Soviética (atual Rússia): concentravam-se em sistemas inglês-russo e na análise morfológica e sintática para a transformação da LP por meio de regras gramaticais. Apesar de o inglês ser o principal idioma estudado devido questões políticas entre os países, as pesquisas ampliaram para uma diversidade maior de idiomas do que nos EUA. Os principais grupos de pesquisa foram: o grupo de D. Y. Panov no <i>Institute of Precision Mechanic</i> (hoje <i>Lebedev Institute of Precision Mechanics and Computer Engineering</i>), o grupo de Ljapunov, Kulagina e Mel'cuk no <i>Steklov Mathematical Institute</i> e o grupo de Nikolaj Andreev na Universidade Estadual de Leningrado.
Londres, Reino Unido: o grupo de Margaret Masterman na <i>Language Research Unit</i> de Cambridge adotou a linha de abordagem "interlíngua" que consiste no método de transformação do texto fonte numa representação abstrata de língua independente (a interlíngua) e, em seguida, o texto de chegada é gerado a partir da interlíngua. No caso das pesquisas de Cambridge, o desenvolvimento do método interlíngua foi baseado em palavra por palavra produzindo traduções denominadas "pidgin".
Milão, Itália: o grupo de Silvio Ceccato na Universidade de Milão também concentrou suas pesquisas no método da interlíngua e o desenvolveu com base em processos cognitivos, especialmente no conceito de análise de palavras e suas possíveis correlações com outras palavras no texto.

(Fonte: elaborado pelos autores)

Outros grupos fora dos EUA avançaram com protótipos de sistemas de TA nas décadas seguintes. Como é o caso do grupo TAUM liderado por John Chandioux na Universidade de Montreal, no Canadá, responsável pelo desenvolvimento do sistema Meteo adotado em 1977 pelo Centro de Meteorologia Canadense. Além disso, grupos de pesquisa no território europeu e no Japão deram seguimento ao desenvolvimento de ferramentas de TA. Dentre estes: o grupo que desenvolveu o sistema GETA-Ariane na França; o MU no Japão; o SUSY na Alemanha; o AMPAR (inglês-russo) e NERPA (alemão-russo) na ex-União Soviética.

Em 1970, na Universidade de Montreal, a pesquisa do grupo TAUM sobre o sistema Meteo (para tradução inglês-francês de previsões meteorológicas) aplicava vocabulário e sintaxe específicos dos relatórios meteorológicos. O sistema Meteo, baseado em tradução direta, empregava dois dicionários bilíngues para processamento da TA. Sendo, a) um de expressões meteorológicas, incluindo por volta de 300 verbetes e b) outro com categorizações lexicais, ou seja, para cada par inglês-francês teria sua correspondência semântica e variações morfológicas; este incluía por volta de 1.200 verbetes. O processamento da TA no Meteo era feito por meio da aplicação de um *parser* (analisador) capaz de analisar uma sequência de frases na LP e buscar sua correspondência na LC, de acordo com as regras especificadas pelos dicionários aplicados.

Processamento de TA a partir de tradução direta, característico do programa de mecanização, pode ser observado também no desenvolvimento de sistemas como Metal (1970) - parceria entre a Universidade do Texas (EUA) e a empresa Siemens (Alemanha) - para tradução do inglês-alemão, e Sysram (1976) - desenvolvido pela empresa norte-americana Sysram - para tradução de diferentes pares de idiomas, que consistia na tradução das sentenças analisadas na LP em suas "formas normais", ou seja, sem elementos lexicais interlínguas. Porém, problemas nos níveis de análise detectados eram difíceis de serem corrigidos.

A importância desses grupos é fundamental. Foram as elites universitárias responsáveis pela difusão de uma teoria e legitimação no ambiente acadêmico. A

rede de influências foi consolidada por esses grupos, que promoveram circulação da produção acadêmica e aplicação profissional de suas propostas.

A divulgação do trabalho desses grupos avançou, e em 1954 a primeira revista sobre TA intitulada "*Mechanical Translation*" foi lançada por William Locke e Victor Yngve, que sucederam Bar-Hillel no Instituto de Tecnologia de Massachusetts (MIT) a partir de 1953. Em 1954, foi concluída a primeira tese de doutorado em TA intitulada "*A study for the design of an automatic dictionary*", de Anthony G. Oettinger (Harvard), que estudou um dicionário mecânico em russo que pudesse ser atualizado automaticamente. No ano seguinte, em 1955, o primeiro livro sobre TA intitulado "*Machine Translation of Languages: fourteen essays*" foi lançado e era uma coleção editada por William Locke e Andrew Booth, incluindo artigos acadêmicos publicados por Yehoshua Bar-Hillel (MIT), Leon Dostert (Georgetown), Anthony Oettinger (Harvard), Erwin Reifler (Universidade de Washington) e Victor Yngve (MIT), além do memorando de Warren Weaver de 1947.

Apesar de várias iniciativas para a divulgação da TA, as pesquisas publicadas ainda eram poucas. Não encontramos trabalhos na área das linguagens nessa época que discutam TA. A produção era fundamentalmente sobre estudos técnicos, sem interesse para as humanidades na discussão sobre impactos sociais da revolução tecnológica que iniciativas como as de TA provocariam.

A publicação de um relatório em 1966 criticando investimentos em TA provocou desmonte dos grupos norte-americanos. Redigido pela ALPAC, resultou no redirecionamento do apoio financeiro para produção científica voltada a aparatos bélicos e às ciências espaciais (os EUA passavam pela Guerra do Vietnã [1955-1975] e começavam sua Corrida Espacial [1960-1975]).

Os efeitos do relatório também foram percebidos em grupos fora do território norte-americano, mas estes conseguiram atingir resultados importantes, o que motivou uma retomada gradativa de grupos de pesquisa nos EUA a partir da década de 1980.

A década de 1980 também foi marcada pelo estouro nas pesquisas em tecnologia no Japão e, conseqüentemente, aumentou o interesse desse país no desenvolvimento de sistemas de TA. Essas pesquisas também influenciaram o

interesse pela TA de outros países daquele continente como Coréia, Taiwan, China e Malásia.

Nesse ponto, os grupos começaram a se interessar pelo processo do pensamento humano (o que futuramente seriam os estudos em IA) no PLN e acreditavam que naquele caminho encontrariam as soluções para os problemas da TA. Esses grupos impulsionaram a consolidação da TA como uma área de pesquisa importante. No entanto, ainda se tratava de um campo de pesquisa complexo, tanto do ponto de vista técnico como de gestão organizacional e de infraestrutura, e caro em relação ao esforço de formação de equipes e tempo para obter resultados.

Influenciado por projetos globais, o grupo de pesquisa GETA, em Grenoble na França, começou a desenvolver o sistema Ariane nos anos 1980. O sistema destacou-se pela flexibilidade, modularidade e algoritmos avançados para manipulação de estruturas de árvores. O grupo introduziu conceitos de gramáticas estáticas e dinâmicas, permitindo diferentes níveis de representação em estruturas de árvore rotuladas, o que proporcionou flexibilidade em transferências multiníveis.

O mercado de TA teve uma expansão e vários sistemas entraram em funcionamento em instituições públicas e grandes corporações privadas. Multinacionais investiram no desenvolvimento de seus próprios sistemas de TA e, posteriormente, empresas de *softwares* entraram no mercado oferecendo opções personalizadas também. Alguns desses primeiros sistemas serviram de base para o desenvolvimento de programas futuros.

QUADRO 4 - Primeiros sistemas comercializados nos anos de 1980

SISTEMA	GRUPO DE PESQUISA
GAT	Universidade de Georgetown (USA)
METEO	Universidade de Montreal (Canadá)
TITUS	<i>Institut Textile de France</i> (França)
CULT	Universidade de Hong Kong (China)
Systran	Systran (USA)
Logos	Logos Corporation (USA)
METAL	Universidade do Texas em parceria com a empresa Siemens (USA e Alemanha)

(Fonte: elaborado pelos autores)

Alguns projetos deixaram importantes legados, sendo precursores da retomada dos estudos em TA nos anos de 1980, pavimentando vias para a grande virada na área a partir da década seguinte. Como exemplo, está o grupo GETA (e seu sistema Ariane), que desenvolveu o projeto Eurotra, em que interações entre a TA e a linguística computacional começaram a ser mais efetivos

Nesse período, pesquisas em TA impulsionaram diálogos com áreas na linguística formalista (semântica formal, linguística matemática, coleta de dados quantitativos) e na linguística computacional.

No Japão, por exemplo, o sistema MU desenvolvido na Universidade de Kyoto por Makoto Nagao, seguindo modelo semelhante ao GETA-Ariane, destacou-se pelo uso da análise gramatical, de representações de árvores de dependência (representação visual empregada em teorias linguísticas formalistas) e desenvolvimento de um ambiente de programação para escrita gramatical. Sua pesquisa influenciou vários projetos de TA no Japão, incluindo sistemas comerciais.

Já a pesquisa conduzida na Universidade de Saarland, em Saarbrücken na Alemanha, desenvolveu o sistema SUSY que empregava diferentes técnicas, como regras de estrutura frasal, transformações gramaticais e análise de casos (elementos presentes em linguísticas formalistas como a gramática gerativa). Seu foco principal era o tratamento detalhado de línguas como russo e alemão, mas também investigou inglês e francês.

O programa de mecanização ficou marcado com grande importância para diferentes campos no que diz respeito à aplicação de computadores para realização de tarefas "intelectuais" humanas. A demanda pela TA nesse período significou uma das grandes motivações para pesquisas em computação, servindo como fundamento para o desenvolvimento de linguagens de programação.

3. O tradutor humano profissional durante o programa de mecanização

A necessidade de atender a volume crescente de traduções que não poderia ser resolvido apenas com tradução humana, bem como a demanda por agilidade e economia de custos, levou ao surgimento do programa de mecanização na TA.

Primeiros protótipos de sistemas de TA foram testados na tradução de artigos científicos, documentos que tinham grande volume impresso e exigiam tradução rápida. Com isso, a TA começou a ser idealizada porque, apesar de entregar um texto de pouca qualidade, baixa fluência e no formato de tradução palavra por palavra, ainda o fazia de maneira mais rápida em comparação ao trabalho de um tradutor humano. Consequentemente, aceitava-se uma tradução de menor qualidade, visto que se buscava entendimento geral do assunto para manter, em contexto de conflitos políticos, por exemplo, a observação dos movimentos do inimigo.

Com novas possibilidades de automatização da tradução, observa-se nova função atribuída aos tradutores: o trabalho de pós-edição da tradução gerada pela máquina. Essa função ainda é bastante executada por tradutores, revisores, linguistas, entre outros, atualmente, mesmo com uma TA de qualidade bem superior às geradas durante o programa de mecanização.

A perspectiva de idealização de uma TA perfeita (conceito de FAHQMT [*Fully Automatic High Quality Machine Translation*]) defendida por Bar-Hillel (1951) afirmava que a necessidade de pré/pós-edição realizada por um tradutor humano no processo não representaria uma TA. Essa retórica desmoralizou a TA e o trabalho de pesquisa desenvolvido naquele período, ocasionando a suspensão dos estudos por uma década, em especial nos EUA.

O programa de mecanização na TA, tal como aqui proposto, não apresentou especificamente base teórica articulada em um grupo teórico em diálogo com estudos de tradução (a rigor, estudos dessa área se institucionalizaram na década de 1970). Foi escasso o envolvimento de profissionais da linguagem (tradutores e linguistas) no desenvolvimento de ferramentas, que ficaram a cargo de especialistas das ciências exatas e da ciência da computação.

Já nos anos 1980, tradutores começaram a reconhecer, em alguns tipos de trabalho, as vantagens dos programas de computação no tratamento de dados e de dicionários. No entanto, um avanço mais significativo com engajamento de profissionais da linguagem em diálogo com aqueles já envolvidos no programa de mecanização se daria apenas nos anos 1990.

Como resultado social da presença de ferramentas computacionais durante os anos 1940-1980, o processo de tradução como atividade intelectual elitizada, gradativamente, começou a se transformar em conhecimento de domínio comum ao longo dos anos seguintes. Essa alteração social exigiu cada vez mais do tradutor o desenvolvimento de habilidades em novas tecnologias para atender a uma demanda que despontaria nas próximas décadas no mercado para serviços linguísticos em TA.

Colocar em discussão o trabalho do tradutor profissional, como proposto neste artigo em visão historiográfica, auxilia-nos na compreensão de como a relação tradutor e ferramentas foi sendo moldada em diferentes tempos históricos nos quais a tecnologia possibilitou novos horizontes e novos mecanismos. O passado, nesse sentido, é meio de compreensão de nosso presente. Na nossa época, as tecnologias de informação e comunicação passam por inovações constantes e necessitam ser objeto de reflexão a respeito, mais uma vez na história, da relação homens e máquinas.

Considerações finais

Durante os anos 1940-1980, a TA foi concebida e caracterizou-se pelo desenvolvimento de ferramentas tecnológicas pensadas para mecanização da tradução com base no processo de transformação de estruturas lógicas. Empregava-se o conceito de tradução direta, executada com base em regras gramaticais e técnica de tradução palavra por palavra. Esse processo de tradução era processado a partir do desenvolvimento de dicionários, com apresentação de terminologia específica para área e equivalentes únicos na língua de chegada que tentavam abranger ampla variedade de sentidos da língua de partida.

Num cenário intelectual de estudos linguísticos predominantemente estruturalista e formalista, grupos de pesquisa ao redor do mundo se concentraram no desenvolvimento de ferramentas de TA a partir da perspectiva das relações estruturais e regulares da língua, com técnicas para codificação e aplicação de regras gramaticais formais.

Uma necessidade de mecanização da tradução e, conseqüentemente, uso da TA nesse período buscava atender à demanda por rapidez na tradução de grandes volumes de textos, especialmente, por se tratar de contexto de guerra (Guerra Fria [1947-1991]), em que a decodificação rápida de informações e vigilância de um potencial inimigo eram cruciais.

Com isso, a baixa qualidade da tradução entregue pela máquina era compensada pela velocidade e economia de custos que proporcionava, visto que o objetivo era obter um entendimento geral de textos, em sua maioria técnico-científicos.

O processo de TA introduziu a necessidade de pós-edição de traduções geradas pela máquina, função desempenhada por tradutores humanos, na época, e ainda demandada para certos tipos de conteúdo atuais. Com a inserção do trabalho realizado pelo tradutor humano num contexto de uso de ferramentas tecnológicas, a partir dos anos 1980, presenciamos o desenvolvimento de ferramentas como dicionários personalizados, sistemas de gerenciamento de texto, memórias de tradução que facilitaram o trabalho de tradutores humanos, melhorando eficiência e consistência no trabalho.

A necessidade de mecanização da tradução e o surgimento da TA foram motivadores para desenvolvimentos futuros de estudos computacionais, com avanços significativos em linguagens de programação, consolidando bases para evolução tecnológica de diferentes áreas de estudos nos anos seguintes.

Referências

- ALPAC. (1966). *Language and machines: computers in translation and linguistics*. Division of Behavioral Sciences, National Academy of Sciences, National Research Council.
- Altman, C. (2021). *A guerra fria estruturalista: estudos em historiografia linguística brasileira*. Parábola.
- Altman, C. (2019). História, estórias e historiografia da linguística brasileira. In R. de O. Batista (Org.). *Historiografia da Linguística* (pp. 19-44). Contexto.
- Arnold, D., Balkan, L., Meijer, S., Humphreis, R. L., & Sadler, L. (1994). *Machine Translation: an introductory guide*. Blackwell.
- Bar-Hillel, Y. (1951). The present state of research on mechanical translation. *American Documentation*, 2(1), 153-165.
- Batista, R. de O. (2013). *Introdução à historiografia da linguística*. Contexto.

- Batista, R. de O. (2020). *Fundamentos da pesquisa em historiografia da linguística*. Mackenzie.
- Debonis, L. (2021). *Um estudo historiográfico da evolução da tradução automática* [Dissertação de Mestrado, Universidade Presbiteriana Mackenzie]. Adelpa Repositório Digital. <https://dspace.mackenzie.br/bitstreams/2b41d2ab-3122-4569-8111-a3286e28784e/download>
- Dostert, L. (1957). Systems of logic in machine translation. *Eighth Annual Round Table Meeting on Linguistics and Language Studies*. Washington D.
- Holmes, J. (2000). The name and nature of translation studies. [republicação de texto original de 1972] In L. Venuti (Ed.), *The Translation Studies Reader* (pp. 172-185). Routledge.
- Hutchins, W. J. (2001). Machine translation over fifty years. *Histoire Épistémologie Language*, 23(1), 7-31.
- Koerner, K. (2014). *Quatro décadas de historiografia linguística: estudos selecionados*. Centro de Estudos em Letras - Universidade de Trás-os-Montes e Alto Douro.
- Masterman, M., & Kay, M. (1960). *Mechanical pidgin translation: an estimate of the research value of "word-for-word" translation into a pidgin language, rather than into the full normal form of an output language*. Cambridge Language Research Unit.
- Nelson, B. (1967). Machine Translation: Committe Skeptical over Research Support. *Science*, 155(1), 1-2.
- Oswald, V. (1952). Word-by-word translation. *Conference on Mechanical Translation, Massachusetts Institute of Technology*. Massachusetts Institute of Technology.
- Oswald, V. (1953). An idioglossary for mechanical translation. *Modern Language Forum*, 38, 3-4.
- Poibeau, T. (2017). *Machine translation*. The MIT Press.
- Pym, A. (1998). *Method in translation history* (2nd. ed.). Routledge.
- Swiggers, P. (1981). The history writing of linguistics: a methodological note. *General Linguistics*, 21(1), 11-16.
- Swiggers, P. (2019). Historiografia da Linguística: princípios, perspectiva, problemas. In R. de O. Batista (Org.), *Historiografia da Linguística* (pp. 45-80). Contexto.
- Thomas, M. (2020). *Formalism and functionalism in linguistics*. Routledge.
- Weaver, W. (1947). *Correspondência*. The Rockefeller Foundation.
- Weaver, W. (1948). *Correspondência*. The Rockefeller Foundation.
- Weaver, W. (1949). *Correspondência*. The Rockefeller Foundation.
- Yngve, V. H. (1953). *Mechanical Translation. Quarterly Progress Report of the Research Laboratory of Electronics*. Massachusetts Institute of Technology.