

Dossiê

Estudos da tradução e IA:
diálogos (im)possíveis

Tudo o que você sempre quis saber sobre tradução automática: entrevista com a professora doutora Natália Carolina Resende

Entrevistado

Natália Carolina Resende

Trinity College | Dublin | IRL

resenden@tcd.ie

<https://orcid.org/0000-0002-5248-2457>

Entrevistadora

Marina Fonseca Darmaros¹

Universidade de São Paulo (USP) | São Paulo | SP | BR

marina.darmaros@usp.br

FAPESP

<https://orcid.org/0000-0002-9102-7748>

Formada pela UFMG, Natália Carolina Resende é professora no Trinity College, em Dublin, na Irlanda, e teve projeto financiado pelo programa Marie Skłodowska-Curie, da União Europeia, investigando como uma segunda língua pode ser reprogramada na interação com sistemas de tradução automática. Atualmente, conduz pesquisas que exploram o uso da inteligência artificial generativa na tradução literária e os impactos cognitivos dessas ferramentas sobre o tradutor. Na primeira etapa, ela e sua equipe analisaram a tradução de poesia com o uso dessa tecnologia e a conclusão foi surpreendente.

Natália, pode me contar um pouco da sua experiência até chegar ao Trinity College?

Minha experiência é bem diversa, pois sou formada em Letras-Inglês, mas tenho mestrado em *Natural Language Processing*,² uma das subáreas da Inteligência Artificial, que fiz pelo programa Erasmus Mundus, da União Europeia, parcialmente em Portugal e parcialmente na Inglaterra, já que se tratava de um consórcio de universidades. Após isso, fiz meu doutorado na Universidade Federal de Santa Catarina, em psicolinguística, com sanduíche de um ano no Max Plank Institute for Psycholinguistics, na Holanda. Esta pesquisa de doutorado também envolveu técnicas de Inteligência Artificial, principalmente alguns experimentos fazendo simulação do processamento linguístico no cérebro humano. Em meu mestrado, trabalhei

¹ Marina Darmaros é doutora em Cultura e Literatura Russa pela Universidade de São Paulo e finalizou em 2025 um pós-doutorado no PPG LETRA da FFLCH-USP e na Universidade de Bolonha (Itália) apoiado pela Fapesp (processos 2021/00445-9 e 2022/03370-2, respectivamente).

² Processamento de linguagem natural (PLN).



com tradução e também tenho uma pós-graduação em tradução pela UFMG. Minha ligação com a tradução inicialmente vem daí. Depois, vim para a Irlanda para trabalhar em um projeto de avaliação de sistemas de tradução automática, analisando 34 sistemas de tradução que foram usados na plataforma do Parlamento Europeu. Nesse projeto, usei diversas técnicas, tanto com a avaliação humana quanto com a automática dos sistemas. Fiquei um ano nesse projeto na DCU (Dublin City University), mas também trabalhei em vários projetos envolvendo tradução automática. Na DCU também fui PI (*principal investigator*)³ do meu próprio projeto de pesquisa, para o qual obtive um financiamento do programa Marie Curie para observar a influência de um sistema de tradução no processamento cognitivo de um aprendiz de inglês como segunda língua, ou seja, o quanto usar um sistema de tradução impacta ou consegue moldar ou mudar o processamento do inglês como segunda língua. Esse projeto é bastante interdisciplinar, uma mistura de tradução automática com aprendizado de segunda língua, psicolinguística e avaliação do sistema.

Depois disso, trabalhei também em outras universidades como professora. Iniciei na University of Galway como professora assistente em Ciências da Computação, mas logo surgiu uma oportunidade aqui no Trinity College para o centro de tradução literária. Hoje, sou a professora responsável pelas disciplinas com enfoque em tecnologias da tradução. Leciono uma disciplina que envolve programação em linguagem Python, outra onde ensino a trabalhar com vários *softwares* relacionados a sistemas de tradução – por exemplo, *softwares* de legendagem e as *CAT Tools* –, falo de *Machine Translation*⁴ e agora também bastante de Inteligência Artificial generativa e seus modelos, como o *ChatGPT*, treinando os alunos a entender melhor esses modelos. A pesquisa atualmente tem se voltado bastante para o campo da tradução automática, tanto de sistemas de tradução (*Machine Translation*) quanto para IA generativa usada na tradução.

Qual a diferença entre a tradução neural de ferramentas como o Google Translate, o DeepL etc. e a da IA generativa (que também é neural)?

Há uma discussão enorme sobre isso na atualidade, porque os sistemas de IA generativa não foram construídos especificamente para tradução, mas podem traduzir porque têm um vasto conhecimento de várias línguas. Como eles não são feitos para traduzir, mas aprendem a língua e o fazem com uma grande quantidade de dados textuais, eles vão detectando padrões e aprendendo como funciona determinada língua. Ao aprender uma língua (ou várias), eles conseguem entender o que está escrito em uma língua e passar aquele conteúdo para uma outra. Isso é diferente do que faz um sistema de tradução neural. Na verdade, os dois usam a mesma arquitetura de algoritmo, mas com propósitos diferentes. Os dois têm como base as redes neurais, os algoritmos das redes neurais – e existem duas grandes famílias de algoritmos: a dos algoritmos simbólicos e a conexionista, que é a família das redes neurais. Dentro da família conexionista existem várias arquiteturas de algoritmos. Então, ambos [IA generativa e Neural Machine Translation]⁵ usam a tecnologia do *transformer*, que é uma rede neural, ou seja, ambos pertencem à tradição conexionista. Mas o propósito dos sistemas é diferente: um deles é treinado especificamente para tradução e usa um *corpus* paralelo, ou seja, o texto

³ Pesquisador principal.

⁴ Tradução automática.

⁵ Tradução Automática Neural.

escrito na língua original e o mesmo texto traduzido por humanos em uma segunda língua. Portanto, o algoritmo aprende analisando o *source* e o *target*.⁶

E de onde vêm esses corpora?

Existem várias, muitas fontes de *corpus*. Por exemplo, bem lá no início da invenção do sistema de tradução automática estatística, os sistemas iniciais eram treinados, tendo como fonte o *corpus* paralelo do Parlamento Europeu, já que existia muita tradução humana nas várias línguas da União Europeia. Outro exemplo é o Canadá, onde eles também tinham um *corpus* enorme, já que eles precisam traduzir tudo para o inglês e para o francês. Hoje em dia existem vários *shared tasks*⁷ exatamente para construir esses *corpora* paralelos. Portanto, estão disponíveis vários *corpora*, por exemplo, os do WMT,⁸ que são compostos não apenas do texto-fonte e o texto-alvo, mas também do texto pós-editado.

A própria internet é uma fonte de *corpus* e, para as *low resources languages*, ou seja, as línguas de recursos baixos, os sistemas muitas vezes têm que fazer uma tradução indireta. Para traduzir, por exemplo, do russo para o vietnamita, pode ser necessário passar do russo para o inglês e depois para o vietnamita, já que é muito difícil encontrar *corpora* paralelos diretamente entre russo e vietnamita. A construção de um bom sistema depende de vários fatores, entre eles a quantidade de *corpus* que cada empresa usou para treinar e construir aquele sistema, já que quanto mais dados disponíveis, melhor vai ficando o sistema. Também tem a ver com a forma como eles avaliam o sistema, porque a construção do sistema anda lado a lado com sua avaliação. Dependendo de como é feita a avaliação, de quais os tipos de métricas usadas, isso pode interferir, pois às vezes há mais rigor com determinadas línguas.

Eles não divulgam jamais o tipo de corpus que usam?

Geralmente os sistemas comerciais, como o *Google Translate*, usam *corpus* de vários gêneros textuais, já que são *General Domain Systems*, ou seja, Sistemas de Domínio Geral –, pois também existe a possibilidade de haver um sistema especializado em um só tipo de tradução. Mas esses de que estamos falando são os de domínio geral e usam várias fontes.

E o que são os sistemas especializados?

Se você tem uma empresa especializada em tradução de textos na área de saúde, por exemplo, pode construir o seu próprio sistema de tradução automática usando um *corpus* paralelo próprio, voltado somente para aquela área – por exemplo, um *corpus* grande produzido por tradutores que você contratou ao longo da existência da sua empresa. Esse sistema de tradução, portanto, vai ser especializado naquela área e naqueles pares linguísticos.

As empresas fazem isso? Não é um investimento gigantesco?

Fazem! Não será um investimento gigantesco se ela tiver um *corpus* próprio.

⁶ A fonte e o alvo.

⁷ Tarefas compartilhadas.

⁸ O WMT (*Workshop on Machine Translation* ou Workshop de Tradução Automática) é um *workshop* acadêmico que organiza tarefas compartilhadas (*shared tasks*) em tradução automática para a avaliação de sistemas e para a comparação de performance de sistemas.

É uma espécie de CAT (Acrônimo de Computer-Assisted Translation) Tool?⁹

Não, a CAT Tool já é outro tipo de plataforma, bem diferente: é um sistema que ajuda o tradutor na tarefa da tradução e ao qual pode estar integrado um sistema já existente de *machine translation* – mas a tradução automática nele é opcional, pois se o tradutor não quiser traduzir automaticamente, ele pode traduzir tudo do zero. Portanto, na CAT Tool o profissional pode simplesmente ligar o sistema de tradução e obter toda a tradução do texto automaticamente para apenas fazer sua pós-edição ou traduzir usando uma memória de tradução – ferramenta que difere de um *machine translation*. A memória de tradução é um tipo de repositório de textos traduzidos anteriormente. Assim, quando você estiver traduzindo uma frase que dá *match*¹⁰ com algo que já tinha sido traduzido anteriormente, o resultado anterior aparece como sugestão e você pode aceitar essa solução ou não.

Um sistema de tradução especializado difere disso, porque é especializado no tipo de tradução que se quer obter, ou seja, no par linguístico para o qual foi construído e pode ser especializado em diferentes áreas ou tipos de texto.

A empresa que investe em um sistema de tradução próprio começa do zero ou ela tem um software pré-programado onde joga um corpus ou coisa do gênero?

Existem plataformas mais *user-friendly*¹¹ em que é possível fazer o sistema de uma forma bem mais simplificada do que contratar um programador especialista em construção de sistema de tradução.

São muitos os programadores que fazem sistemas de tradução especializados para determinadas empresas?

Eu não sei no Brasil, mas aqui tem bastante. Inclusive, nesse projeto de avaliação de sistemas que trabalhei, as empresas que faziam parte do projeto e que foram contratadas para construir os sistemas empregavam programadores especialistas em *Natural Language Processing* e em *deployment* de *machine translation*, que é a implantação do sistema de tradução.

Você também realizou um webinar no canal do YouTube da FFLCH-USP, em outubro de 2024, sobre tradução de poesia com IA generativa. A IA está ajudando na tradução de poesia mais do que na tradução literária de prosa?

Como nosso centro trata de tradução literária, começamos a nos perguntar aqui se esses sistemas, como parece, são mesmo um *game changer*¹² para a tradução de poesia. Isso porque um sistema de tradução normal, os comerciais, tipo *Google Translate* ou *DeepL*, não traduz rima, por exemplo, e você não pode customizar a sua tradução. Só um adendo aqui: *machine translation*, como *Google Translate* etc. também são sistemas de IA, só que são diferentes da IA generativa. Nesses sistemas de IA generativa, conseguimos, através do *prompt engineering*¹³

⁹ Ferramenta de Tradução Assistida por Computador.

¹⁰ Em tradução assistida (CAT Tools), “match” indica o nível de correspondência entre o segmento a traduzir e a memória de tradução.

¹¹ Intuitivas.

¹² Divisor de águas.

¹³ Engenharia de prompts.

customizar o *output*¹⁴ das traduções. Nos meus experimentos, analisávamos o quanto o sistema conseguia fazer o balanço em preservar o conteúdo do texto-fonte e, ao mesmo tempo, reproduzir o esquema de rimas do texto-fonte. Comparamos vários sistemas e chegamos à conclusão, naquela época,¹⁵ de que o melhor sistema era o *Chat GPT*, com o *GPT-4*, que conseguia reproduzir o esquema de rimas de sonetos que usamos como teste. Mas percebemos uma diferença de tratamento entre línguas. Por exemplo, ele reproduz imediatamente e com facilidade o esquema de rimas do inglês para o espanhol, melhor do que o do inglês ao português. Nossa hipótese é que existe uma menor quantidade de dados em português para o sistema processar e executar associações do que para o espanhol, que certamente tem uma quantidade muito maior de dados no sistema. Assim, percebemos diferenças entre sistemas de IA generativa e entre pares linguísticos. Outra questão que analisamos foi sobre qual técnica de *prompt engineering* estava sendo usada. Fizemos vários níveis de *prompts*: um bem simples, outro mais detalhado e outro ainda mais avançado. Fizemos o *zero shot*, *one shot* e os *few shot*, como chamamos na literatura especializada.¹⁶ Além disso, verificamos uma outra técnica, que se chama *chain of thought*, ou seja, o “encadeamento de pensamento”. Passo a passo, lá no sistema, primeiro você pergunta “você conhece tal poema?”, depois “você pode explicar como é a rima interna deste poema?”, e então “como é a rima final do poema?”. O que você faz, portanto, é comandar o sistema para que analise o poema até que, digamos assim, você dê o bote: “Então traduza para mim”. Para concluir, é preciso dizer “quero que você traduza e quero que você reproduza o esquema de rimas do original da forma como me explicou”. Este é um tipo de técnica de *prompt engineering*. O *prompt* chamado de *zero shot* não inclui exemplo nenhum, o *one shot* inclui um exemplo, a partir do qual o sistema aprende, e no *few shot* você dá vários exemplos ao modelo, várias demonstrações. Nos meus experimentos, por exemplo, no *few shot* eu explicava com exemplos como funcionava a rima do original, dando inclusive exemplos de palavras que estavam rimando, e depois eu mostrava no poema onde estavam essas rimas e pedia para que ele reproduzisse na tradução. Nós pudemos perceber nitidamente uma diferença do *chain of thought* para o *few shot*. Nos meus experimentos com o *Chat GPT*, o *few shot* foi a melhor técnica, a que obteve os melhores resultados.

Então quais foram suas conclusões gerais?

Fizemos outras análises e comparamos a tradução do sistema com a tradução humana. Nas traduções humanas que usamos, o tradutor humano não tenta reproduzir o esquema de rima, focando mais na semântica. Assim, analisando apenas sob essa perspectiva e considerando o que já sabemos da literatura de *machine translation*, percebemos que alguns resultados já reportados na literatura foram inclusive reproduzidos: a diversidade de vocabulário do sistema é menor que a de um humano, mas o sistema pode ser muito útil para o ser humano, já serve como um dicionário bem mais facilitado para colher ideias e inspirar uma tradução que reproduza o esquema de rimas do original, se essa for a opção do tradutor. O sistema faci-

¹⁴ Resultado.

¹⁵ Mais tarde, a professora Natália Resende especificou melhor o período em conversa pelo *Whatsapp* em 24 de junho de 2025: “Fizemos vários experimentos ao longo de 2024, mas acho que esses experimentos foram rodados em agosto de 2024 e resultados foram publicados em um *paper* em setembro de 2024” (Nota da autora).

¹⁶ Não existe uma tradução para o termo em português, que é usado exclusivamente em inglês. Poderíamos verter essas expressões como “nenhum exemplo”, “um exemplo” e “vários exemplos”.

lita a vida do tradutor humano nesse sentido. Algum dia você chegou a pegar uma tradução de poemas que ainda não tinha sido pós-editada e falar: “Nossa, isto está muito bom! Vamos publicar *assim!*”?

Sim, conseguimos bons *outputs*, eu gostei. O que percebemos é que, quando pedimos para o sistema reproduzir o esquema de rimas, ele traduz mais ou menos palavra por palavra as primeiras palavras de cada linha do poema e, para conseguir se adequar ao esquema de rimas, muda a palavra final. É esse o balanço que ele tenta fazer entre significado e forma.

Ele não faz uma troca na ordem das palavras dentro da frase em cada verso para dar a rima, como um humano talvez fizesse?

Não, ele segue a ordem certinho.

E se tem um enjambement,¹⁷ por exemplo, uma quebra da frase entre dois versos, ele reproduz também?

Ele reproduz certinho e reproduz também a forma do poema. Por exemplo, fizemos um teste com um poema escrito na forma de uma árvore de Natal e ele reproduziu na tradução a árvore de Natal. É algo que ele faz que não tínhamos com o *Google Translate* etc. Então, na minha visão, ele é um *game changer*, porque é uma excelente fonte de inspiração. Mas, claro, tudo isso vai depender do ponto de vista do tradutor. Pode ser que um tradutor, ao obter o *output* do sistema, diga: “Isso está maravilhoso, vou publicar assim!” Mas tem outros que vão querer mudar. Portanto, isso depende da escolha do tradutor e do ponto de vista dele, do que ele acha que vale a pena ou não modificar no poema.

Além dos testes em relação ao esquema de rima, fizemos testes também em relação à métrica, mas, na época, não obtivemos bons resultados. Fizemos testes usando diversos tipos de *prompts* para ver se ele reproduzia a métrica exata do original, porém não conseguimos. Mas é possível que as coisas tenham mudado um pouco com os novos modelos mais avançados.

E a tradução literária de prosa na IA generativa?

Eu nunca fiz experimentos com prosa. Estou focando nisso agora, mas com outra finalidade: analisar se a recepção desses textos em prosa traduzidos pela IA se equipara à recepção de um texto traduzido por um humano. Para tanto, estamos usando textos que não são de livre acesso. São textos inéditos, feitos por escritores que contratamos para escrever contos em duas línguas – um para escrever originalmente em inglês e o outro, originalmente em português. Vamos comparar os textos originalmente escritos na língua nativa e a recepção deles com esses mesmos textos traduzidos; depois, esse mesmo texto pós-editado; e, então, esse mesmo texto traduzido por um *Large Language Model*,¹⁸ ou seja, por uma IA generativa.

Por exemplo, temos um texto originalmente escrito em inglês e um texto comparável¹⁹ a esse texto em inglês, mas em português. Então, o texto escrito em português será traduzido

¹⁷ Cavalgamento.

¹⁸ Modelo de Linguagem Grande. Trata-se de um tipo de modelo de inteligência artificial treinado com enormes quantidades de textos, capaz de gerar, compreender e analisar linguagem natural.

¹⁹ A professora Natália Resende explicou, em conversa por *Whatsapp* em 24 de junho de 2025: “O projeto envolve analisar textos originalmente escritos em uma língua (inglês) e traduções (ao inglês). Eles são comparáveis porque estão na mesma língua e, também, porque pertencem ao mesmo gênero ou subgênero. Por exem-

para o inglês por um humano. Depois, ele também será traduzido para o inglês por meio de uma máquina, ou seja, de IA generativa, e vou pedir a um humano para pós-editar. Portanto, serão quatro textos em inglês: 1) o original (não tradução); 2) uma tradução humana; 3) uma pós-edição; 4) uma tradução apenas por IA generativa. A partir daí, analisaremos a recepção desses textos: o que o leitor acha, do que ele gosta mais, em qual ele obtém mais prazer na leitura, ele tem maior engajamento, esse tipo de métrica. Usaremos algumas técnicas utilizadas na psicolinguística, por exemplo *iTracker* etc.²⁰

Alguns estudos mostram que para a tradução literária de prosa esses sistemas estão melhores em alguns aspectos que o *Google Translate*, mas alguns erros ainda permanecem, como omissões, partes não traduzidas da fonte, alguns problemas gramaticais etc. Porém, é visível que os sistemas têm melhorado cada dia mais. Portanto, essas análises acabam tendo que ser feitas de tempos em tempos, já que a tecnologia vai evoluindo e o cenário, mudando.

Na sua opinião, pode-se dizer que tanto a tradução de IA generativa quanto a neural machine translation ainda são como uma criança que necessita estar sempre sob supervisão de um adulto (um profissional da área) na área de tradução literária?

Com certeza elas ainda precisam a supervisão de um humano. Elas ainda não alcançaram aquela *human parity*,²¹ não são páreo para os humanos. Elas não estão ainda no nível de um ser humano, principalmente na tradução literária, como verificamos, pois a tradução de poesia, para ser boa, ainda envolve uma certa edição. Essas traduções também ainda não são muito boas para expressões idiomáticas, dependendo do par linguístico, não são boas para *low resources languages*, ou seja, para línguas de recursos baixos... Existem ainda muitas questões, e uma delas é a questão da segurança: por exemplo, se uma empresa está trabalhando com dados sensíveis, ela não pode colocá-los em um sistema como o *Chat GPT*, porque é perigoso vazarem. Tudo isso envolve a assistência de humanos.

Quer acrescentar algo para finalizarmos?

Há muita coisa ainda ser explorada, muita questão a ser pesquisada. Uma delas é estudar como esses sistemas estão impactando, por exemplo, as mídias sociais. Esta é uma grande questão que ainda tem sido muito pouco abordada no meio.

plo, vou comparar texto literário só com texto literário, *short story* com *short story* [conto com conto]. Se possível, com tópicos semelhantes, o mesmo gênero do autor, porque tudo isso influencia nos padrões linguísticos que encontraremos no texto. Mas com o texto da tradução faremos em duas modalidades: o texto traduzido por um humano e o texto traduzido por uma máquina – no caso, IA generativa, que poderá vir a ser o *ChatGPT*, pois não escolhemos ainda. Este projeto é mais complexo” (Nota da autora).

²⁰ A professora Natália acrescentou ainda por conversa de *Whatsapp* em 24 de junho de 2025: “Neste experimento, o grande motivador é o fato de que os leitores em geral percebem que um texto foi traduzido, pois ele é diferente de um texto originalmente escrito naquela língua. Por quê? Porque sempre existe aquela influência da língua-fonte na língua-alvo, mesmo que o tradutor tente ao máximo fazer com que pareça a língua natural. Na literatura, a tradução é chamada de *third code*, o ‘terceiro código’. Também existe o termo *translationese*, que seria a língua da tradução. Vários experimentos, muitos na realidade, já mostraram que, estatisticamente, os padrões linguísticos de um texto traduzido são diferentes dos de um texto escrito originalmente em uma língua, a ponto de um algoritmo conseguir detectar o que é uma tradução e o que é um texto originalmente escrito em uma língua. Agora, com a emergência de textos e da tradução automática, vemos um padrão ainda mais diferente, porque ele é um *translationese* mais exacerbado ainda, ou seja, um texto traduzido ou um texto pós-editado por uma máquina é ainda mais estranho do que um texto traduzido por um humano. Portanto, ele traz um *translationese* exacerbado” (Nota da autora).

²¹ Desempenho equivalente ao de um ser humano.