

# e|ling<sup>U</sup><sub>P</sub>

Revista Eletrónica de Linguística dos  
Estudantes da Universidade do Porto  
*A Students e-journal of Linguistics  
from University of Porto*



**Diretores:**

António Leal  
Purificação Silvano

**Editores:**

Mariana Ribeiro  
Ana Luísa Fernandes  
Ari Soares  
Fábio Granja  
Glória Oliveira  
Marina Salimon  
Sofia Pinho-Silva  
Tatiana Moura

**Secretariado Editorial:**

Ana Fidelis  
Beatriz Mendes  
Bruna Costa  
Catarina Varela  
Cecília Ortiz  
Cláudia Couto  
Cristiana Dias  
David Gomes  
Deolinda Pinheiro  
Dinis Carqueijó  
Fernando Santos  
Francisca Zagalo  
Francisco Rodrigues  
Gabriel Santos  
Guendalina Gianfranchi  
Inês Lourenço  
Joana Gomes  
João Albergaria  
Jorge Carvalho  
Leonor Oliveira  
Luísa Alves  
Luísa Custódio  
Maria João Sá  
Nana Yu  
Pedro Ráinho  
Pedro Ribeiro  
Renata Rodrigues  
Tiago Azevedo  
Tomás Moreira  
Vera Ramos

**Comissão Científica:**

Alexandra Guedes Pinto  
Ana Maria Brito  
Ângela Carvalho  
Celda Morgado  
Clara Amorim  
Clara Barros  
Elena Zagar Galvão  
Fátima Oliveira  
Fátima Silva  
Idalina Ferreira  
Inês Oliveira  
Isabel Margarida Duarte  
Joana Guimarães  
João Veloso  
Luís Filipe Cunha  
Maria da Graça Pinto  
Rogélio Ponce de León Romeo  
Rui Sousa Silva  
Sónia Valente Rodrigues  
Thomas Hüsgen

## FICHA TÉCNICA

*eling*<sup>UP</sup>

Revista Eletrónica de Linguística dos Estudantes da Universidade do Porto  
*A students' e-journal of Linguistics from University of Porto*

<https://ojs.letras.up.pt/index.php/elingUP/>

Vol. 13 | N.º 2 | 2024

Periodicidade: Bidual

Logótipo e Capa por: Teresa Ribeiro

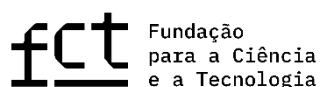
Os artigos publicados estão sujeitos a avaliação.

A Revista está registada na OJS, Bielefeld, Latindex e Worldcat OCLC e está indexada na base de dados Communication Source.

ISSN 1647-4058

DOI: [https://doi.org/10.21747/1647-4058/eling13\\_2](https://doi.org/10.21747/1647-4058/eling13_2)

Esta publicação é financiada por fundos nacionais através da FCT - Fundação para a Ciência e a Tecnologia, I.P., no âmbito do projeto UIDB/00022/2020. DOI: [10.54499/UIDB/00022/2020](https://doi.org/10.54499/UIDB/00022/2020)



## ÍNDICE

<b>Nota Editorial</b>	<b>5</b>
<b>Entrevista ao Professor Doutor Tjerk Hagemeijer</b>	<b>7</b>
<b>Artigos</b>	
<b>Tradução automática vs tradução humana: o caso da homonímia e da homografia portuguesas</b> Margarida Silva e Raquel Moreira	<b>26</b>
<b>Case study: translation from English into European Portuguese using Gender-Neutral Language. Do AI chatbots perform better than MT Systems?</b> Lisbeth Ferreira	<b>56</b>
<b>Language assessment in the normalizing society</b> Diego Oliveira	<b>82</b>

## Nota Editorial

Depois de realizar a segunda edição da *eling<sup>UP</sup> Preview*, a equipa editorial da revista *eling<sup>UP</sup>* termina o ano de 2024 com o sentimento de dever cumprido, esperança e ânimo para o que o futuro lhe reserva. A conclusão de mais um ano é também celebrada pela equipa com a publicação do 2.º número de 2024, que promete ter tanto de estimulante como de desafiador para os interessados pelos estudos linguísticos.

O entrevistado deste número é o Senhor Professor Doutor Tjerk Hagemeyer, a quem nos cabe agradecer profundamente pela disponibilidade, simpatia e, principalmente, partilhas importantes relativas à história, investigação e papel dos crioulos de base portuguesa.

Este número, do qual fazem parte trabalhos que refletem a investigação desenvolvida por estudantes de diferentes ciclos de estudos da Faculdade de Letras da Universidade do Porto e da Universidade Estadual de São Paulo, traz consigo algumas questões atinentes à Linguística Computacional, Tradução e Política Linguística.

No artigo “Tradução Automática *versus* Tradução Humana: o caso da homonímia e homografia portuguesas”, de Margarida Silva e Raquel Moreira, podemos encontrar casos de palavras homónimas e homógrafas em português cuja tradução em inglês, realizada por ferramentas de tradução automática, parece não ser satisfatória nem suficiente sem a intervenção de um tradutor humano, sendo flagrante que as palavras homógrafas são as que mais erros na tradução entre as duas línguas apresentam.

Lisbeth Ferreira assina o artigo “Case Study: Translation from English into European Portuguese using Gender-Neutral Language. Do AI chatbots perform better than MT Systems?”. Nele, a autora procura expor a complexidade que é a procura por uma escrita inclusiva relativamente ao género e, simultaneamente, as limitações e as soluções que a Tradução Automática, através da Inteligência Artificial, atualmente apresenta.

Por último, o artigo “Language Assessment in the Normalizing Society”, da autoria de Diego Oliveira, problematiza normas linguísticas presentes no *Quadro Comum Europeu*, enfatizando que traços fonético-fonológicos da língua nativa de um determinado falante são penalizados no desempenho oral de uma língua estrangeira.

Agradecemos muito a todos os avaliadores que continuam a prestar um precioso apoio à revista *eling<sup>UP</sup>*, a todos os autores que submetem entusiasticamente trabalhos para publicação e a toda a equipa, que trabalha empenhadamente para que este projeto continue a colecionar leitores interessados, a quem também endereçamos um agradecimento particular. A

contribuição de todos permite que a *eling<sup>UP</sup>* atinja agora a maioria - de facto, esta revista científica conta já com dezoito anos de existência.

Feçamos com chave de ouro e o nosso desejo é que 2025 seja repleto de tantas realizações e conquistas como o ano que termina. Esperamos que se deleitem com a leitura deste número e que procurem sempre saber mais sobre os estudos da linguagem humana e das línguas, que não se esgotam nunca e onde há sempre muito por explorar, questionar e deixar-se maravilhar.

## Entrevista ao Professor Doutor Tjerk Hagemeyer



O Professor Doutor Tjerk Hagemeyer é Professor Associado do Departamento de Linguística Geral e Românica da Faculdade de Letras da Universidade de Lisboa e investigador do Centro de Linguística da Universidade de Lisboa. Possui Mestrado em Linguística pela Universidade de Lisboa e obteve, na mesma instituição, o Doutoramento em Linguística com especialização em Linguística Geral, em 2007, com a dissertação *Clause structure in Santome*. Ao longo da sua carreira, tem desenvolvido vários projetos de investigação sobre crioulos de base lexical portuguesa - em especial os do Golfo da Guiné - e variedades do português em África. Atualmente, é membro do Conselho de Escola da Faculdade de Letras da Universidade de Lisboa e representante dos docentes da Área de Ciências da Linguagem no Conselho Pedagógico da Faculdade de Letras da Universidade de Lisboa. Desde 2019, ocupa, também, o cargo de tesoureiro na Associação de Crioulos de Bases Lexical Portuguesa e Espanhola (ACBLPE), tendo já sido Vice-Presidente de 2003 a 2007 e Presidente de 2007 a 2011. É ainda membro da Comissão de Curso de Estudos Africanos da Faculdade de Letras da Universidade de Lisboa. Foi, de 2014 e 2016, investigador responsável do grupo de investigação ANAGRAMA do Centro de Linguística da Universidade de Lisboa.

A entrevista ao Professor Doutor Tjerk Hagemeijer teve lugar no dia 26 de junho de 2024, no Centro de Linguística da Universidade do Porto (CLUP), tendo sido conduzida pelos estudantes Fábio Granja, Glória Reis, Rute Rebouças e Tatiana Moura. A entrevista foi preparada pelos estudantes Ana Fidelis, Fábio Granja, Marina Araújo, Pedro Raínho, Renata Rodrigues e Tatiana Moura e, posteriormente, transcrita pelos estudantes Catarina Varela, Cláudia Couto, David Garcia, Fernando Santos, Maria João Sá, Pedro Raínho, Tiago Azevedo e Tomás Moreira.

A Equipa Editorial gostaria de expressar o seu agradecimento ao Senhor Professor pela disponibilidade e generosidade na concessão desta entrevista. Foi deveras gratificante a oportunidade de testemunhar a partilha de conhecimentos linguísticos, memórias e preciosos conselhos.

*Qual considera que foi a sua motivação inicial para estudar os crioulos de base portuguesa do Golfo da Guiné? Houve algum evento ou acontecimento na sua vida que o tenha levado a interessar-se por esta área especificamente?*

Pensei um pouco sobre essa questão. Desde sempre, desde muito jovem, quando vivia nos Países Baixos, tive sempre um grande interesse por África. Não sabia que existiam crioulos no Golfo da Guiné, não sabia que existiam crioulos sequer, mas interessava-me. Lia obras de literatura africana, especialmente da literatura sul-africana, mas também de outros países. Lá em casa recebíamos uma revista que era da Cooperação Internacional, que continha muitas matérias que tratavam de África nas suas múltiplas vertentes.

Depois da minha escola secundária, fiz uma espécie de *gap year* em Portugal, a aprender português para estrangeiros na FLUL. No regresso aos Países Baixos, fiz o primeiro ano da licenciatura em Língua e Cultura Portuguesa (*Portugese Taal- en Letterkunde*) na Universidade de Utrecht, um curso que mais tarde foi extinto devido a cortes no ensino superior. Depois desse ano, regressei definitivamente a Portugal para fazer a licenciatura em Língua e Cultura Portuguesa na FLUL. Durante a licenciatura, tinha colegas de diferentes países de língua oficial portuguesa, a quem me aliei em alguns momentos para fazer trabalhos sobre crioulos de base lexical portuguesa. Tinha colegas que eram falantes de línguas crioulas. Uma dessas colegas, que continua a ser uma pessoa muito amiga, era a Professora Beatriz Afonso, de São Tomé e Príncipe. Penso que terá sido com ela que fiz o primeiro trabalho

sobre crioulos, na cadeira de Linguística Histórica, e especificamente sobre o forro (crioulo de São Tomé), que continua a ser a minha língua de eleição. Com outros colegas, que incluíam um colega guineense, fizemos um trabalho sobre o kriol (crioulo da Guiné-Bissau) na cadeira de Fonologia. Na licenciatura, também fiz a cadeira de Crioulos de Base Lexical Portuguesa com a Professora Dulce Pereira.

Quando terminei a licenciatura, hesitei em seguir na área das literaturas ou da linguística. Optei pela linguística, por achar que era uma área mais objetiva do que a literatura. Candidatei-me ao mestrado em Linguística Teórica e, no dia da entrevista, perguntaram-me em que gostaria de trabalhar. Atirei barro à parede e disse ao júri que tinha pensado na aquisição do português por falantes do neerlandês ou talvez em algo relacionado com crioulos. Os professores do júri, Inês Duarte, Ernesto d'Andrade e Alina Villalva entusiasmaram-se com a ideia dos crioulos, porque não havia muito trabalho nessa área. E foi assim! Durante o mestrado, no fim do milénio passado, fui de mochila às costas para São Tomé durante dois meses, sem ter muito conhecimento das línguas que se falavam por lá. Com um gravador, ainda com cassetes, e com a ajuda de pessoas locais, comecei a gravar o forro, com muitas peripécias pelo meio. Pouco a pouco, comecei a aprender a língua, a transcrever dados, com a ajuda dos meus “capitães de campo”, primeiro Luís Morais e, depois, Caustrino Alcântara, e a descrever e analisar estruturas sintáticas da língua. Apesar de ser de base lexical portuguesa, é uma língua muito afastada do português e inacessível para um falante do português. (Para ouvir e saber mais sobre crioulos, recomendo vivamente uma visita a: <https://apics-online.info>)

Foi assim que comecei a trabalhar na área do contacto de línguas e sobre crioulos, os crioulos do Golfo da Guiné em particular. Esta é a história abreviada da minha experiência, da minha vivência e do meu interesse por estas línguas.

*Além da investigação, o Senhor Professor tem também dedicado muito do seu tempo à docência, lecionando atualmente disciplinas como Crioulos de Bases Lexical e Linguística Comparada: Português em Contacto. No seu entender, qual é a importância da relação docência-investigação? Como tem sido a sua experiência como docente-investigador até ao momento?*

A primeira coisa que é preciso dizer é que a docência e a investigação são duas atividades que claramente se complementam. Para sermos docentes atualizados e também

motivados, precisamos de nos manter ativos na área da investigação, porque muitas vezes o que acontece é ouvir-se: “Bem, lá vou ter que dar de novo esta matéria”. A investigação faz parte da solução para superar este problema. Queremos descobrir mais e, ao mesmo tempo, transmitir novos conhecimentos aos estudantes, pelo que um trabalho ativo na investigação é também um trabalho que contribui para nos mantermos motivados na área da docência. Também é verdade que, sendo docente, é preciso conciliar as aulas, diversas outras responsabilidades e a investigação. E há várias formas de fazer investigação: uma coisa é irmos a São Tomé com a mochila às costas e fazer trabalho de campo; outra coisa, também importante quando somos docentes no ensino superior, é conhecer os trabalhos dos colegas. Faz parte da investigação estarmos atualizados. Há ainda uma outra faceta, na parte da docência, que são as questões, fáceis ou difíceis, que os estudantes nos colocam. Por exemplo, o vosso guião tem algumas questões que não são tão simples de responder e que podiam ser perguntas feitas em contexto de sala de aula (*risos*). Isto obriga-nos a estar em alerta e obriga-nos também a investigar e a reinvestigar, para estarmos preparados para o desafio que é enfrentar turmas com diferentes perfis de estudantes. Portanto, creio que não há ensino superior sem esta relação de complementaridade entre a investigação e a docência. Respondendo à segunda parte da questão, tem sido, para mim, uma experiência boa e gratificante, porque Crioulos de Base Lexical Portuguesa ou Português em Contacto, na área da Linguística Comparada, ou mesmo Multilinguismo e Política Linguística, são unidades curriculares que me interessam diretamente e me permitem fazer este trabalho de investigação. Obviamente, às vezes temos que ensinar disciplinas para as quais não estamos tão habilitados ou que não nos despertam tanto interesse, mas isso também faz parte. Aí sim, é mais difícil manter esta relação estreita entre a docência e a investigação. Portanto, quando se trata - e falo por mim e em geral - de unidades curriculares mais da nossa especialização, acho que funciona bem esta relação entre a docência e a investigação. Se nos mantivermos ativos através da investigação e na relação com os estudantes, é mais fácil manter a motivação em alta.

*Tendo em conta a extensa produção do Senhor Professor na área dos crioulos de base lexical portuguesa e do português em África e atendendo também ao facto de que este é, do ponto de vista linguístico, um campo de investigação relativamente recente e complexo, gostaríamos de saber quais considera serem os maiores desafios dessa área? Que tendências linguísticas, no domínio da sintaxe, tem a investigação revelado?*

Esta é uma dessas perguntas difíceis e também mais pessoais. Eu diria que, quando fui para o mestrado em Linguística e quando comecei a trabalhar sobre crioulos, tive sempre como uma das motivações o facto de ser uma área relativamente pioneira. Através da minha própria investigação e do trabalho sobre crioulos, comecei a perceber que estas línguas são muitas vezes línguas minorizadas no contexto em que são faladas. Portanto, um dos grandes desafios nesta área é contribuir, de alguma maneira, para a valorização e a manutenção da vitalidade destas línguas. Não me refiro tanto a Cabo Verde ou ao cabo-verdiano, nem ao kriol da Guiné-Bissau, que têm uma grande vitalidade e são o orgulho dos falantes. Mas os outros crioulos de base lexical portuguesa, os do Golfo da Guiné, falados em São Tomé e Príncipe e Annobón, e todos os crioulos da Ásia são línguas ameaçadas.

Obviamente, cada investigador terá as suas motivações. Há quem queira simplesmente fazer descrição linguística, sem pensar muito no resto, mas, quando trabalhamos com a comunidade e quando fazemos trabalho de campo, também compreendemos a importância que uma língua tem para a comunidade e, percebendo que essa língua está de alguma maneira ameaçada, nós, linguistas, podemos ter um papel na sua vitalização e documentação. Portanto, eu acho que esse é um dos grandes desafios em relação aos crioulos de base lexical portuguesa, enquanto ainda formos a tempo, porque algumas dessas línguas vão extinguir-se, sem qualquer dúvida. Vou dar-vos o exemplo de um crioulo da Índia, o crioulo de Cananor, no Sul da Índia, que tem neste momento cinco ou seis falantes. Uma língua que tem cinco ou seis falantes é uma língua que vai desaparecer. Mas há outras situações, como o caso dos crioulos do Golfo da Guiné: o forro, por exemplo, é uma língua ainda falada por, talvez, um terço ou metade - os dados nem sempre são fiáveis - da população de São Tomé. Aqui ainda há oportunidades para tentar promover, revitalizar e dar uma nova vida a uma língua que é ameaçada pelo português. É também muito importante que estados como São Tomé e Príncipe percebam que é perfeitamente possível criar mecanismos para gerir o multilinguismo, que o multilinguismo faz bem e que não é, como se pensava no passado, algo negativo para os próprios falantes. Portanto, acho que este é, talvez, o principal desafio para quem trabalha na área dos estudos sobre crioulos.

Um outro grande desafio, mais pessoal, talvez, é tentar chegar às origens destas línguas. Como surgiram estas novas línguas, que resultam do contacto entre o português e outras línguas? E como foi este processo historicamente? De onde vieram as pessoas? Porque é que a língua é como é? Felizmente, tenho tido a oportunidade, através do meu trabalho e através de um projeto que coordenei há uns anos, de colaborar com historiadores e

geneticistas, para tentarmos perceber melhor como é que surgiram os crioulos do Golfo da Guiné, como foi este processo a partir do momento do primeiro povoamento, em finais do século quinze, na ilha de São Tomé. É uma matéria fascinante, porque se cruzam pessoas, línguas e culturas, e o resultado é algo de totalmente novo, mas que não deixa de ter origens. Os são-tomenses pensam muitas vezes que têm uma origem mais ligada a Angola, mas, na verdade, do ponto de vista linguístico, genético e histórico, a sua origem está mais ligada à Nigéria. O nosso trabalho interdisciplinar colaborativo foi uma oportunidade para aprofundar o que são estas línguas, de onde vieram, e porque são como são. Enfim, como é que uma língua de base lexical portuguesa como, por exemplo, o forro, consegue ser tão diferente do português?

Portanto, estes são alguns dos desafios, mas há quem tenha como desafio fazer uma boa descrição linguística, uma área em que continua a haver, para os crioulos, falta de investigação. Há inúmeros trabalhos sobre o português europeu, sobre o português do Brasil e sobre línguas mais conhecidas em geral. Mas, se alguém quiser, de facto, abraçar um desafio um bocadinho diferente e que tem uma ligação com o português, ainda que às vezes um pouco remota, os crioulos de base lexical portuguesa oferecem muitas oportunidades. Sabemos muito mais do que há umas décadas, mas ainda falta saber muito.

*Sabemos que o Senhor Professor desempenhou o cargo de Presidente da Assembleia Geral da Associação de Crioulos de Bases Lexical Portuguesa e Espanhola (ACBLPE), entre 2007 e 2011. Gostaríamos de saber que balanço faz desta experiência? Que projetos gostaria ainda de ver a Associação concretizar?*

Na verdade, a minha vida tem quase um cordão umbilical que me liga a esta Associação (<https://acblpe.com>). Eu sou também um dos fundadores da Associação, que foi constituída em 2003. Durante os primeiros anos, fui vice-presidente, na presidência do Professor Alan Baxter, agora sou tesoureiro e desempenhei também o cargo referido na pergunta. Tenho estado muito presente na Associação ao longo de duas décadas. Inicialmente, houve quem achasse que a Associação não tinha futuro, mas, na prática, tem havido bastante interesse nos encontros que organizamos em diferentes locais do mundo, na Europa, em África e também nas Américas, por vezes em colaboração com outras associações, como a Associação Brasileira de Estudos do Contacto Linguístico (ABECS)<sup>1</sup>, no Brasil, e a Society

---

<sup>1</sup> <https://aovivo.abralin.org/lives/estudos-da-sociolinguistica-de-contato/>

for Pidgin and Creole Linguistics (SPCL)<sup>2</sup>, nos Estados Unidos. Portanto, a Associação tem vitalidade e há pessoas que se interessam pelos crioulos de base lexical portuguesa e espanhola e pelas situações de contacto que envolvem o português e o espanhol, embora sejam os crioulos portugueses que têm atraído mais investigadores e interessados. Também se verifica que há sempre jovens investigadores que estão a ingressar nesta área e procuram a Associação para apresentar o seu trabalho, o que mostra que existe renovação. Portanto, a Associação, não sendo grande nem muito profissional, é uma associação de investigadores que partilham motivações semelhantes: muitos fazem trabalho de campo e querem partilhar o que descobriram. A Associação também aceita trabalhos sobre as variedades do português em África ou sobre variedades do espanhol na América, porque é sempre o contacto com estas línguas que nos move.

Qual é o futuro da Associação ou o que é que a Associação pode fazer mais ou melhor? Talvez criar uma ligação mais forte às comunidades, isto é, tentar trabalhar de uma forma mais direta com as comunidades, mostrando que as suas línguas têm valor e que estamos a desenvolver trabalho sobre elas. Mas é também importante levar o nosso trabalho aos falantes, à comunidade. Este ano, por exemplo, o encontro vai ser na ilha de São Tomé (julho de 2024). Da minha própria experiência com São Tomé e Príncipe, posso dizer que os resultados dos trabalhos de investigação nem sempre voltam às ilhas. Na verdade, as pessoas em São Tomé não sabem muito bem o que se tem feito sobre as suas línguas. Acho que se pode fazer um maior esforço para mostrar que esse trabalho existe e que o conhecimento produzido pode ser aproveitado e utilizado localmente. Isto também vai dar mais força à comunidade e espera-se que sirva de estímulo aos agentes políticos e educativos locais, no sentido de se fazer mais em prol das línguas, nomeadamente daquelas que estão ameaçadas.

*Qual considera ser a importância de trabalhar para a recolha, descrição e dicionarização de línguas crioulas de origem portuguesa? No caso das línguas do Golfo da Guiné, que contam hoje com poucos falantes, há expectativa de revitalização ou o foco concentra-se em conservação e documentação?*

Bom, é uma questão muito ligada às respostas anteriores. Em primeiro lugar, no caso dos crioulos do Golfo da Guiné, falando mais especificamente de São Tomé e Príncipe, estamos a falar de três crioulos autóctones: o forro, o angolar e o lung'ie (literalmente, a língua da ilha, o crioulo do Príncipe). Estas línguas apresentam desafios diferentes. O crioulo

---

<sup>2</sup> <https://www.spclsociety.com/>

do Príncipe é, destas línguas, aquele que se encontra em maior risco de extinção. Tem, no máximo, algumas centenas de falantes, mas também é, ao mesmo tempo, objeto de uma revitalização mais ativa. Tem havido algumas iniciativas por parte do governo regional da ilha do Príncipe no sentido de promover e tentar restabelecer a transmissão entre gerações, porque as pessoas que falam essa língua são geralmente pessoas mais velhas. Neste momento, estas pessoas trabalham com os jovens nas escolas e creio que até mesmo nas creches, no sentido de os mais novos voltarem a ter contacto com uma língua que tem uma presença histórica na ilha do Príncipe, sendo património único e exclusivo desta ilha.

Às vezes, tenta-se revitalizar uma língua, mas nem sempre se consegue. Nem sempre a população ou comunidade mostra vontade e interesse, porque quer falar a língua maioritária, que, neste caso, seria o português, a língua mais amplamente falada em São Tomé e Príncipe. Portanto, uma língua com poucos falantes constitui um desafio exigente, porque obviamente se pode colocar a questão: “Para quê falar o lung’ie?”. Por outro lado, também temos cada vez mais consciência da importância que o património linguístico tem. Acho que estamos num mundo em que o bom senso dita que nos devemos pautar pela diversidade. A diversidade linguística é algo que devemos preservar, tal como queremos preservar a diversidade biológica, da fauna e da flora. Costumo dizer aos estudantes que há mais documentários sobre animais em vias de extinção do que sobre línguas em risco de extinção; todos os anos morrem muitas línguas, mas raramente é notícia. A língua é algo intrinsecamente humano e nem sempre lhe damos a devida atenção.

Portanto, mesmo havendo uma pequena comunidade de falantes, como no Príncipe, existem possibilidades de revitalização. Será talvez mais fácil fazer o mesmo em relação às outras duas línguas que se falam em São Tomé e Príncipe, o forro, que tem mais falantes, e também o angolár. No entanto, neste momento, infelizmente, não há iniciativas tão dedicadas como há na ilha do Príncipe. Muitas vezes o que acontece é que estas iniciativas só aparecem quando as pessoas percebem verdadeiramente que a sua língua está a desaparecer. Em São Tomé, costuma dizer-se “os crioulos são o nosso bilhete de identidade”, mesmo aqueles que não falam ou mal falam a língua. Portanto, é, antes de mais nada, uma frase feita. Quando a oíço, costumo dizer às pessoas que, então, o seu bilhete de identidade está caducado e que é preciso ir renová-lo (*risos*). Enfim, o tempo dirá o que vai acontecer, mas as iniciativas em prol do forro e do angolár têm sido muito poucas. Eu próprio colaborei numa proposta de escrita para os três crioulos autóctones de São Tomé e Príncipe, mas, na prática, muito poucas

pessoas escrevem nessas línguas, apenas alguns músicos e algumas pessoas ligadas à área da cultura. A maioria nem sequer sabe que há um decreto-lei que oficializou uma escrita para estas línguas. É uma realidade muito diferente da cabo-verdiana, onde a escrita da língua tem marcado o debate público e ela é utilizada na prática.

*No seu projeto “Posse e localização: microvariação em variedades africanas do português (PALMA)”, refere que o objeto de estudo do mesmo se foca na contribuição da nativização do português na emergência de traços linguísticos em variedades urbanas contemporâneas faladas em Angola, Moçambique e São Tomé e Príncipe. Pode explicar, mais concretamente, o funcionamento deste processo?*

Bom, agora saltamos dos crioulos para as variedades do português que são faladas em África, que também, na verdade, constituem uma área de investigação muito interessante e em que tem havido, sobretudo desde a década de 1980, cada vez mais trabalho. Especificamente nestes três países, assistimos a um processo de nativização. Quer dizer que, historicamente, as populações nesses países eram falantes de outras línguas que não o português. No caso de Angola e Moçambique, sobretudo línguas do grupo bantu; no caso de São Tomé e Príncipe, havia sobretudo falantes dos crioulos. Mas, por razões históricas, que são diversas e diferentes para cada um destes estados, o português tem vindo a ocupar um papel cada vez mais relevante. Primeiro, como língua não materna (L2), mas, sobretudo no decorrer do século XX e depois das independências, também como língua materna. Vou-vos dar o exemplo de São Tomé e Príncipe. Neste país, os crioulos autóctones de que falámos acima eram historicamente as línguas mais faladas, mas, no fim do século XIX, duas culturas ganham grande relevância: a cultura do café e, especialmente, a cultura do cacau. Nessa altura, a escravatura já tinha sido abolida e o regime colonial não conseguia mão-de-obra local suficiente para trabalhar nas plantações de cacau e de café. Foram buscar trabalhadores, os chamados serviçais ou contratados, a outros países, em particular a Cabo Verde, Angola e Moçambique. Os contratados que foram para São Tomé eram falantes de outras línguas, mas, em São Tomé, no contexto das chamadas roças (empresas agrícolas), propriedade de portugueses, começaram a utilizar o português como língua veicular. O número de contratados era tão grande que a população de São Tomé mais do que duplicou. Devido ao contexto das roças, os contratados geralmente não adotaram o forro como a sua nova língua, mas sim o português, que, desta forma, se começa a afirmar cada vez mais no contexto de São

Tomé e Príncipe. Já durante a chamada Segunda República, nos anos de 1940/1950, houve alguma repressão oficiosa dos crioulos. Há muitos relatos de pessoas mais velhas que diziam que, em casa, os filhos não poderiam falar crioulo com os pais. A independência significou o fim do regime colonial, mas o acesso à língua portuguesa aumenta substancialmente, sobretudo através da massificação do ensino. E a população começa a transitar cada vez mais do português L2 para L1. Hoje em dia, em São Tomé e Príncipe, a larga maioria da população tem o português como língua materna, em consequência de uma reviravolta social e linguística que teve início em finais do século XIX.

Em Angola, também se observa uma forte nativização. Encontramos, hoje em dia, muitos monolíngues em português, sobretudo nos centros urbanos. Em primeiro lugar, no tempo colonial, a presença portuguesa foi sempre bastante significativa. Estima-se que, em 1973, 10% da população angolana era daqui, de Portugal, o que significa que havia muitos falantes do português. Depois da independência, rebentou o conflito armado que durou de 1975 até 2003, obrigando muita gente a refugiar-se nas cidades. Houve um êxodo do interior para as cidades, sobretudo Luanda, onde o português já funcionava como língua franca para os falantes de línguas bantu. No fundo, há aqui um efeito de bola de neve, com cada vez mais falantes do português L2 e novas gerações para as quais o input L2 passa a ser de L1. Este processo está em curso, com ganhos para o português e custos para as outras línguas. Em São Tomé, o processo de nativização está praticamente concluído, ou seja, poucos falantes ainda têm como L1 um dos crioulos. Em Angola, o português é hoje falado por mais de 70% da população como L1 ou L2. Por comparação, o umbundo, a língua bantu mais falada, é falada por cerca de 22% da população angolana. Tudo isto mostra bem que, num período de tempo relativamente curto, a situação sociolinguística de um país pode mudar drasticamente. Apesar de o processo de nativização em Moçambique ser mais lento, os censos mostram que há cada vez mais falantes do português, incluindo cada vez mais falantes L1.

Portanto, temos de estar preparados para um futuro que é sobretudo em língua portuguesa nestes estados, o que significa que é importante investigar o que os falantes fazem com o português do ponto de vista linguístico. Isto pode ser importante para efeitos educativos, por exemplo. Como é que abordamos as diferenças relativamente à norma do português europeu em vigor, mas que não é, geralmente, a norma praticada? Como é que se pode sanar esta tensão que existe entre uma norma de fora e o uso da língua? Enfim, temos de olhar com especial atenção para estes novos falantes nativos e fazer o nosso trabalho de investigação, para tentar perceber o que são e o que vão ser estas novas variedades no futuro.

Será que se vão aproximar do português europeu? Será que se vão afastar? Será que podem tornar-se independentes, como aconteceu, por exemplo, com o inglês americano? Já agora, o português do Brasil é também um ótimo exemplo que mostra que os estados e a sua língua, mesmo que seja uma língua com um passado colonial, podem adquirir a sua independência linguística. Havendo cada vez mais falantes nativos do português em África, o debate em torno desta questão também se vai tornar cada vez mais premente: será que Angola deve seguir eternamente a norma do português europeu? Ou será que quer olhar para o que é o português de Angola e, eventualmente, com o tempo (ainda há muitos constrangimentos), considerar a possibilidade de conquistar uma maior autonomia para a sua própria realidade linguística? Em Angola, estas questões começam a marcar a agenda, mas só o futuro dirá qual será o destino da língua portuguesa em Angola. Esta questão da descentralização das constelações pluricêntricas é uma questão global que não se esgota na língua portuguesa.

*Apesar das variantes regionais distintas, o crioulo cabo-verdiano é considerado como uma única língua. Que critérios linguísticos são normalmente usados para determinar se as diferentes variantes do crioulo, como as encontradas em Cabo Verde, devem ser classificadas como dialetos separados ou como parte de um único idioma?*

Como sabem tão bem como eu, o conceito de língua é o que nós queremos que seja. Portanto, uma língua é uma construção, é um conceito bastante abstrato - para definir o que é uma língua, não podemos simplesmente basear-nos em critérios linguísticos. Provavelmente, se a Galiza fizesse parte de Portugal, iríamos dizer que se falava ali uma variedade do português. Não é aceitável, porque o galego é uma língua oficial, consagrada na constituição de um outro estado. Quando olhamos para um contexto como o de Cabo Verde, vemos que há nove ilhas habitadas com variedades do cabo-verdiano por vezes bastante distintas, sobretudo entre o Barlavento e o Sotavento, por exemplo a variedade da ilha de São Vicente e a da ilha de Santiago. Em Cabo Verde, estas diferenças têm inclusive resultado num debate sobre qual deve ser a escrita e a norma para o cabo-verdiano. Esta questão já está resolvida para muitas línguas, como o português, mas línguas que começaram a ser planeadas em tempos mais recentes ainda estão a passar por um processo de normalização. Como é que se vai escrever a língua, qual é a variedade da língua que se vai adotar, se é a desta ou daquela ilha, se é da capital ou de áreas mais rurais, se é o crioulo, como as pessoas às vezes dizem, mais “puro”, mais “conservador”, mais “fundo”, etc. Portanto, colocam-se desafios que não se colocam da

mesma maneira para línguas como a minha (o neerlandês), o português e muitas outras. Mas, obviamente, Cabo Verde é um estado e, apesar da variação, que existe em todas as línguas, tem interesse em manter uma certa unidade em relação à língua materna das pessoas: todos os cabo-verdianos têm uma dessas variedades como a sua língua materna e é do interesse do Estado cabo-verdiano chamar ao conjunto das variedades língua cabo-verdiana. É uma decisão política. Podemos imaginar muitos cenários: poderíamos imaginar que os Açores conseguissem a independência em relação a Portugal e que talvez quisessem fundar a sua própria língua - o açoriano. Este exemplo é um bocadinho absurdo, mas pensem, por exemplo, na antiga Jugoslávia, onde falavam uma língua que era rotulada de servo-croata e que todos designavam como servo-croata. Mas, desde a dissolução da Jugoslávia, falamos cada vez mais em croata, umas das línguas oficiais da União Europeia, e sérvio, que um dia também poderá vir a ser uma das línguas oficiais da União Europeia. O sérvio e o croata estão num contínuo dialetal e são praticamente o mesmo, porém separados por uma fronteira política. No caso de Cabo Verde, existe um contínuo dialetal insular complexo, mas do ponto de vista político é do interesse do Estado manter esta unidade que é a língua cabo-verdiana e ter uma única norma.

*É conhecido que Cabo Verde tem debatido o processo de oficialização da língua crioula como um grande dilema político e linguístico com forte divisão de opiniões. O Senhor Professor gostaria de comentar este assunto e seus possíveis impactos sócio-educacionais?*

Não gostaria, mas vou (*risos*). Em primeiro lugar, é importante valorizar que os cabo-verdianos têm debatido muito entre eles em relação à sua língua materna - acho que isso é um sinal da própria vitalidade da língua (é, realmente, o crioulo de base lexical portuguesa com mais vitalidade, tanto em Cabo-Verde como na diáspora). É um claro sinal positivo em relação à própria língua. É também verdade que se trata de instituir algo novo: a língua oficial é, neste momento, o português, mas prevê-se a oficialização do cabo-verdiano num futuro não tão distante. E colocam-se questões (de certa forma já ultrapassadas) como: “Como escrever a língua?”, “Qual é a variedade que devemos adotar?”. É positivo que haja, de facto, este debate, que as pessoas tenham orgulho de saber falar crioulo, de serem cabo-verdianos, falantes do crioulo.

Em São Tomé, por exemplo, existe uma relação diferente entre as pessoas e as línguas. Por vezes os são-tomenses mostram ter alguma vergonha de serem falantes do crioulo (e do

português que falam). É uma realidade completamente distinta. Na minha opinião, o mais importante nestes processos é agir. Não vale a pena debater eternamente sobre se devemos escrever uma determinada palavra com “c” ou com “k”, ou de uma determinada forma ou de outra, mas sim tomar uma decisão política. É isso que faz progredir as línguas. Ninguém espera que o cabo-verdiano hoje seja igual ao cabo-verdiano de amanhã, tal como o português de hoje não é igual ao português de ontem. As línguas mudam e a normalização das línguas tem de ser revista de tempos a tempos.

Em Portugal, passou-se (ou passa-se?) por uma reforma ortográfica polémica, mas isto faz parte da vida das línguas. Na minha opinião, haver um debate sem que se tomem decisões para um futuro melhor da língua não faz muito sentido. É preciso agir, é preciso implementar e, se não correr tão bem, revê-se, faz-se um trabalho de revisão, adapta-se: este parece-me ser sempre o caminho. Portanto, eu ficaria muito contente se, em São Tomé e Príncipe, houvesse uma atitude muito mais proativa em relação às línguas, como existe em Cabo Verde, para que as pessoas tomassem mais consciência dos desafios que a língua enfrenta e da importância que é tomar medidas efetivas para resolver esta questão da sua minorização.

Portanto, sim, o debate é bom, mas não pode ser um eterno politiquês. É preciso que se avance, para que, de facto, a língua possa beneficiar, que não seja cada uma das capelinhas que quer que seja assim ou assado, mas que, de facto, a própria língua colha os frutos das decisões coletivas, da oficialização. No caso de Cabo Verde, provavelmente também haverá instituição do cabo-verdiano no ensino: já houve uma experiência-piloto de ensino bilingue, mas faz muito sentido que se faça mais. Sabe-se que a alfabetização e o ensino na língua materna produzem sempre resultados melhores do que a alfabetização e o ensino em línguas que não sejam as línguas maternas da população.

Ocorre-me o exemplo da Guiné-Bissau. Neste país, mais ou menos 90% da população fala o crioulo (ou guineense), de acordo com o censo de 2009. E apenas uns 30% da população indicou saber falar português. O português é a língua oficial, é a língua das escolas, mas, na verdade, muitos guineenses não têm, a não ser no contexto da escola, contacto com a língua portuguesa. A Guiné-Bissau é, na minha opinião, um exemplo paradigmático de um estado onde funciona uma política linguística errada. É um país com muitos problemas, de vários níveis, mas que poderia seguir um caminho diferente se adotasse o crioulo como língua oficial e se comesse a alfabetizar e a ensinar nessa língua. Isso não significa que o português não possa manter-se como uma língua estrangeira especial, como uma língua que também faça parte do currículo. Talvez não como o inglês aqui em Portugal, mas como uma língua que tem

um estatuto histórico especial no contexto da Guiné-Bissau. Portanto, não se trata de uma proposta de “acabar com a língua portuguesa”. É, na verdade, uma política linguística que pensa mais na população e também no desenvolvimento, porque a questão da língua está ligada ao desenvolvimento. É preciso criar mecanismos para as pessoas poderem beneficiar da alfabetização e do ensino na sua língua materna.

Em Moçambique, depois da independência, adotou-se o português como única língua oficial e os problemas de insucesso escolar e de abandono escolar não demoraram a manifestar-se. Ou seja, Moçambique percebeu que o papel exclusivo do português colocava em causa o desenvolvimento. Na década de 1990, implementou-se um projeto de ensino bilingue, que envolvia o português e 16 línguas bantu que se falam em Moçambique. Nos primeiros anos da escolaridade, as crianças começam a aprender a ler, a escrever, a fazer contas, etc., na sua língua materna. O português também faz parte do currículo, porque, depois, a partir do quarto ano, há uma transição para o ensino exclusivamente em português. Portanto, é uma política de transição da L1 para a L2, que é sempre preferível àquela política em que um aluno, um jovem de seis anos, chega à escola e, de repente, vai começar a escolaridade numa língua com a qual teve, às vezes, pouco ou nenhum contacto prévio.

Em suma, no caso dos países de língua oficial portuguesa em África, vemos diferenças, por exemplo, do ponto de vista da nativização e do ponto de vista das políticas linguísticas. Penso que é importante que os diferentes Estados reflitam sobre o que é que é de facto melhor para a população, como é que a linguística pode contribuir para o desenvolvimento. Esta questão é muitas vezes marginalizada ou secundarizada. Mesmo no caso de Moçambique, é discutível se é ideal seguir o modelo adotado por várias ex-colónias britânicas em África, que é de começar a escolaridade na língua materna e depois transitar para a língua oficial, o inglês, porque é preciso reconhecer que, apesar de tudo, o objetivo final é a assimilação à língua europeia. Por outras palavras, é legítimo questionar se não poderia haver um outro tipo de modelo bilingue mais continuado, para promover não só, em última instância, o inglês ou o português, mas também as línguas maternas que as populações falam, promovendo, assim, a diversidade linguística. Como provavelmente sabem, perto de metade das línguas faladas atualmente encontra-se ameaçada. É muito fácil, olhando para os espaços onde o português é a língua oficial, sinalizar línguas que vão desaparecer ou que sobrevivem com cada vez menos falantes. Basta pensar na lição de história do impacto negativo que as línguas europeias tiveram no continente americano em relação às línguas ameríndias.

*De que forma aborda a questão da variabilidade linguística entre os crioulos que compartilham uma mesma região? Na sua análise, qual é a extensão em que essa variabilidade pode ser atribuída a influências históricas, sociais e culturais?*

Esta é daquelas perguntas para as quais só há uma resposta longuíssima (*risos*). Bom, é claro que, olhando para os crioulos, só conseguimos compreender o cenário completo olhando para diferentes aspetos linguísticos, históricos, sociais, etc. das diferentes línguas que entraram em contacto. Falando mais concretamente dos crioulos do Golfo da Guiné, sabemos que se formou uma protolíngua, o protocrioulo do Golfo da Guiné, que surgiu na Ilha de São Tomé e que, depois, deu origem a quatro línguas diferentes ao longo do século XVI. Ou seja, é um pouco como a história do latim (vulgar) e das línguas românicas, mas num passado mais recente. Compreender esta protolíngua implica que vamos ter de olhar para todos os ingredientes: para o português - não o português do século XXI, mas sim o português médio e clássico dos séculos XV, XVI -, para diversas línguas africanas e, idealmente, também para a história destas línguas, o que muitas vezes é difícil, para tentar reconstituir o que levou à formação da protolíngua. Outras áreas de conhecimento, como a genética e a história, também podem ajudar a explicar as origens populacionais, mas a linguística tem, de facto, um papel primordial. Em termos de variação linguística, as línguas diferentes, que resultam da difusão de uma protolíngua no tempo e no espaço, são sempre o resultado de uma conjugação de fatores, linguísticos e não só, que também se vão manifestar linguisticamente. O protocrioulo do Golfo da Guiné, que deu origem a quatro crioulos, existiu e pode ser reconstruído. Enquanto linguistas, podemos olhar para os quatro crioulos do Golfo da Guiné, compará-los e analisar o que têm de semelhante e o que têm de diferente: o que têm de semelhante provavelmente tem a ver com uma origem partilhada, ou seja, se há um determinado som que é comum aos quatro crioulos nos mesmos contextos, provavelmente este som terá existido na protolíngua; se há uma determinada palavra que é diferente num dos quatro crioulos para referir um determinado conceito, provavelmente esta palavra não fez parte da protolíngua. Reconstituir traços gramaticais e lexicais da protolíngua a partir da variação existente entre línguas contemporâneas que sabemos que partilham uma origem comum permite-nos pensar um pouco naquilo que terá sido o antepassado comum destes crioulos.

As diferenças entre línguas que partilham a mesma origem aumentam com o seu isolamento no espaço e no tempo, devido a diferentes caminhos de evolução interna e devido

a fatores externos, como um novo contacto, por exemplo, com a língua lexificadora ou outras línguas, mas também devido a fatores sociais difíceis de prever. Os crioulos do Golfo da Guiné descendem de um tronco comum, mas são hoje línguas que não são mutuamente inteligíveis, porque cada uma foi construindo a sua própria história.

*Alguns nomes importantes dentro da crioulística não seguem o paradigma de diferenciação tipológica entre crioulos e não-crioulos, difundido sobretudo por Bickerton. Michel DeGraff, por exemplo, considerara o crioulo haitiano uma língua românica. O Professor aborda esta discussão no artigo “The Gulf of Guinea Creoles: a case-study of syntactic reconstruction”. Poderia discutir sobre a potencial visão dos crioulos de base portuguesa como pertencentes a este ramo da família indo-europeia?*

Sim, é uma questão complexa e polémica, com um historial entre os estudiosos das línguas crioulas e muitas vezes marcada por visões opostas ou diferentes. Costumo dizer nas aulas que o conceito “crioulo” não nos ajuda muito. O facto de aplicarmos o termo “crioulo” a um conjunto de línguas é, na verdade, uma criação histórica, que não corresponde, em primeiro lugar, a uma realidade sócio-histórica comum: as línguas a que chamamos crioulos têm muitas vezes histórias sociais muito diferentes. No caso do Atlântico, temos frequentemente uma história ligada à escravatura e às plantações; no caso da Ásia, uma história mais ligada ao comércio (que não de seres humanos) e à religião. Portanto, não constituem, desse ponto de vista, uma classe, mas também não constituem, na minha opinião, uma classe tipológica do ponto de vista da linguística, mas aí reside a controvérsia. Sabemos hoje o suficiente de línguas crioulas para perceber que tipologicamente são línguas, por vezes, muito distintas entre si. O forro de São Tomé e o cabo-verdiano em Cabo Verde são crioulos de base lexical portuguesa em África com léxico maioritariamente derivado do português, porém, tipologicamente, apresentam características muito distintas. Os criadores ou fundadores destas línguas não eram os portugueses, eram falantes de outras línguas que deixaram, em menor ou maior grau, marcas das suas línguas maternas.

A segunda parte da resposta tem a ver com a classificação genética. E aí eu não tenho muitas dúvidas, porque os crioulos de base lexical portuguesa, se nós aplicarmos o Método Comparativo que se utiliza tradicionalmente na linguística histórica, são claramente línguas filiadas na sua língua lexificadora, o português. Em geral, mais de 90% do léxico de um crioulo português tem origem no português. Assim, cumpre-se o primeiro passo: a existência

de cognatos, palavras que ligam estes crioulos ao português. O segundo passo é identificar correspondências de som regulares, o que é fácil de comprovar. Vemos que há regularidade na forma como as palavras do português foram integradas no léxico do cabo-verdiano e no léxico do forro. O forro caracteriza-se pela despalatalização: *sola* ‘chorar’; *sabi* ‘chave’ e *son* para ‘chão’; na Alta Guiné, em Cabo Verde e na Guiné-Bissau, a solução nos mesmos contextos é uma africada: *txora*, *txabi*, *txon*.

O Método Comparativo aplica-se às línguas, mas ignora a história: não leva em conta, por exemplo, se os crioulos foram adquiridos como L2. No caso do inglês, também há um forte historial de aquisição L2 e reestruturação profunda, mas ninguém diz que o inglês não é uma língua germânica, pois não? Seguindo os factos e aplicando o Método Comparativo que utilizamos para classificar as línguas no mundo, temos de dizer de uma forma muito clara que os crioulos de base lexical portuguesa são línguas que estão na linhagem do indo-europeu, das línguas itálicas, ou românicas, se quiserem, descendendo em última instância do português. A mesma genealogia aplica-se ao haitiano, que, neste caso, descende do francês; se for o jamaicano, descende das línguas germânicas, mais especificamente do inglês. Na minha opinião, não há dúvidas na questão da filiação. Queria também dizer que um dos argumentos na base deste debate sobre a filiação tem a ver com o facto de os crioulos terem uma história peculiar de contacto e de serem gramaticalmente bastante diferentes das suas línguas lexificadoras. Embora isso seja verdade, dentro das diferentes famílias linguísticas também encontramos línguas que são muito diferentes entre si: o islandês e o inglês são duas línguas germânicas muito diferentes e, em outras famílias, há línguas cujas diferenças são ainda muito maiores do que as diferenças entre o islandês e o inglês. Portanto, esse argumento não funciona. Para efeitos de classificação, temos que seguir o mesmo método que aplicamos às outras línguas e não cair na tentação de fazer dos crioulos um grupo diferente das outras línguas do mundo. Aliás, o próprio termo sócio-histórico “crioulo” parece ser um obstáculo à integração das línguas com este “rótulo” nas línguas naturais do mundo. Vejam que, curiosamente, temos um atlas online para “crioulos” - Atlas of Pidgin and Creole Language Structures (APiCS)<sup>3</sup> - e um atlas para as línguas do mundo - World Atlas of Language Structures (WALS)<sup>4</sup>, sem “crioulos”.

---

<sup>3</sup> <https://apics-online.info/>

<sup>4</sup> <https://wals.info/>

*Tendo o Senhor Professor construído uma longa e relevante carreira trabalhando com os crioulos de base portuguesa, quais são os conselhos que poderia dar aos futuros linguistas que também almejam seguir uma vida académica, em especial àqueles que se querem dedicar ao estudo destas línguas?*

Em primeiro lugar, eu diria, como já tinha dito num outro momento, que, se alguém tiver realmente interesse em abraçar algo diferente, uma área que é mais pioneira, em que se podem fazer trabalhos pioneiros, a área dos crioulos é excelente para iniciar a investigação. Da minha própria experiência, eu diria que o melhor que se pode fazer é trabalho de campo. No início da entrevista, falei da viagem que fiz a São Tomé no fim do milénio passado. Essa e as viagens seguintes para São Tomé foram momentos que me marcaram, porque nos sentamos com pessoas que, de forma muitas vezes voluntária, oferecem algo de si, que é a sua língua, o objeto da nossa investigação. São momentos em que criamos uma ligação com o nosso objeto de estudo. Não só com a língua como também com as pessoas. Considero importante tentar devolver algo do meu trabalho às comunidades e sobretudo pensar em formas de manter essas línguas minoritárias. Eu não quero, mas não digo isso de forma egoísta, que o meu objeto de estudo desapareça. A melhor forma de criarmos um laço forte com esta área de investigação é fazer trabalho de campo, que é sempre uma experiência única. Não é fazer aquela tese, aquele trabalho teórico em que vamos abrir todos os livros que estão na prateleira e tentar fazer algo com isso. Não. Isso também é importante, mas o trabalho de campo, a recolha de dados, o olhar para os dados e não olhar primeiro para todas as teorias e ver se conseguimos encaixar os dados, tudo isso é fundamental para fazermos algo diferente e algo que nos mantenha sempre motivados para o nosso trabalho.

Na primeira vez que fui a São Tomé, entrevistei várias pessoas, mas trabalhei mais com um senhor mais velho que era um terapeuta tradicional. Era uma pessoa que tinha um grande conhecimento das plantas e ajudava as pessoas que iam de diferentes partes de São Tomé em busca da sua ajuda: esse senhor chamava-se *Sun* (Sr.) Sabino. Entre as suas massagens e terapias, concedia-me essas entrevistas. Ele contava histórias tradicionais, intercaladas com canto. Era um profundo conhecedor das histórias tradicionais de São Tomé e Príncipe (as *soya*). Numa das sessões, disse uma frase (em forro) em relação às gravações de que nunca me esqueci: “eu posso morrer, mas a minha voz continuará”. Ou seja, deixava-se gravar de forma abnegada, em prol da sua voz que era a sua língua. Ele, entretanto, já faleceu, já tinha uma idade avançada naquela altura, mas, de facto, tal como ele disse, a sua voz

continuará. A sua voz está presente no *corpus* que recolhi, nas análises que eu faço dos dados e é, de facto, uma voz que permanece na minha investigação e que me tem ajudado a retribuir com o meu trabalho. Portanto, insisto no trabalho de campo. Ter coragem de pegar em equipamento de gravação, preparar tudo, levar questionários - ou não: eu fui um bocadinho à toa a primeira vez que fui a São Tomé - e meter as mãos na massa, conhecer uma nova realidade, porque uma língua é muito mais do que só um conjunto de palavras. É toda uma cultura. Enfim, é através do trabalho de campo que vamos descobrir mais, não só sobre a língua, mas também sobre tudo o que envolve a língua. No meu caso, fui como linguista para recolher dados, para fazer investigação linguística, mas acabei não só por fazer esse trabalho como também comecei a investigar as origens do forro e dos outros crioulos do Golfo da Guiné, colaborando inclusive com outros linguistas, geneticistas e historiadores, participei numa proposta de escrita para os crioulos de São Tomé e Príncipe e sou coautor de um pequeno dicionário do forro.

Quando criamos um laço com a língua e com a comunidade que a fala, percebemos, com o passar do tempo, que valeu a pena investir em algo que é um pouco diferente daquilo que habitualmente se faz. Incentivo todos os jovens estudantes a arriscar um bocadinho, a pensar e agir fora da caixa, a sair da sua zona de conforto (é preciso para o trabalho de campo!) e adotar, para investigação e para a vida, um crioulo de base lexical portuguesa.

# TRADUÇÃO AUTOMÁTICA *VERSUS* TRADUÇÃO HUMANA: O CASO DA HOMONÍMIA E HOMOGRAFIA PORTUGUESAS

Margarida Silva<sup>5</sup>

[up202308227@edu.lettras.up.pt](mailto:up202308227@edu.lettras.up.pt)

FACULDADE DE LETRAS DA UNIVERSIDADE DO PORTO (PORTUGAL)

Raquel Moreira<sup>6</sup>

[up201709580@edu.med.up.pt](mailto:up201709580@edu.med.up.pt)

FACULDADE DE LETRAS DA UNIVERSIDADE DO PORTO (PORTUGAL)

**RESUMO.** Um dos problemas das ferramentas de tradução automática é a ambiguidade do texto, nomeadamente, frases com palavras homónimas e/ou homógrafas. Neste trabalho, procuramos verificar em que situações ocorrem erros na tradução automática destas palavras. Num estudo qualitativo do desempenho de duas ferramentas na tradução para inglês de palavras homónimas e homógrafas portuguesas, é feito um levantamento dos erros ocorridos. Por meio da metodologia utilizada, detetaram-se 86 erros, a maioria relacionada com terminologia ou alteração de palavras. As palavras homógrafas foram as que geraram mais erros, demonstrando limitações na tradução automática deste tipo de palavras e reforçando a importância da revisão humana.

**PALAVRAS-CHAVE.** Tradução Automática, Homonímia, Homografia, Português, Inglês.

**ABSTRACT.** One of the problems with machine translation tools is text ambiguity, namely sentences using homonyms and/or homographs. In this study, we intended to verify in which situations machine translation generates errors when using these words. We developed a qualitative study of the performance of two machine translation tools when translating Portuguese homonyms and homographs into English. This methodology detected 86 errors, most of them related to terminology or word changes. Homographs were the words associated with more mistakes, revealing that there is a limitation in the automatic translation of these words, which means that human translation cannot yet be replaced by these systems.

**KEYWORDS.** Machine Translation, Homonymy, Homography, Portuguese, English.

---

<sup>5</sup> Estudante do 1.º ano do Mestrado em Tradução e Serviços Linguísticos.

<sup>6</sup> Estudante do 1.º ano do Mestrado em Tradução e Serviços Linguísticos.

## 1. Introdução

As definições de homonímia e homografia encontram-se já bem estabelecidas. A homonímia diz respeito a “palavras que têm a mesma forma gráfica e fonética mas sentidos diferentes” (Chaves 2013). Por sua vez, as palavras homógrafas referem-se a “significantes diferentes (com pronúncia diferente) mas com igual grafia [excetuados os sinais diacríticos (acentos gráficos)]” (Fernandes 2005).

Um dos principais problemas relacionados com o processamento automático da linguagem que ocorrem quando se utilizam ferramentas de tradução automática é a presença de algum tipo de ambiguidade no texto, quer seja estrutural ou lexical. A ambiguidade lexical relaciona-se com a atribuição de mais do que um significado a uma palavra e, inerentemente, com a polissemia, homonímia e vagueza lexical (Abdullah, Sarno, Purwitasari & Akhsani 2023; Gomes 2010).

As ferramentas de tradução automática têm evoluído devido à quantidade e qualidade do material linguístico que utilizam nos seus algoritmos, recorrendo, por exemplo, a dicionários eletrónicos e *corpora*. Apesar da facilidade no acesso a estas ferramentas e da possível utilidade dos seus resultados, a qualidade das traduções é ainda questionável e os seus utilizadores reconhecem-no (Barreiro, Ling, Coheur, Batista & Trancoso 2014). Todavia, estas ferramentas são alvo de constantes melhorias, particularmente com a atual oferta de serviços de inteligência artificial gratuitos que pretendem mimetizar o processamento da linguagem natural e aproximar-se da cognição humana (Borsatti, Santorum & Costa 2022). É possível distinguir diferentes tipos de tradução automática, a saber, baseada em regras, baseada em conhecimento e baseada em estatística (Leal, Maia & Maia 2014). Estas ferramentas são práticas e de utilização económica, no entanto, não parecem ser capazes de traduções completamente autónomas sem revisão ou pós-edição (Jamilah 2012).

Alguns tradutores automáticos estão disponíveis gratuitamente, apresentando como vantagem o facto de terem à sua disposição uma quantidade ilimitada de dados. A desvantagem associada prende-se com a qualidade desses mesmos dados que entram nos algoritmos, sendo variável de acordo com os pares linguísticos a trabalhar (Barreiro *et al.* 2014).

As palavras homónimas correspondem a um fenómeno semântico no qual uma palavra está associada a mais do que um significado, formando-se através de um fenómeno de

convergência fonética no qual ocorrem modificações de sons em palavras sem qualquer relação semântica aparente (Câmara 1977; Ullmann 1953).

O número excessivo deste tipo de palavras numa determinada língua dificulta a sua distinção e contribui para a ambiguidade semântica. As palavras homógrafas são, por sua vez, uma subcategoria das palavras homónimas, logo todas as características anteriormente mencionadas também se aplicam a este grupo de palavras. A diferença está na fonética e no significado, sendo que nesta subcategoria apenas a grafia se mantém igual. As homógrafas falsas são palavras cuja única diferença é gráfica e marcada por um acento (Mengjie 2020).

O comportamento e a perceção das palavras homónimas e homógrafas portuguesas em contexto de tradução automática não foi largamente estudado, sendo a literatura disponível escassa. Assim, o objetivo deste estudo consistiu em verificar em que situações ocorrem os erros na tradução automática deste tipo de palavras, quais foram os fatores que mais contribuíram para a sua ocorrência e de que forma o tradutor humano pode interpretar o erro e melhorar a eficiência e a utilidade destas ferramentas de apoio à tradução.

## **2. Metodologia**

### **2.1. Desenho de estudo e processo de amostragem**

Trata-se de um estudo qualitativo cuja amostra em análise resulta da seleção de palavras a partir de uma listagem exaustiva de todas as palavras homógrafas, homógrafas falsas e homónimas identificadas em português europeu, listadas no trabalho de Mengjie (2020).

### **2.2. Recolha de dados e variáveis em análise**

Para a análise do desempenho das duas ferramentas de tradução automática, *Google Translate* e *DeepL*, na tradução de português europeu para inglês britânico, foram formuladas de forma autónoma e independente frases que contivessem uma ou ambas as palavras identificadas nos pares de homógrafas, homógrafas falsas e homónimas. A seleção destas duas ferramentas em particular deve-se à sua amplitude de uso e à sua disponibilidade gratuita.

Todas as frases foram traduzidas, em março de 2024, com recurso a ambas as ferramentas mencionadas, seguindo-se a análise e a classificação das traduções obtidas em

três níveis: Nível I - totalmente corretas, Nível II - aceitáveis, apesar de serem possíveis alterações e melhorias, e Nível III - incorretas.

### 2.3. Metodologia de análise

Para a análise pretendida, verificaram-se quais os casos que geraram erros de tradução e/ou linguísticos nos resultados obtidos através das ferramentas de tradução automática. Foi dada preferência à primeira tradução sugerida quando a ferramenta de tradução automática *DeepL* oferecia mais do que uma opção de tradução. A utilização de duas ferramentas permitiu também a comparação dos seus respetivos desempenhos.

Foram selecionadas três frases formuladas a partir das palavras de cada uma das categorias em estudo, homógrafas, homógrafas falsas ou homónimas, incluindo uma tradução classificada como correta, uma aceitável e uma incorreta. Os critérios sintáticos utilizados para a análise dos resultados incluíram a identificação das classes de palavras nos textos de partida e de chegada. Os critérios semânticos visaram a determinação da função semântica das palavras de origem e das correspondentes nas traduções. Por fim, foram sugeridas correções e reformulações às traduções obtidas, consideradas como aceitáveis ou incorretas (Níveis II e III, respetivamente).

## 3. Resultados

### 3.1. Amostra em estudo

Um total de 109 palavras foi selecionado para tradução nas ferramentas em estudo. Destas, 92 eram homógrafas (incluindo 33 homógrafas falsas) e 17 eram homónimas.

Obtiveram-se erros em 76 palavras ao longo de toda a amostra (69,7%), distribuídos da seguinte forma: 17 erros de Nível II e 36 de Nível III no grupo das palavras homógrafas, (n=53, 89,8% do grupo, 48,6% do total amostral), 9 erros de Nível II e 4 de Nível III no grupo das homógrafas falsas (n=13, 39,4% do grupo, 11,9% do total amostral) e 1 erro de Nível II e 9 de Nível III no grupo das palavras homónimas (n=10, 58,8% do grupo, 9,2% do total amostral).

Em anexo, encontram-se todas as palavras analisadas, as frases produzidas e as respetivas traduções nas duas ferramentas. Sempre que se atribui uma função semântica a um verbo, pretende-se que a interpretação seja “*O Sujeito desempenha a função de*”.

### 3.2. Análise de palavras homógrafas

Correta: *Acordo*

Origem	Texto	Classe de palavras	Funções semânticas
Frase original	Certo dia <b>acordo</b> (1) e vejo que a ONU chegou, finalmente, a um <b>acordo</b> (2).		(1) Experienciador
<i>Google Translate</i>	One day <b>I wake up</b> (1) and see that the UN has finally reached an <b>agreement</b> (2).	(1) Verbo → Verbo	→ Experienciador
<i>DeepL</i>	One day <b>I wake up</b> (1) and see that the UN has finally reached an <b>agreement</b> (2).	(2) Nome → Nome	(2) Tema → Tema

Aceitável: *Aceno*

Origem	Texto	Classe de palavras	Funções semânticas
Frase original	Um <b>aceno</b> (1) é um sinal de reconhecimento, por isso <b>eu aceno</b> (2) ao João sempre que o vejo.	(1) Nome → Nome	(1) Estímulo → Estímulo →
<i>Google Translate</i>	A <b>wave</b> (1) is a sign of recognition, which is why <b>I wave</b> (2) to João whenever I see him.	→ Sugestão: <i>nod</i> (Nome)	Sugestão: <i>nod</i> (Estímulo)
<i>DeepL</i>	A <b>wave</b> (1) is a sign of recognition, which is why <b>I wave</b> (2) to João every time I see him.	(2) Verbo <i>Acenar</i> → Verbo ( <i>To wave</i> )	(2) Agente → Agente

Incorreta: Coro

Origem	Texto	Classe de palavras	Funções semânticas
Frase original	Sempre que tenho apresentações do <b>coro</b> (1), os meus colegas dizem-me que <b>coro</b> (2).	(1) Nome → Nome	(1) Tema → Tema
Google Translate	Whenever I have <b>choir</b> (1) performances, my colleagues tell me <b>what choir</b> (2).	(2) Verbo <i>Corar</i> → Nome → Sugestão:	(2) Experienciador → Tema →
DeepL	Whenever I have <b>choir</b> (1) performances, my colleagues tell me <b>which choir</b> (2).	<i>I blush</i> (Verbo <i>To blush</i> )	Sugestão: <i>I blush</i> (Experienciador)

Incorreta (DeepL): Aperto

Origem	Texto	Classe de palavras	Funções semânticas
Frase original	Sinto que <b>aperto</b> (1) demasiado a tua mão. Estamos num <b>aperto</b> (2) financeiro desde junho.	(1) Verbo <i>Apertar</i> → Verbo <i>To squeeze</i>	(1) Agente → Agente
Google Translate	I feel like I'm <b>squeezing</b> (1) your hand too much. We have been in a financial <b>crunch</b> (2) since June.	(2) Nome <i>Aperto</i> → Nome <i>Squeeze</i>	(2) Paciente → Paciente →
DeepL	I feel like I'm <b>squeezing</b> (1) your hand too much. We've been in a financial <b>squeeze</b> (2) since June.	→ Sugestão: <i>crunch</i> (Nome)	Sugestão: <i>crunch</i> (Paciente)

### 3.3. Análise de palavras homógrafas falsas

Correta: *Alívio / Alívio*

Origem	Texto	Classe de palavras	Funções semânticas
Frase original	Vou tomar a medicação para ver se sinto algum <b>alívio</b> (1) neste joelho e <b>alívio</b> (2) a mente também.	(1) Nome → Nome	(1) Estímulo → Estímulo
<i>Google Translate</i>	I'm going to take the medication to see if I feel any <b>relief</b> (1) in this knee and <b>ease</b> (2) my mind as well.	(2) Verbo <i>Aliviar</i> → Verbo <i>To ease</i>	(2) Experienciador → Experienciador
<i>DeepL</i>	I'm going to take the medication to see if I can get some <b>relief</b> (1) from this knee and <b>ease</b> (2) my mind too.		

Aceitável: *Médico / Medico*

Origem	Texto	Classe de palavras	Funções semânticas
Frase original	Nunca me <b>medico</b> (1) sem consultar um <b>médico</b> (2) antes.	(1) Verbo <i>Medicar</i> → Perífrase verbal <i>To take medication</i> → Sugestão: <i>self-medicate</i> (Verbo <i>To self-medicate</i> )	(1) Paciente afetado → Tema → Sugestão: <i>self-medicate</i> (Paciente afetado)
<i>Google Translate</i>	I never <b>take medication</b> (1) without consulting a <b>doctor</b> (2) first.		
<i>DeepL</i>	I never <b>take medication</b> (1) without consulting a <b>doctor</b> (2) first.	(2) Nome → Nome	(2) Origem → Origem

Incorreta: Relês / Reles

Origem	Texto	Classe de palavras	Funções semânticas
Frase original	Este livro é <b>reles</b> (1), não entendo por que motivo o <b>relês</b> (2).	(1) Adjetivo → Nome → Sugestão: <i>cheap</i> (Adjetivo)	
Google Translate	This book is <b>rubbish</b> (1), I don't understand why <b>I read</b> (2) it again.	(2) Verbo <i>Reler</i> (2. <sup>a</sup> pessoa do singular, Presente do Modo Indicativo)	(1) Tema → Tema → Sugestão: Tema
DeepL	This book is <b>rubbish</b> (1), I don't understand why <b>you're re-reading</b> (2) it.	→ Verbo <i>To read</i> (1. <sup>a</sup> pessoa do Singular, <i>Present Simple</i> ) / Verbo <i>To Re-read</i> (2. <sup>a</sup> pessoa do singular, <i>Present Continuous</i> ) → Sugestão: <i>you are re-reading it</i> (2. <sup>a</sup> pessoa do singular, <i>Present Continuous</i> )	(2) Experienciador → Experienciador / Experienciador → Sugestão: Experienciador

Incorreta (DeepL): Anúncio / Anuncio

Origem	Texto	Classe de palavras	Funções semânticas
Frase original	Ao ver o <b>anúncio</b> (1) na Internet, tomei a minha decisão: <b>anuncio</b> (2) a venda do meu carro ainda hoje.	(1) Nome → Nome	(1) Estímulo → Estímulo
Google Translate	When I saw the <b>ad</b> (1) on the Internet, I made my decision: I <b>will announce</b> (2) the sale of my car today.	(2) Verbo <i>Anunciar</i> → Verbo <i>To sell</i> → Sugestão: <i>to announce the sale of</i> (Verbo)	(2) Agente → Agente → Sugestão: Agente
DeepL	When I saw the <b>advert</b> (1) on the Internet, I made up my mind: I'm going <b>to sell</b> (2) my car today.		

### 3.4. Análise de palavras homónimas

Correta: São

Origem	Texto	Classe de palavras	Funções semânticas
Frase original	As crianças estão a <b>são</b> (1) e salvo, <b>são</b> (2) muito resistentes.	(1) Nome → Nome	(1) Experienciador → Experienciador
Google Translate	The children are safe and <b>sound</b> (2), they <b>are</b> (2) very resilient.	(2) Verbo <i>Ser</i> →	(2) Experienciador
DeepL	The children are safe and <b>sound</b> (2), they <b>are</b> (2) very resilient.	Verbo <i>To be</i>	→ Experienciador

Aceitável: Morro

Origem	Texto	Classe de palavras	Funções semânticas
Frase original	<b>Morro</b> (1) de cansaço sempre que subo aquele <b>morro</b> (2).	(1) Verbo <i>Morrer</i> →	(1) Paciente afetado → Paciente afetado /
Google Translate	I <b>die</b> (1) of exhaustion every time I climb that <b>hill</b> (2).	Verbo <i>To die</i> / Colocação <i>To get tired</i> → Sugestão: <i>die</i> (Verbo <i>To die</i> )	Paciente afetado → Sugestão: <i>die</i> (Paciente afetado)
DeepL	I <b>get tired</b> (1) every time I go up that <b>hill</b> (2).	(2) Nome → Nome	(2) Destino → Destino

Incorreta: Cedo

Origem	Texto	Classe de palavras	Funções semânticas
Frase original	O meu filho tem de se levantar <b>cedo</b> (1) todos os dias, nisso não <b>cedo</b> (2).	(1) Advérbio → Advérbio	(1) Estímulo → Estímulo
Google Translate	My son has to get up <b>early</b> (1) every day, <b>that's not early</b> (2).	(2) Verbo <i>Ceder</i> → Construção com advérbio	(2) Agente → Tema
DeepL	My son has to get up <b>early</b> (1) every day, <b>not early at that</b> (2).	→ Sugestão: <i>I don't give in on that (Phrasal verb) / I don't compromise on that</i>	→ Sugestão: <i>give in / compromise (Agente)</i>

Incorreta: (Dar um) Trago

Origem	Texto	Classe de palavras	Funções semânticas
Frase original	<b>Trago</b> (1) sempre comigo uma garrafa de água, para que possa <b>dar um trago</b> (2) caso precise.	(1) Verbo → Verbo	(1) Agente → Agente
Google Translate	I always <b>carry</b> (1) a bottle of water with me, so I can <b>have a drink</b> (2) if I need to.	(2) Construção com verbo leve → Construção com verbo <i>To have</i> e Nome	(2) Agente → Agente
DeepL	I always <b>carry</b> (1) a bottle of water with me so that I can <b>take a drink</b> (2) if I need to.	<i>Drink</i> → Sugestão: Construção com verbo <i>To take</i> e Nome <i>Sip</i>	Agente → Sugestão: <i>take a sip (Agente)</i>

### 3.5. Achados linguísticos

A Tabela 1 sumaria os achados linguísticos geradores de erro. A contabilização de erros diz respeito a ocorrências, sendo considerados para contagem múltiplos tipos de erros que possam ocorrer na mesma palavra ou frase.

Tabela 1 – Achados linguísticos das traduções geradas automaticamente.

<b>Achados linguísticos</b>	<b>Total de casos</b>	<b>Google Translate (n.º de casos)</b>	<b>DeepL (n.º de casos)</b>
Problemas de terminologia	30	16	14
Substituição de verbos por perífrases verbais / construções com verbos auxiliares	7	2	5
Alteração da pessoa da conjugação verbal	3	1	2
Omissão de determinantes	2	2	0
Pluralização de elementos frásicos	2	1	1
Dificuldade na seleção do tempo verbal	4	2	2
Alteração de palavras e/ou classes de palavras	31	17	14
Adição ou omissão de palavras	5	1	4
Dificuldade com colocações	2	1	1
<b>Total, n (%)</b>	<b>86 (100%)</b>	<b>43 (50%)</b>	<b>43 (50%)</b>

#### 4. Discussão

Tal como expectável, as duas ferramentas de tradução automática utilizadas para a realização deste estudo apresentaram limitações no seu desempenho perante a integração de palavras homónimas e homógrafas. Foram detetados, no total, 86 erros, distribuídos equitativamente (em termos percentuais e absolutos) pelas duas ferramentas. A maioria dos erros relacionou-se com questões pequenas de terminologia inadequada ou alteração de palavras com repercussões no sentido das frases. As palavras homógrafas foram as que, em proporção, geraram uma maior ocorrência de erros: aproximadamente, 90% das palavras deste grupo.

Nas traduções corretas (Nível I), as classes de palavras permaneceram as mesmas, bem como as funções semânticas. Nas aceitáveis (Nível II), apesar da compreensão do sentido das frases, tanto a classe de palavras como a função semântica, por vezes, se alteraram na frase de chegada, tal como nas traduções incorretas (Nível III). Este erro, algumas vezes, deve-se à elipse do sujeito antes do verbo, sendo esta uma construção não reconhecida pelo tradutor automático. Quando identificado o sujeito, a palavra podia alterar-se, mas continuava a perder o significado original. Assim, verifica-se a falta de capacidade do tradutor automático para traduzir palavras com a mesma grafia e fonia, sendo que, quando a grafia é alterada por um acento gráfico, constatou-se uma melhor capacidade de processamento da informação. Os exemplos em análise que originaram erros de Nível II demonstraram sempre falta de concordância de significado entre os textos de partida e de chegada.

A tradução automática tem apresentado um crescimento significativo de qualidade; no entanto, os seus algoritmos usam informação sobre os textos de chegada descurando a sua precisão face ao texto de partida (Gonçalves, Buchicchio & Moniz 2022). No que diz respeito à desambiguação lexical, quando em comparação com os tradutores humanos, há autores que afirmam que a componente humana e a consciência linguística associada melhoram os resultados, particularmente no caso de palavras homógrafas, as quais já demonstraram exigir por parte do tradutor automático um processamento superior ao que é necessário para a tradução de outras palavras, uma vez que exige a seleção de significados, podendo ou não existir equivalentes terminológicos entre línguas (Gadelha & Toassi 2021; Karnal 2019). No estudo de Avelino, Macketanz, Avramidis & Möller (2022), os autores demonstraram que a ambiguidade continua a ser um dos maiores problemas para estas ferramentas, a par da seleção de terminologia adequada e da escolha dos tempos verbais corretos.

Em 2016, a *Google* substituiu o seu sistema estatístico por um sistema neural, através do qual as traduções são processadas por sistemas baseados em palavras e não em frases. Estas traduções, em teoria, tornaram-se mais precisas não só em termos de terminologia, mas também em termos frásicos, uma vez que se tornou mais fácil construir um contexto depois de determinados os significados individuais mais adequados (Borsatti *et al.* 2022). No entanto, o tipo de erros que parece continuar a ocorrer e que se destaca inclui, indubitavelmente, a polissemia e a homonímia. Diferentes ferramentas lidam com este tipo de partículas linguísticas de formas distintas, destacando-se o papel do tradutor humano, o qual obtém de forma destacada os melhores resultados (Abdulaal 2022).

Outros autores já compararam o desempenho do *Google Translate* com o do *DeepL*, relatando uma pior experiência com o segundo, ao contrário daquela que é a perceção empírica da maioria dos tradutores profissionais, trabalhando com o par linguístico português-ínglês, que recorrem a esta ferramenta. Todavia, há estudos prévios que reconhecem dificuldades terminológicas em ambas as plataformas (Borges & Pimentel 2020). Neste estudo, o desempenho das ferramentas foi semelhante em termos percentuais; contudo, nas categorias mais afetadas por erros (terminologia e alteração de palavras), o *Google Translate* apresentou mais ocorrências. A melhor forma de avaliar a qualidade destas traduções consiste, invariavelmente, em recorrer a um tradutor humano. Não obstante, este método é dispendioso, pelo que se devem procurar processos mais automatizados para lidar com os resultados da tradução automática (Avelino *et al.* 2022; Barreiro & Ranchhod 2005).

## 5. Limitações

Este estudo apresenta algumas limitações, nomeadamente: (i) a metodologia adotada para obtenção do *corpus* em análise, que limita a sua aplicação a amostras de maiores dimensões; (ii) a formulação das frases, que talvez pudesse ser melhorada e tornada mais neutra para a análise; e (iii) o número de palavras e a variedade de contextos de utilização das mesmas, que poderiam ser alargados. Assim, pelo facto de o *corpus* não ser autêntico, a validação destes resultados poderá ser limitada aos contextos de utilização criados através da presente metodologia. Apesar disso, consideramos que a opção por este método de obtenção de *corpus* foi a mais adequada para este estudo, dada a especificidade e a escassa ocorrência de frases com ambas as formas deste tipo de palavras (isto é, homónimas e homógrafas) em materiais publicados disponíveis para constituição de um *corpus* autêntico. Partindo da premissa de que as ferramentas de tradução automática se revelariam menos eficazes quando este tipo de palavras ocorresse em contiguidade, optou-se por criar situações com pouco contexto e de interpretação mais complexa, de forma a avaliar o desempenho destas ferramentas em situações de uso extremas e, conseqüentemente, pouco frequentes.

Considera-se que mais estudos devem ser realizados neste âmbito, uma vez que a evidência disponível acerca do tema é ainda escassa. A evolução das ferramentas de tradução automática é contínua e encontra-se em crescimento exponencial, pelo que o desenvolvimento de estudos que abordem este tópico poderá contribuir para a melhoria do seu desempenho, acelerando o processo atual assente na crescente utilização destas ferramentas associadas a algoritmos.

## 6. Conclusão

As ferramentas de tradução automática apresentam ainda algumas limitações. Por vezes, o seu desempenho é adequado e fornece sugestões de tradução úteis; no entanto, estas podem facilmente tornar-se demasiado literais, incompletas ou desadequadas ao estilo do texto de partida, pelo que a componente humana na tradução não poderá, por ora, ser suplantada pela máquina, sobretudo no que diz respeito à utilização de palavras homónimas e homógrafas.

A percentagem de ocorrência de erros na tradução deste tipo de palavras de português para inglês foi significativa, apesar de as ferramentas utilizadas apresentarem um melhor desempenho com palavras homónimas do que com palavras homógrafas. Várias traduções

foram consideradas aceitáveis, conquanto não fossem as que apresentavam melhor fluência, o que demonstra que estas ferramentas ainda não estão suficientemente desenvolvidas para compreenderem diferenças contextuais quando este tipo de palavras é utilizado.

## Referências

- Abdulaal, M. A. A. D. 2022. Tracing machine and human translation errors in some literary texts with some implications for EFL translators. *Journal of Language and Linguistic Studies* 18(1): 176-191.
- Abdullah, R.; Sarno, R.; Purwitasari, D.; Akhsani, A. I. 2023. Homonym and Polysemy Approaches in Term Weighting for Indonesian-English Machine Translation. 14th International Conference on Information and Communication Technology and System, ICTS 2023. Surabaya: IEEE, 232-237.
- Avelino, M.; Macketanz, V.; Avramidis, E.; Möller, S. 2022. A Test Suite for the Evaluation of Portuguese-English Machine Translation. In *Computational Processing of the Portuguese Language. PROPOR 2022. Lecture Notes in Computer Science(13208)*. Springer, 15-25.
- Barreiro, A.; Ling, W.; Coheur, L.; Batista, F.; Trancoso, I. 2014. Projetos sobre tradução automática do português no laboratório de sistemas de língua falada do INESC-ID. *Linguamática* 6(2): 75-85.
- Barreiro, A.; Ranchhod, E. 2005. Machine translation challenges for Portuguese. *Linguisticae Investigationes* 28(1): 3-18.
- Borges, T. M. e.; Pimentel, J. M. M. 2020. Avaliação humana da tradução automática de combinações lexicais especializadas: o caso do Google Translate e do DeepL. *Belas Infiéis* 9(4): 21-43.
- Borsatti, D. A.; Santorum, K. A. T.; Costa, A. R. 2022. Tradução automática de palavras polissêmicas: tecnologias atuais à luz da Linguística Cognitiva. *Entrepalavras* 11(3): 52-74.
- Câmara, J. M. 1977. *Dicionário de lingüística e gramática, referente à língua portuguesa*. Petrópolis: Editora Vozes.
- Chaves, R. P. 2013. Organização do Léxico. In *Gramática do português (Vol. I, 190-192)*. Lisboa: Fundação Calouste Gulbenkian.
- Fernandes, A. G. 2005. Gramática da língua portuguesa: uma nova abordagem (86-89). Edições Caixotim.
- Gadelha, L. M. d. S.; Toassi, P. F. P. 2021. Acesso lexical de palavras homógrafas interlinguísticas português brasileiro: Inglês em uma tarefa de decisão linguística e de tradução. *Revista Linguagem em Foco* 13(4): 56-77.
- Gomes, A. L. M. 2010. Tradução Automática e Linguagens Controladas: Contributos para um Português Controlado. Tese de mestrado, Universidade de Lisboa.
- Gonçalves, M.; Buchicchio, M.; Moniz, H. 2022. Impacto da qualidade de textos de partida criados por utilizadores e agentes e a propagação de erros em sistemas de Tradução Automática. *Revista da Associação Portuguesa de Linguística* (9): 133-149.

- Jamilah, S. N. 2012. Semantic translation errors as a result of Google Translate (a case of errors of translation on homonymous and polysemous words in Bahasa Indonesia). *Diglossia: Jurnal Kajian Ilmiah Kebahasaan dan Kesusastraan* 4(1).
- Karnal, A. R. 2019. Um Estudo Comparativo da Desambiguação Lexical Realizada pelo Tradutor Eletrônico e Tradutores Humanos. *Cadernos de Tradução (Número Especial)*: 24-34.
- Leal, M. A.; Maia, A.; Maia, T. 2014. Um Estudo Comparativo de Três Técnicas de Tradução Automática para Processamento de Textos. Disponível na Internet em: <https://www.researchgate.net/publication/265374753>
- Mengjie, Z. 2020. O Domínio das Palavras Homógrafas e Parónimas: um estudo com alunos chineses. Dissertação de Mestrado, Universidade de Aveiro.
- Ullmann, S. 1953. Descriptive semantics and linguistic typology. *Word* 9(3): 225-240.

ANEXO – LISTA DAS PALAVRAS E FRASES ANALISADAS

Lista de palavras homógrafas, das frases produzidas e respetivas traduções (Nível I - Correta; Nível II - Aceitável; Nível III - Incorreta):

Palavra	Frase	Tradução <i>Google Translate</i>	Nível	Tradução <i>DeepL</i>	Nível
Aceno	Um aceno é um sinal de reconhecimento, por isso eu aceno ao João sempre que o vejo.	A wave is a sign of recognition, which is why I wave to João whenever I see him. <b>(Sugestão: A wave → A nod)</b>	II	A wave is a sign of recognition, which is why I wave to João every time I see him. <b>(Sugestão: A wave → A nod)</b>	II
Acerto	Houve um acerto na correção do teste, o que é uma pena porque eu acerto sempre as respostas.	There was a mistake in correcting the test, which is a shame because I always get the answers right. <b>(Sugestão: mistake in → rearrangement when)</b>	II	There was a hit in the correction of the test, which is a shame because I always get the answers right. <b>(Sugestão: hit → rearrangement)</b>	III
Acordo	Certo dia acordo e vejo que a ONU chegou, finalmente, a um acordo.	One day I wake up and see that the UN has finally reached an agreement.	I	One day I wake up and see that the UN has finally reached an agreement.	I
Almoço	Sempre que almoço fico feliz e este almoço está muito bom.	Whenever I have lunch I feel happy and this lunch is very good.	I	Every time I have lunch I'm happy and this lunch is very good.	I
Apelo	O João fez um apelo aos amigos: - Apelo a todos que me ouçam.	João made an appeal to his friends: - I appeal to everyone to listen to me.	I	João appealed to his friends: - I appeal to everyone to listen to me.	I
Apoio	A minha avó necessita de um apoio para subir as escadas. Eu apoio a minha equipa de futebol.	My grandmother needs support to climb the stairs. I support my football team. <b>(Sugestão: support → a support/an aid)</b>	II	My grandmother needs a support to climb the stairs. I support my football team.	I
Aterro	Detesto a vista do avião sempre que aterro neste	I hate the view of the plane every time I land at this airport	I	I hate the view from the aeroplane every time I land at this	I

Palavra	Frase	Tradução Google Translate	Nível	Tradução DeepL	Nível
	aeroporto tão perto do aterro do lixo da cidade.	so close to the city's garbage dump.		airport so close to the city's rubbish dump.	
Aperto	Sinto que aperto demasiado a tua mão. Estamos num aperto financeiro desde junho.	I feel like I'm squeezing your hand too much. We have been in a financial crunch since June.	I	I feel like I'm squeezing your hand too much. We've been in a financial squeeze since June. <b>(Sugestão: squeeze → crunch)</b>	III
Besta	O camelo é uma besta de carga. Aquela besta é ótima para caçar.	The camel is a beast of burden. That beast is great for hunting. <b>(Sugestão: That beast → That crossbow)</b>	III	The camel is a beast of burden. That beast is great for hunting. <b>(Sugestão: That beast → That crossbow)</b>	III
Borra	Odeio ter de limpar a borra de café. Esta caneta borra tudo.	I hate having to clean up coffee grounds. This pen smudges everything.	I	I hate having to clean up coffee grounds. This pen smears everything.	I
Choro	Já ninguém pode ouvir o meu choro porque todas as semanas choro por uma razão diferente.	No one can hear my crying anymore because every week I cry for a different reason. <b>(Sugestão: me crying → my cry)</b>	II	Nobody can hear me cry anymore because I cry for a different reason every week. <b>(Sugestão: me cry → my cry)</b>	II
Colher	A minha avó foi colher alfaces do quintal e esqueceu-se de me passar a colher que pedi.	My grandmother went to collect lettuce from the yard and forgot to give me the spoon I asked for.	I	My grandmother went to pick lettuce from the backyard and forgot to give me the spoon I asked for.	I
Começo	Quando começo a dançar toda a gente fica a olhar para mim. Adoro quando tenho a oportunidade de um novo começo.	When I start dancing everyone looks at me. I love it when I have the opportunity for a new beginning.	I	When I start dancing, everyone stares at me. I love it when I have the opportunity for a new beginning.	I
Cor	Essa cor de cabelo fica muito bem. Sei de cor as ruas da minha cidade natal.	This hair color looks really good. I know the streets of my hometown by heart.	I	That hair colour looks great. I know the streets of my hometown by heart.	I

Palavra	Frase	Tradução Google Translate	Nível	Tradução DeepL	Nível
Coro	Sempre que tenho apresentações do coro, os meus colegas dizem-me que coro.	Whenever I have choir performances, my colleagues tell me what choir. <b>(Sugestão: what choir → I blush)</b>	III	Whenever I have choir performances, my colleagues tell me which choir. <b>(Sugestão: which choir → I blush)</b>	III
Corte	Antigamente, os cavaleiros faziam a corte às donzelas, principalmente quando gostavam do seu corte de cabelo.	In the past, knights courted maidens, especially when they liked their haircut.	I	In the past, knights used to court maidens, especially when they liked their haircut.	I
Desse	Tirei o livro desse monte. Se me desse a oportunidade, agradecia.	I took the book out of that pile. If you gave me the opportunity, I would appreciate it.	I	I took the book out of that pile. If you'd give me the chance, I'd appreciate it.	I
Desses	Ele é um desses. Voltava se me desses uma oportunidade.	He is one of those. I would come back if you gave me a chance.	I	He's one of those. I'd go back if you gave me the chance.	I
Deste	O projeto é deste género. Já me deste a almofada.	The project is like this. You already gave me the pillow. <b>(Sugestão: like this → of this kind)</b>	II	The project is like this. You've already given me the cushion. <b>(Sugestão: like this → of this kind)</b>	II
Destes	Escolha um destes quadros. Ontem não me destes a vossa lista.	Choose one of these frames. Yesterday you didn't give me your list.	I	Choose one of these pictures. You didn't give me your list yesterday.	I
Encosto	Estraguei o encosto da cama. Normalmente, encosto a vassoura à parede.	I ruined the back of the bed. Normally, I lean the broom against the wall. <b>(Sugestão: back of the bed → headboard)</b>	III	I've broken the back of the bed. I usually put the broom against the wall. <b>(Sugestão: back of the bed → headboard)</b>	III
Endereço	Endereço esta carta ao teu irmão, mas tens que me lembrar o seu endereço.	I address this letter to your brother, but you have to remind me of his address.	I	I'm addressing this letter to your brother, but you'll have to remind me of his address.	I
Enterro	Sinto que enterro o passado sempre que vou a um enterro.	I feel like I bury the past every time I go to a funeral.	I	I feel like I'm burying the past every time I go to a funeral.	I

Palavra	Frase	Tradução Google Translate	Nível	Tradução DeepL	Nível
Este	Ele virou a este. Este livro é meu.	He turned east. This book is mine.	I	He turned to the east. This is my book.	I
Força	É preciso ter força de vontade para continuar. Ele força sempre um sorriso.	You need willpower to continue. He always forces a smile.	I	You have to have the willpower to keep going. He always forces a smile.	I
Forma	Adoro a nova forma de bolachas que comprei, mas se continuar a comê-las não vou continuar em forma.	I love the new cookie cutter I bought, but if I keep eating them I won't stay in shape.	III	I love the new biscuit tin I bought, but if I keep eating them I won't stay in shape. <b>(Sugestão: tin → cutter)</b>	III
Gelo	O gelo está no congelador e eu gelo sempre que abro a porta para o ir buscar.	The ice is in the freezer and I ice it every time I open the door to get it. <b>(Sugestão: ice it → freeze)</b>	III	The ice is in the freezer and I freeze every time I open the door to get it.	I
Gosto	Esse arroz tem o gosto da minha infância e eu gosto mesmo destes almoços de família.	This rice tastes like my childhood and I really like these family lunches.	I	This rice tastes like I was growing up and I really enjoy these family lunches. <b>(Sugestão: I was growing up → my childhood / when I was growing up)</b>	III
Governo	O governo dissolveu-se este ano, o primeiro-ministro disse: - Eu governo mal.	The government dissolved this year, the prime minister said: - I govern badly.	I	The government dissolved itself this year, the Prime Minister said: - I govern badly.	I
Gozo	Dá-me mesmo gozo quando faço partidas, principalmente quando gozo contigo e tu ris-te.	It really makes me happy when I play pranks, especially when I make fun of you and you laugh. <b>(Sugestão: make fun of → mock)</b>	II	I really enjoy it when I play pranks, especially when I make fun of you and you laugh. <b>(Sugestão: make fun of → mock)</b>	II
Interesse	É do meu interesse estudar, espero que a Maria se interesse também.	It is in my interest to study, I hope Maria is interested too.	I	I'm interested in studying, and I hope Maria is too. <b>(Sugestão: is too → is interested too)</b>	II

Palavra	Frase	Tradução Google Translate	Nível	Tradução DeepL	Nível
Jogo	Quando chego a casa jogo sempre o mesmo jogo.	When I get home I always play the same game.	I	When I get home I always play the same game.	I
Leste	Quando chegaste à ponta leste do jardim, leste aquele livro que te dei.	When you reached the eastern end of the garden, you read that book I gave you.	I	When you got to the east end of the garden, you read that book I gave you.	I
Lobo	Aquele lobo é lindíssimo, mas feriu o lobo da orelha do seu tratador.	That wolf is beautiful, but he injured his handler's earlobe.	I	That wolf is beautiful, but he hurt his keeper's ear. <b>(Sugestão: ear → earlobe)</b>	III
Meta	Ele chegou à meta. Não meta a roupa em cima da cama.	He reached the goal. Don't put your clothes on the bed. <b>(Sugestão: goal → finish line)</b>	II	He's reached the finish line. Don't put your clothes on the bed.	I
Molho	Usei um molho de salsa para fazer este molho.	I used a salsa sauce to make this sauce. <b>(Sugestão: salsa sauce → bunch of parsley)</b>	III	I used a bunch of parsley to make this sauce.	I
Olho	Quando olho para o lado dói-me o olho direito.	When I look to the side my right eye hurts.	I	When I look away, my right eye hurts.	I
Pegada	Esta página está pegada. A pegada ecológica é muito importante.	This page is broken. The ecological footprint is very important. <b>(Sugestão: broken → stuck)</b>	III	This page has a footprint. The ecological footprint is very important. <b>(Sugestão: has a footprint → is stuck)</b>	III
Peso	Ceguei ao meu peso de sonho, agora peso 55 kg.	I reached my dream weight, now I weigh 55 kg.	I	I've reached my dream weight, I now weigh 55kg.	I
Pregar	Sempre que é para pregar um prego em casa, o meu pai está pronto para pregar um sermão.	Whenever there is a need to preach a nail at home, my father is ready to preach a sermon. <b>(Sugestão: preach → nail down)</b>	III	Whenever there's a nail to be nailed down at home, my father is ready to preach a sermon.	I

Palavra	Frase	Tradução Google Translate	Nível	Tradução DeepL	Nível
Porto	Cheguei finalmente ao porto do Porto.	I finally arrived at the port of Porto.	I	I finally arrived at the harbour in Porto.	I
Remo	Sempre que remo perco um remo do barco.	Every time I row I lose an oar from the boat.	I	Every time I row, I lose an oar from the boat.	I
Rego	Tenho umas plantas num rego que nunca rego.	I have some plants in a ditch that I never water.	I	I have some plants in a watering hole that I never water. <b>(Sugestão: watering hole → ditch)</b>	II
Rola	A rola é um tipo de pássaro. Ele rola os pedaços de madeira para a lenha.	The turtledove is a type of bird. He rolls the pieces of wood into firewood.	I	The turtledove is a type of bird. It rolls pieces of wood for firewood. <b>(Sugestão: It → He)</b>	II
Rota	Não posso usar uma saia rota para a rota que planeamos.	I can't use a broken skirt for the route we have planned. <b>(Sugestão: broken → ripped)</b>	III	I can't wear a skirt for the route we're planning. <b>(Sugestão: a skirt → a ripped skirt)</b>	III
Seca	Acho uma seca que o país esteja a passar por uma estação seca.	I think it's a drought that the country is going through a dry season. <b>(Sugestão: it's a drought → it sucks / it's dreadful)</b>	III	I think it sucks that the country is going through a dry season.	I
Sede	Nunca há água na sede, fico sempre com sede.	There is never water when thirsty, I am always thirsty. <b>(Sugestão: when thirsty → at headquarters)</b>	III	There's never any water at headquarters, I'm always thirsty.	I
Segredo	Quero contar-te um segredo. Nunca segredo, falo sempre alto.	I want to tell you a secret. I never secret, I always speak loudly. <b>(Sugestão: secret → whisper)</b>	III	I want to tell you a secret. I never secrete, I always speak up. <b>(Sugestão: secrete → whisper)</b>	III
Sobre	Espero que não sobre comida, se	I hope not about food, otherwise my	III	I hope it's not about food, otherwise my	III

Palavra	Frase	Tradução Google Translate	Nível	Tradução DeepL	Nível
	não a minha mãe não se vai calar sobre isso.	mother won't shut up about it. (Sugestão: <b>not about food</b> → <b>there are no leftovers</b> )		mum won't shut up about it. (Sugestão: <b>it's not about food</b> → <b>there are no leftovers</b> )	
Solto	Eu solto o cabo quando sentir que já está solto o suficiente.	I release the cable when I feel it is loose enough.	I	I release the cable when I feel it's loose enough.	I
Sufoco	Ainda sufoco de verdade com este sufoco financeiro.	I still really struggle with this financial hardship. (Sugestão: <b>I still really struggle</b> → <b>I might suffocate</b> )	III	I'm still really struggling with this financial situation. (Sugestão: <b>I'm</b> → <b>I</b> ; <b>still really struggling</b> → <b>might suffocate</b> )	II
Termos	Já passaram os termos para entrega dos trabalhos. Tenho 2 garrafas-termos.	The terms for delivery of the work have already been passed. I have 2 thermos bottles. (Sugestão: <b>terms</b> → <b>due dates/deadlines</b> ; <b>already been</b> → <b>∅</b> ; <b>thermos bottles</b> → <b>thermos flasks</b> )	II	The deadline for submitting the work has passed. I have 2 bottle-terms. (Sugestão: <b>bottle-terms</b> → <b>thermos flasks</b> )	III
Testo	Sempre que testo novas panelas perco o testo.	Whenever I test new pans I lose the test. (Sugestão: <b>the test</b> → <b>the lid</b> )	III	Whenever I test new pans, I lose it. (Sugestão: <b>lose it</b> → <b>lose the lid</b> )	III
Transtorno	Desculpa, mas a tua visita hoje vai causar-me algum transtorno, até me transtorno a dizer-te isto.	I'm sorry, but your visit today is going to cause me some inconvenience, it even inconveniences me to tell you this. (Sugestão: <b>inconveniences</b> → <b>upsets</b> )	III	I'm sorry, but your visit today is going to cause me some inconvenience, I'm even inconveniencing myself by telling you this. (Sugestão: <b>inconveniencing</b> → <b>embarrassed with</b> )	III

Palavra	Frase	Tradução Google Translate	Nível	Tradução DeepL	Nível
Torre	O fumo que vês naquela torre é de um produtor de café magnífico, só espero que a fábrica não torra demasiado o café.	The smoke you see in that tower is from a magnificent coffee producer, I just hope the factory doesn't roast the coffee too much.	I	The smoke you see in that tower is from a magnificent coffee producer, I just hope the factory doesn't roast the coffee too much.	I
Tola	Que tola que ela é, sempre de cabeça no ar, nunca sei o que se passa naquela tola.	What a fool she is, always with her head in the air, I never know what's going on with that fool. <b>(Sugestão: air → clouds; with that fool → in that head)</b>	III	What a fool she is, always with her head in the air, I never know what's going on with that fool. <b>(Sugestão: air → clouds; with that fool → in that head)</b>	III
Troço	Troço sempre do ministro das infra-estruturas quando passo este troço na A29, está tão mal projetado.	I always see the Minister of Infrastructure when I pass this section on the A29, it is so poorly designed. <b>(Sugestão: see → mock)</b>	III	I always make fun of the infrastructure minister when I pass this stretch of the A29, it's so badly designed. <b>(Sugestão: make fun of → mock)</b>	II
Vede	Vede aquele rapaz. Espero que vede tudo o que está contaminado.	Look at that boy. I hope he seals everything that is contaminated.	I	Look at that boy. I hope you see everything that's contaminated. <b>(Sugestão: you → he; see → seals)</b>	III
Zelo	Para mim, que zelo sempre pelos meus, nunca me parece excesso de zelo.	For me, who always takes care of my own, it never seems overly zealous.	I	For me, who always takes care of my own, it never seems like overzealousness.	I

**Lista de palavras homógrafas falsas, das frases produzidas e respetivas traduções (Nível I - Correta; Nível II - Aceitável; Nível III - Incorreta):**

Palavras	Frase	Tradução <i>Google Translate</i>	Nível	Tradução <i>DeepL</i>	Nível
Alívio Alívio	Vou tomar a medicação para ver se sinto algum alívio neste joelho e alívio a mente também.	I'm going to take the medication to see if I feel any relief in this knee and ease my mind as well.	I	I'm going to take the medication to see if I can get some relief from this knee and ease my mind too.	I
Análise Analise	Doutor, analise o resultado desta análise o mais rapidamente possível.	Doctor, analyze the result of this analysis as soon as possible.	I	Doctor, please analyse the results of this test as soon as possible. <b>(Sugestão: please → Ø; results → result)</b>	I
Ânimo Animo	Com o teu ânimo constante até me animo nos piores momentos.	With your constant encouragement, I even cheer myself up in the worst moments.	I	With your constant encouragement, I even cheer up in the worst of times.	I
Andámos Andamos	Andamos cansados porque na semana passada andámos muito atarefados.	We are tired because last week we were very busy.	I	We're tired because we were so busy last week.	I
Anúncio Anuncio	Ao ver o anúncio na Internet, tomei a minha decisão: anuncio a venda do meu carro ainda hoje.	When I saw the ad on the Internet, I made my decision: I will announce the sale of my car today.	I	When I saw the advert on the Internet, I made up my mind: I'm going to sell my car today. <b>(Sugestão: sell my → announce the sale of my)</b>	III
Àquele Aquele	Aquele gato é muito semelhante àquele outro ali.	That cat is very similar to that other one over there.	I	That cat is very similar to the other one over there.	I
Assédio Assedio	Sou totalmente contra o assédio, seja de que tipo for, nunca assedio ninguém.	I am completely against harassment, of any kind, I never harass anyone.	I	I'm totally against harassment of any kind, I never harass anyone.	I
Avó Avô	A minha avó e o meu avô vêm jantar cá a casa hoje.	My grandmother and grandfather are coming to our house for dinner today.	I	My grandmother and grandfather are coming round for dinner tonight.	I
Baía	As baias de	Safety bays in the	III	The safety bays in	III

Palavras	Frase	Tradução Google Translate	Nível	Tradução DeepL	Nível
Baía	segurança na baía garantiram que não houvesse acidentes durante a tempestade.	bay ensured there were no accidents during the storm. <b>(Sugestão: Safety bays → The barricades)</b>		the bay ensured that there were no accidents during the storm. <b>(Sugestão: safety bays → barricades)</b>	
Beneficiária Beneficiaria	Ao ser beneficiária desse investimento, sempre idealizei o que a minha filha beneficiaria no futuro com os resultados.	Being the beneficiary of this investment, I always imagined what my daughter would benefit from the results in the future.	I	As a beneficiary of this investment, I always imagined how my daughter would benefit from the results in the future.	I
Crítico Crítico	Sou muito crítico, pelo que critico tudo o que consumo artisticamente.	I'm very critical, so I criticize everything I consume artistically.	I	I'm very critical, so I criticise everything I consume artistically.	I
Centrífuga Centrifuga	O laboratório de química tem uma centrífuga, por favor centrifuga lá as amostras de sangue que recolheste.	The chemistry laboratory has a centrifuge, please centrifuge the blood samples you collected there.	I	The chemistry lab has a centrifuge, please centrifuge the blood samples you've taken there.	I
Coágulo Coagulo	Eu coagulo muito rápido, o que justifica o coágulo que estás a ver no filtro de sangue.	I clot very quickly, which justifies the clot you see in the blood filter.	I	I clot very quickly, which explains the clot you see on the blood filter.	I
Contínuo Continuo	Eu continuo sem perceber a relação entre o fluxo contínuo e a resistência que o professor de física explicou.	I still don't understand the relationship between continuous flow and resistance that the physics teacher explained.	I	I still don't understand the relationship between continuous flow and resistance that the physics teacher explained.	I
Contrário Contrario	Sabes bem que nunca te contrário, muito pelo contrário, estou sempre do teu lado.	You know that I never contradict you, quite the contrary, I am always on your side.	I	You know that I never go against you, on the contrary, I'm always on your side. <b>(Sugestão: go against → contradict)</b>	II
Distância Distancia	Por favor, distancia o nosso carro do da frente, repara na distância mínima	Please distance our car from the one in front, notice the minimum distance	I	Please distance our car from the one in front, look at the minimum distance	I

Palavras	Frase	Tradução Google Translate	Nível	Tradução DeepL	Nível
	que nos separa.	that separates us.		between us.	
Denúncia Denuncia	Aquela denúncia, denuncia muito mais do que apenas um crime.	That complaint denounces much more than just a crime. <b>(Sugestão: complaint → accusation)</b>	II	That accusation denounces much more than just a crime.	I
Depositária Depositaria	Gostaria que fosses a minha depositária, depositaria em ti toda a minha confiança.	I would like you to be my trustee, I would place all my trust in you.	I	I'd like you to be my trustee, I'd put all my trust in you.	I
Dúvida Duvida	O Pedro duvida que a resposta esteja correta, pelo que vai enviar a sua dúvida ao professor.	Pedro doubts that the answer is correct, so he will send his doubt to the teacher.	I	Peter doubts that the answer is correct, so he sends his question to the teacher. <b>(Sugestão: sends → will send)</b>	II
Está Esta	Está muito frio, leva antes esta camisola mais quente.	It's very cold, so take this warmer sweater first. <b>(Sugestão: first → instead)</b>	II	It's too cold, take this warmer jumper instead.	I
Fábrica Fabrica	Esta fábrica fabrica chocolates.	This factory makes chocolates.	I	This factory makes chocolates.	I
Fotocópia Fotocopia	Fotocopia os meus apontamentos e, nessa fotocópia, poderás sublinhar à vontade.	Photocopy my notes and, on that photocopy, you can underline as you wish.	I	Photocopy my notes and on that photocopy you can underline at will.	I
Fotógrafo Fotografo	Como fotógrafo que sou, inevitavelmente, fotografo muitos casamentos.	As a photographer, I inevitably photograph a lot of weddings.	I	As a photographer, I inevitably shoot a lot of weddings.	I
Hábito Habito	Habito na última casa da minha rua, por isso tenho o hábito de admirar as decorações das outras casas.	I live in the last house on my street, so I have a habit of admiring the decorations in other houses.	I	I live in the last house on my street, so I have a habit of admiring the decorations on other houses.	I
Matrícula Matricula	Enquanto ela se matricula na	As she enrolls in college, we can look	II	While she's enrolling at college, we can	I

Palavras	Frase	Tradução Google Translate	Nível	Tradução DeepL	Nível
	faculdade, podemos procurar informações sobre a matrícula do seu irmão no centro de estudos.	for information about her brother's enrollment at the study center. <b>(Sugestão: enrolls → is enrolling)</b>		look for information about your brother's enrolment at the study centre.	
Médico Medico	Nunca me medico sem consultar um médico antes.	I never take medication without consulting a doctor first. <b>(Sugestão: take medication → self-medicate)</b>	II	I never take medication without consulting a doctor first. <b>(Sugestão: take medication → self-medicate)</b>	II
Pêlo Pelo	Pelo que o dermatologista viu, isto não passa de um pêlo encravado.	From what the dermatologist saw, this is nothing more than an ingrown hair.	I	From what the dermatologist saw, this is nothing more than an ingrown hair.	I
Prática Pratica	Ela tem muita prática neste tipo de exercícios porque pratica muito.	She has a lot of practice in this type of exercise because she practices a lot.	I	She has a lot of practice in this type of exercise because she practises a lot.	I
Pronúncia Pronuncia	A tua pronúncia em inglês é fenomenal, pronuncia <i>Yorkshire</i> para eu ouvir de novo.	Your pronunciation in English is phenomenal, pronounce Yorkshire so I can hear it again.	I	Your English pronunciation is phenomenal, pronounce Yorkshire so I can hear it again.	I
Relês Reles	Este livro é reles, não entendo porque motivo o relês.	This book is rubbish, I don't understand why I read it again. <b>(Sugestão: rubbish → cheap; I → you are; read it again → re-reading it)</b>	III	This book is rubbish, I don't understand why you're re-reading it. <b>(Sugestão: rubbish → cheap)</b>	II
Rótula Rotula	Ele partiu a rótula. O médico rotula este tipo de acidentes desportivos como evitáveis.	He broke his kneecap. Doctors label these types of sports accidents as preventable. <b>(Sugestão: Doctors → The doctor; these types → this type)</b>	II	He broke his kneecap. The doctor labels this type of sporting accident as avoidable.	I

Palavras	Frase	Tradução <i>Google Translate</i>	Nível	Tradução <i>DeepL</i>	Nível
Sábia Sabia	Ela é tão sábia que já sabia tudo o que lhe ia explicar.	She is so wise that she already knew everything I was going to explain to her.	I	She's so wise that she already knew everything I was going to explain to her.	I
Secretária Secretaria	Dei o teu recado à secretária na secretaria da escola.	I gave your message to the secretary in the school office.	I	I gave your message to the secretary in the school office.	I

**Lista de palavras homónimas das frases produzidas e respetivas traduções (Nível I - Correta; Nível II - Aceitável; Nível III - Incorreta):**

Palavra	Frase	Tradução <i>Google Translate</i>	Nível	Tradução <i>DeepL</i>	Nível
Amo	Amo que o meu cão me considere o seu amo.	I love that my dog considers me his master.	I	I love that my dog thinks of me as his master.	I
Cabo	O cabo militar está muito habituado a manejar o cabo dessa ferramenta.	The military corporal is very used to handling the handle of this tool.	I	The military corporal is very used to handling the handle of this tool.	I
Caminho	Caminho o meu próprio caminho, não sigo as pegadas de ninguém.	I walk my own path, I don't follow in anyone else's footsteps.	I	I walk my own path, I don't follow in anyone else's footsteps.	I
Canto	Gosto quando canto em casas de fado nas quais o palco fica num canto.	I like it when I sing in fado houses where the stage is in a corner.	I	I like it when I sing in fado houses where the stage is in a corner.	I
Cedo	O meu filho tem de se levantar cedo todos os dias, nisso não cedo.	My son has to get up early every day, that's not early. <b>(Sugestão: that's not early → I don't give in on that / I don't compromise on that)</b>	III	My son has to get up early every day, not early at that. <b>(Sugestão: not early at that → I don't give in on that / I don't compromise on that)</b>	III
Era	Na era dos descobrimentos era tudo muito diferente.	In the age of discovery everything was very different.	I	In the age of discovery, everything was very different.	I
Fecho	Antes de sair de casa fecho a porta devidamente e o fecho das janelas também.	Before leaving the house, I close the door properly and the windows as well. <b>(Sugestão: windows as well → windows' latches as well)</b>	III	Before I leave the house, I close the door properly and close the windows too. <b>(Sugestão: windows too → windows' latches too)</b>	III
Fumo	Eu fumo, mas o fumo não me incomoda.	I smoke, but the smoke doesn't bother me.	I	I smoke, but it doesn't bother me. <b>(Sugestão: it → the smoke)</b>	III
Leve	Leve este lenço também, combina com tudo e é muito	Take this scarf too, it goes with everything and is	I	Take this scarf too, it goes with everything and is	I

Palavra	Frase	Tradução Google Translate	Nível	Tradução DeepL	Nível
	leve.	very light.		very light.	
Manga	As mangas da camisola são amarelas como a polpa de uma manga.	The sleeves of the sweater are yellow like the pulp of a mango.	I	The sleeves of the shirt are yellow like the flesh of a mango.	I
Morro	Morro de cansaço sempre que subo aquele morro.	I die of exhaustion every time I climb that hill.	I	I get tired every time I go up that hill. <b>(Sugestão: get tired → die of exhaustion)</b>	II
Rio	Rio de alegria quando vejo o rio.	I laugh with joy when I see the river.	I	I laugh with joy when I see the river.	I
São	As crianças estão a são e salvo, são muito resistentes.	The children are safe and sound, they are very resilient.	I	The children are safe and sound, they are very resilient.	I
Ser	O ser humano caracteriza-se por ser racional.	Human beings are characterized by being rational.	I	Human beings are characterised by being rational.	I
Trago	Trago sempre comigo uma garrafa de água, para que possa dar um trago caso precise.	I always carry a bottle of water with me, so I can have a drink if I need to. <b>(Sugestão: drink → sip)</b>	III	I always carry a bottle of water with me so that I can take a drink if I need to. <b>(Sugestão: drink → sip)</b>	III
Vão	Vão para o carro, tenho que garantir que o meu filho não desce a deslizar pelo vão das escadas.	Go to the car, I have to make sure my son doesn't slide down the stairs. <b>(Sugestão: stairs → stairwell)</b>	III	Get in the car, I have to make sure my son doesn't slide down the stairwell. <b>(Sugestão: Get in → Go to)</b>	III
Verão	Verão toda a minha coleção de biquínis no próximo Verão.	You'll see my entire bikini collection next summer.	I	You'll see my entire bikini collection next summer.	I

# CASE STUDY: TRANSLATION FROM ENGLISH INTO EUROPEAN PORTUGUESE USING GENDER-NEUTRAL LANGUAGE

## DO AI CHATBOTS PERFORM BETTER THAN MT SYSTEMS?

Lisbeth Ferreira<sup>7</sup>

[lferreira@letras.up.pt](mailto:lferreira@letras.up.pt)

FACULDADE DE LETRAS DA UNIVERSIDADE DO PORTO (PORTUGAL)

**ABSTRACT.** There is a growing demand for gender-neutral writing in translation tasks motivated by the need for inclusion, neutrality or economic factors. Studies suggest that current Machine Translation (MT) and Artificial Intelligence (AI) systems tend to mirror the source text's (and real-world) gender bias or default generic masculine. It is important to understand whether, with documents in a gender-neutral style, the MT and AI tools provide better gender-neutral translations by resorting to standard and established grammatical solutions. Additionally, the effectiveness of AI systems, with customised prompts, in surpassing MT limitations regarding gender-neutral translation is also analysed.

**KEY-WORDS.** Gender-Neutral Language, Gender-Neutral Translation, Machine Translation, Artificial Intelligence (AI), Prompts.

**RESUMO.** A procura pela escrita de género neutro na tradução tem aumentado por vários motivos que vão desde a inclusão e a neutralidade até aos fatores económicos. Os estudos sugerem que a Tradução Automática (TA), fornecida por motores de tradução e sistemas de Inteligência Artificial (IA), tende a refletir o género do texto de partida ou a utilizar o masculino genérico. É importante compreender se, com documentos escritos num estilo de género neutro, a TA oferece melhores resultados em termos de escrita de género neutro, recorrendo a soluções gramaticais padrão e estabelecidas. Analisa-se, também, a eficácia de *prompts* personalizados nos sistemas de IA para ultrapassar as limitações da TA neste tipo de escrita.

**PALAVRAS-CHAVE.** Linguagem Neutra de Género, Tradução Neutra de Género, Tradução Automática, Inteligência Artificial (IA), *Prompts*.

---

<sup>7</sup> Estudante do 3.º ano do 3.º Ciclo de Estudos em Ciências da Linguagem.

## 1. Introduction

Over the last few years, there has been an increasing demand for the use of Gender-Neutral Language (GNL) (Cho, Kim, J. Kim, S. & Kim, N. 2019b; Kostikova, Daems & Lazarov 2023; Sun, Webster, Shah, Wang & Johnson 2021). One of the main reasons for this is a broader social awareness in terms of equality, inclusivity and recognition of different gender identities so as to respect the identity of the speaker/author and to avoid misgendering text recipients. Moreover, the use of GNL is also particularly useful when the gender of the target audience is unknown and when there is a need to make communications more universal and suitable for a diverse audience.

In institutional communications such as official documents, business communications, or academic papers, the use of GNL is becoming standard practice (Piergentili, Fucci, Savoldi, Bentivogli & Negri 2023; Sun *et al.* 2021). It helps to maintain a professional and inclusive tone, avoiding assumptions about the gender identity of the recipients while also avoiding the use of the generic masculine. In theory, the use of GNL also simplifies the translation process, especially between languages that have similar approaches to gender conventions. It can allow a more efficient, straightforward and cost-effective translation process across different linguistic and cultural contexts.

It is not the aim of this research to assess the reasons that may lead to this demand for Gender-Neutral Translation (GNT) by customers, but rather to analyse the quality of GNT provided by the currently available Machine Translation (MT) and Artificial Intelligence (AI) systems. In this context, our primary emphasis is on GNL, particularly resorting exclusively to the use of rephrasing and neutralisation strategies, thus Indirect Non-Binary Language (INL)<sup>8</sup>.

Even though gender bias affects several monolingual tasks in natural language processing (NLP), it is even more noticeable in bilingual tasks, such as translation. When

---

<sup>8</sup> It is important to note that the incorporation of neopronouns and neomorphemes (Direct Non-Binary Language – DNL) falls outside the scope of this research.

While languages such as English have strategies that are gaining widespread acceptance, Romance languages like Portuguese, Spanish and Italian cannot resort to the “they” singular solution, for example. These languages are incorporating neopronouns (e.g., EN *ze/zir* instead of *he/she/him/his/her*), neomorphemes (e.g., ES *-e/-es* instead of *-o/-os* and *-a/-as*), and other alternatives (e.g., graphemic devices such as IT *-@* instead of *-a/-o/-e/-i*), which allow reference to individuals without resorting to generic words (Piergentili *et al.* 2023).

To bridge this gap, the Portuguese “Elu” system was developed. It consists on the introduction of neopronouns, such as *elu/delu* instead of the traditional *ele/dele* or *ela/dela*, as well as neomorphemes, thus changing traditional morphological structures. For instance, the Elu system employs neutral forms like *-e* or *-es*, as in *bonite/bonites*, for example, instead of *bonito/bonitos* (masculine endings “o” and “os”) or *bonita/bonitas* (feminine endings “a” and “as”).

translating from a notional gender language (a language in which grammatical gender is inferred or implied rather than marked by grammatical forms, as is the case in English) into a grammatical gender language, the challenge becomes even greater. This process requires neutralisation strategies that can sometimes compromise the effectiveness and neutrality of the texts and always involve making choices.

In notional gender languages, gender is primarily conveyed through personal pronouns and possessive adjectives (e.g., he/him/hers; she/her/hers) and lexically gendered words (e.g., man; woman). On the other hand, grammatical gender languages, such as Portuguese (as well as Spanish, French and Italian, among others), exhibit a morphosyntactic agreement system, where various parts of speech, including verbs, determiners, and adjectives, have gender inflexions. There are also genderless languages (e.g. Finnish and Turkish) in which gendered words are limited to certain lexical pairs.

Current MT systems, such as Google Translate<sup>9</sup> and DeepL<sup>10</sup>, operate at the sentence level and are based on large corpora that include texts from different subjects and styles. As a result, they tend to follow the morphological number and gender cues used in the source text. In general, this can cause information to be lost, namely the lack of number and gender agreement when translating from a gender-marked to a gender-poorer language, often leading to morphological mistakes that ultimately affect the quality of the translation. When it comes to gender, the issue extends beyond potential grammatical agreement problems; the absence of anaphoric or cataphoric reference can also result in the use of the generic masculine (which may reveal gender bias according to recent studies) instead of a non-marked form (Cho, Kim J., Kim, M. & Kim, N. 2019a; Sun *et al.* 2021). While human translators use contextual information to ensure gender agreement or to avoid bias, most current MT systems do not (Vanmassenhove, Hardmeier & Way 2019), and they depend mainly on large corpora and statistical information.

Although not specifically designed for translation tasks, AI systems leverage extensive contextual understanding and nuanced language generation capabilities, which can potentially enhance gender-neutral translations. This potential improvement forms the basis for investigating whether AI systems can outperform traditional MT systems in producing gender-neutral translations.

---

<sup>9</sup> <https://translate.google.com/?sl=en&tl=pt&op=translate>

<sup>10</sup> <https://deepl.com>

In this case study, we compare the translations provided by MT systems (Google Translate and DeepL) and AI systems (ChatGPT<sup>11</sup> and Google Bard<sup>12</sup>) for sentences extracted from gender-neutral writing guidelines<sup>13</sup> to assess whether the latter provide better Gender-Neutral Translation (GNT) results than the former. We hypothesise that more detailed prompts will yield better results with AI systems, as recent studies (Jiao, Wang, Huang, Wang, Shi & Tu 2023; Yamada 2023) have shown that customising prompts with specific instructions and context improves the quality of AI-generated translations. In the provided hyperlink<sup>14</sup>, it is possible to access the 38 selected sentences, the international guidelines used, and the translations obtained for the 38 sentences that constitute the dataset. The AI chatbots were provided with prompts ranging from simple instructions with no information on GNL to richly informed ones meant to function as a sort of ‘translation brief’ aimed at producing fully gender-neutral translations. It should be considered that the sentences tested in this study, as they are taken from international guidelines already written using GNL, have greater potential for generating good results in terms of gender-neutral writing. A similar case study with a completely different corpus would likely have rather different results.

As this is not a real translation task with a specific purpose, audience and application, but rather a study to assess if GNL is used in two types of automatic translation tools, no post-editing was done on any of the translations provided by the MT or AI systems. The assessment was made: a) taking into account only the GNL request; and b) based on the assumption that translation support tools as the ones used for this study require post-editing; therefore, a final version cannot be used “as-is”, i.e., without the intervention of a translator.

---

<sup>11</sup> <https://chatgpt.com>

<sup>12</sup> <https://gemini.google.com>

On 8 February 2024, Google Bard changed its name to Gemini. As the case study was developed before, the former designation was maintained - <https://blog.google/products/gemini/bard-gemini-advanced-app/>.

<sup>13</sup> The following guidelines were used as references for Table 1 and for the translation data analysis. English: E1 - United Nations, 2018; E2 - General Secretariat, Council of the European Union, 2018; E3 - European Parliament, 2018; E4 - North Atlantic Treaty Organisation, 2020 / Portuguese: P1 - Commission for Citizenship and Gender Equality (CIG); P2 - European Parliament, 2018; P3 - IGEN - Fórum; P4 - University of Porto.

<sup>14</sup> [https://github.com/LF-DCL/Gender\\_Neutral\\_Translation\\_Case\\_Study](https://github.com/LF-DCL/Gender_Neutral_Translation_Case_Study) - Includes a downloadable Excel file with the complete dataset, as well as the prompts/interactions used in the study.

## 2. Related work

Research on gender bias in translation technologies has gained significant prominence over the last few years, and the presence of bias in MT has been extensively studied. This bias has a negative impact on NLP (Sun, Gaut, Tang, Huang, ElSherief, Zhao, Mirza, Belding, Chang & Wang 2019), namely MT. According to Friedman & Nissenbaum (1996), bias in computer systems can be divided into three categories: pre-existing bias, technical bias, and emergent bias. These categories can also be applied to gender bias in MT (Savoldi, Gaido, Bentivogli, Negri & Turchi 2021).

Mitigating gender bias in MT requires a comprehensive approach encompassing historical, technical, and emergent dimensions. Refining data curation, making informed architectural choices, and aligning systems with evolving societal values are essential for inclusive and equitable language processing.

Underrepresentation and stereotyping in MT involve biased visibility and negative generalisations. Underrepresentation occurs when MT misrepresents women or overlooks non-binary individuals, which has an impact on communication. Stereotyping perpetuates biased views, affecting self-esteem and reinforcing societal prejudices. The widespread integration of MT in web applications magnifies these challenges, emphasising the interconnected nature of underrepresentation and stereotyping (Savoldi *et al.* 2021). Studies of GNL either approach the incorporation of meta-information in the source text or focus on rewriting techniques (Vanmassenhove, Emmery & Shterionov 2021). Most studies in this field adopt a binary perspective on gender, with a considerably smaller prevalence of studies taking a non-binary approach.

### 2.1. Binary approach

The binary approach refers to the use of gender-specific terms rather than gender-neutral alternatives. This approach reflects traditional linguistic norms, where masculine forms are often used generically, reinforcing stereotypes and perpetuating inequalities (Piergentili *et al.* 2023). Translation with a generic masculine approach predominates due to volume-related and social bias-related aspects. This is why “cops are usually men and nurses are usually women” (Cho *et al.* 2019b).

Dinu, Mathur, Federico & Al-Onaizan (2019) explore a method for eliminating bias by augmenting the MT input with additional information: a black-box approach where MT learns

how to use external terminology provided. Meanwhile Prates, Avelar & Lamb (2020) analyse the phenomenon of machine bias in the translation of pronouns with a comprehensive list of job positions in 12 different gender-neutral languages that they translate into English using Google Translate. They conclude that Google Translate has a higher probability of translating sentences with masculine pronouns than with feminine or neutral pronouns, especially in STEM (Science, Technology, Engineering and Mathematics) areas. Vanmassenhove *et al.* (2019), for their part, present a compilation of vast datasets with information from speakers of 20 languages and develop a set of experiments that incorporate the gender information through tags into neural machine translation for different language pairs, concluding that this can greatly improve translation quality in some pairs.

Stanovsky, Smith & Zettlemoyer (2019) present a challenge set and assessment protocol for analysing gender bias in MT for eight target languages with grammatical gender, based on a morphological analysis (Spanish, French, Italian, Russian, Ukrainian, Hebrew, Arabic and German). They conclude that all tested MT systems (Google Translate, Microsoft Translator, Amazon Translate and Systran) reveal bias. Aligned with these findings, Kocmi, Limisiewicz & Stanovsky (2020) use the WinoMT challenge set to analyse coreference and bias in translation from English into grammatically gendered languages, also observing that all systems they tested show worse performance when the professional in question is female. Saunders & Byrne (2020) introduce a novel perspective on addressing gender bias through a domain adaptation framework. Their study demonstrates significant enhancements in the WinoMT challenge set when adapting to datasets across three language pairs, advocating small-domain adaptations as a more efficient strategy for mitigating gender bias in MT.

Rescigno, Vanmassenhove, Monti & Way (2020) present the results of an assessment of MT systems, namely Google Translate, DeepL and Bing Microsoft Translation, and the rendering of grammatical gender, providing a qualitative, quantitative and statistical analysis of the “frequency of female, male and neutral forms in the translation of a list of personality adjectives, nouns related to professions and bigender nouns”. Pursuing a different direction, Bentivogli, Savoldi, Negri, Di Gangi, Cattoni & Turchi (2020) question what happens with speech translation in particular, aiming to understand if input as audio signals can provide additional information resulting in a reduction of bias.

Research has shown that the presence of gender bias in MT is strong. This bias reinforces inequalities as does the use of the generic masculine, which is a manifestation of linguistic normativity (Piergentili *et al.* 2023).

## 2.2. Non-binary approach

A non-binary approach involves strategies to avoid gender-specific language, promoting inclusivity by using gender-neutral terms. Non-binary language can be divided into two categories: direct and indirect. Direct non-binary language (DNL) uses neopronouns and neomorphemes explicitly designed to avoid gender-specific language. Indirect non-binary language (INL) resorts to pluralisation or sentence restructuring to avoid gender-specific terms without explicitly changing the original words (Vanmassenhove *et al.* 2021).

Recent investigations focus on non-binary approaches, such as gender-neutral rewriting. These studies analyze the capacity of MT systems to adapt to GNL through rewriting and post-editing steps.

Sun *et al.* (2021) propose a rewriting and post-processing approach for generating sentences that incorporate the singular “they” and an evaluation benchmark of 500 sentences from Twitter, Reddit, jokes and movie quotes. They define a new task where the system is required to change to neutral forms (e.g. changes such as fireman to firefighter) with “no human-labelled training data”. The authors emphasise the scarcity of natural corpora featuring gender-neutral sentences, significantly hindering non-binary GNL studies. Vanmassenhove *et al.* (2021) develop a similar study but with larger synthetic and natural benchmarks.

Kostikova *et al.* (2023) study the capabilities of Adaptive Machine Translation (AMT) when using GNL, namely focusing on non-binary language in English-German translation using the Modern MT engine. They investigate two gender-neutral writing strategies in German, both with Direct Non-binary Language (DNL): the DeE-System and the gender star system. Despite signs of masculine bias in Modern MT, their findings suggest AMT systems can adapt to gender-neutral writing, and increased exposure to gender-neutral corpora can facilitate their evolution.

## 3. Gender-Neutral Language

Languages contemplate three categories of gender: natural gender, grammatical gender, and social gender (Rescigno *et al.* 2020). Natural gender is typically linked to the biological sex of a person or animal. Grammatical gender, on the other hand, may not align with a word's semantic category and varies across languages, being shaped by cultural

contexts. Finally, social gender involves associating the natural gender of a person with a word based on specific properties, often influenced by societal stereotypes.

The use of GNL proves to be a valuable solution with a wide range of benefits. Besides fostering inclusivity and respecting different identities, it enables the elimination of misgendering, especially when gender identification is a challenge due to insufficient data.

Piergentili *et al.* (2023) highlight three main desiderata for determining if a gender-neutral output is required: a) refraining from expressing gender when it cannot be reliably assumed from the source, which is particularly relevant when translating from a notional gender language into a grammatical gender one; b) using appropriate expressions of gender if they are indirectly conveyed in the source – gender cues (such as third-person pronouns, terms of address, and gender-specific nouns) play a crucial role in determining whether or not a gender-neutral translation is needed; c) refraining from propagating generic masculine in the source language – even though recognizing gender cues may not be straightforward, especially with generic masculine.

### **3.1. Grammatical gender**

All linguistic systems encode or represent gender in some form, whether through grammatical structures, lexical choices, or social conventions. In English (as in Danish, for instance), a notional gender language, gender is primarily conveyed through personal pronouns and possessive adjectives (e.g., he/him/his; she/her/hers) and lexically gendered words (e.g., man; woman). On the other hand, grammatical gender languages like Portuguese (as well as Spanish, French and Italian, among others) exhibit a morphosyntactic agreement system, where various parts of speech, including verbs, determiners, and adjectives, have gender inflexions. There are also “genderless” languages (e.g., Finnish and Turkish) that lack grammatical gender but still convey gender through lexical or cultural mechanisms. In these languages, gender distinction is often based on “animacy” in some lexical pairs.

These differences are particularly noteworthy in translation, especially when the source language lacks gender information about a referent and the target language is a grammatical gender language, as is the case in translation from English into Portuguese. In terms of automatic translation, this language pair is prone to mistakes related to gender bias.

### 3.2. Guidelines for neutral writing

Several international institutions provide guidelines on gender-neutral writing. Despite the fact that these guidelines address monolingual contexts, they can provide useful insights for translation tasks that require GNL.

The United Nations (2018), the European Union (2018), the European Parliament (2018) and NATO (2020) (along with others) provide guidelines for writing in GNL in English. The guidelines share a common emphasis on the importance of GNL in challenging stereotypes and promoting inclusivity. As for Portuguese, the guidelines made available by the European Parliament (2018), the Portuguese Economic and Social Council (2021) and the IGEN Forum (2018) similarly encourage the use of gender-inclusive language, aligning with international efforts.

### 3.3. Neutralisation strategies

Efforts to promote linguistic fairness and inclusivity for all gender identities can be categorised into two primary approaches: i) INL, the adoption of gender-neutral formulations, and ii) DNL, the introduction of novel linguistic elements (Piergentili *et al.* 2023).

DNL solutions, such as the use of neopronouns and neomorphemes, cannot be neglected and are indeed necessary in translation in specific contexts, especially when the speaker deliberately uses a neutral pronoun. Despite the importance of these solutions, their usage is not universally deemed grammatically correct or socially acceptable in many linguistic environments and, as mentioned earlier, this linguistic solution falls outside the scope of this text. To employ gender-inclusive language in more formal settings, the use of INL with established gender-neutral conventions within the standard language is expected. Neutralisation strategies can be as simple as choosing gender-neutral words (*firefighter* instead of *fireman*, for example) or may require a total sentence rephrasing without neoelements.

Table 1 presents a summary of neutral solutions and strategies that comply with institutional guidelines as well as with gender neutralisation techniques proposed by Vanmassenhove *et al.* (2021) and Piergentili *et al.* (2023), and follows the presentation proposed by the latter.

<b>A. Epicene synonyms</b>		
EN	E4	<i>Chairman</i> → <u>Chair(person)</u>
PT	P4	<i>Os enfermeiros/As enfermeiras</i> → <u>O pessoal de enfermagem</u>
<b>B. Pluralisation/Singularisation (towards generic or epicene forms)</b>		
EN	E1	A judge must certify that <i>he</i> has familiarised <i>himself</i> with... → All <u>judges</u> must certify that <u>they</u> have familiarised <u>themselves</u> with...
PT	P2	<i>Os diretores...</i> / → <u>A direção...</u>
<b>C. Relative and indefinite pronouns</b>		
EN	E4	If a staff member is not satisfied..., <i>he</i> can ask for a rehearing. → Any staff member <u>who</u> is not satisfied... can ask for a rehearing.
PT	P2	<i>Os requerentes</i> devem... / → <u>Quem</u> proceder à apresentação de um requerimento deve...
<b>D. Collective and role nouns</b>		
EN	§	Please contact one of the <i>waiters</i> . / → Please contact our <u>staff</u> .
PT	P2	<i>Os supervisores...</i> / → <u>As pessoas responsáveis pela supervisão...</u>
<b>E. Omission</b>		
EN	§	A person must reside... before <i>he</i> may apply for permanent residence. → ...before <u>  </u> applying for permanent residence.
PT	P2	<i>O juiz</i> proferirá a sentença. / → A sentença será proferida.
<b>F. Repetition</b>		
EN	E2	A manager may apply... if permission has been granted by <i>his</i> institution. → ...if permission has been granted by <u>that manager</u> 's institution.
PT	§	A direção pode recorrer da decisão caso <i>esta</i> assim o pretenda. → A direção pode recorrer da decisão caso <u>a direção</u> assim o pretenda.
<b>G. Passive voice</b>		
EN	E4	Each action officer must send <i>his</i> document. / → Documents <u>must be sent</u> .
PT	P1	<i>O candidato</i> deve enviar o formulário de candidatura até dia x. → O formulário de candidatura deve <u>ser enviado</u> até dia x.
<b>H. Imperative forms</b>		
EN	E4	Each staff member is requested to submit <i>his</i> information. → Please <u>submit</u> all information.
PT	§	<i>O candidato</i> deve enviar o seu CV. / → <u>Envie</u> o seu CV.
<b>I. Impersonal forms</b>		
PT	§	<i>O cidadão</i> perde o direito... / → <u>Perde-se</u> o direito...

Table 1 - Examples of neutralisation strategies<sup>15</sup>.

<sup>15</sup> Table format follows Piergentili *et al.* (2023) adapted for EN-PT-PT. In *red, italic*, the generic masculine formulations; in *green, underlined*, the gender-neutralisation. Column 2 identifies source guidelines; self-created examples (§) or omissions, when not applicable.

The neutralisation strategies presented in the guidelines are targeted towards monolingual tasks. However, they can also be applied in bilingual tasks, especially when gender-neutral writing is necessary in the target text, regardless of whether the source text adheres to this type of writing or not.

#### **4. Proposed method**

To develop this comparative case study, 38 sentences in English were taken from the international guidelines on gender-neutral writing, as outlined by the strategies in Table 1. The sentences that make up the analysed dataset<sup>16</sup> are written according to these neutralisation strategies. The aim of this study is to compare the performance of MT systems and AI platforms in the translation of these gender-neutral texts, particularly when AI models are provided with customised prompts. For that, all the sentences were translated with two MT systems and two AI platforms to determine which technology offers better results in maintaining gender-neutral language.

The selected MT systems were DeepL and Google Translate, which are MT tools, while the other two systems used, ChatGPT and Google Bard<sup>17</sup>, are Artificial Intelligence (AI) platforms that, although capable of generating translations, were not specifically designed for this purpose. In the case of the AI platforms, three versions of the translations were collected: the first version was generated from a simple and direct prompt; the second from a slightly more GNL-specific prompt; and the third prompt was based on the concept of the “translation-brief” (Nord 1997).

##### **4.1. Translation prompts**

One of the features of AI chatbots is the ability to provide detailed instructions, that is to say, fine-tuning prompts so as to render the desired output. The use of AI platforms is still recent, and users are exploring ways to make the most of their features. It is here questioned whether it is possible to achieve a better GNL translation when the AI features are combined with customised prompts that include the translation “skopos” (i.e., the purpose or intended function of the translation) and further details that work as a “translation brief” (understood as

---

<sup>16</sup> The full dataset is available at: [https://github.com/LF-DCL/Gender\\_Neutral\\_Translation\\_Case\\_Study](https://github.com/LF-DCL/Gender_Neutral_Translation_Case_Study).

<sup>17</sup> ChatGPT 3.5 and Google Bard Pro 1.0 – both versions available for free online at the time of the study.

a set of instructions outlining the purpose, target audience, and other specific requirements of the translation) (Nord 1997, 2018; Pym 1993; Reiss & Vermeer 2014).

Other recent studies have been exploring the possibility of enhancing translations generated by AI systems through the use of detailed and customised prompts.

Yamada (2023) explores how to optimise MT in ChatGPT by customising prompts, namely with the translation's purpose and target audience, and concludes that this approach can, in fact, greatly increase the output quality. His assessment, developed from a professional translator's perspective, finds that this method improves translation quality based on industry standards, especially for marketing documents and idioms tied to specific cultures.

Jiao *et al.* (2023) also researched ChatGPT's capabilities in terms of MT and concluded that, when strong prompts are used, the quality of the output is equivalent to that of commercial MT systems, but it shows greater limitations with low-resource language pairs. Gao, Wang & Hou (2023) similarly propose several translation prompts, including details on the translation task, context domain data, and part-of-speech (POS) tags, to further unlock ChatGPT's translation potential. Like Yamada (2023) and Jiao *et al.* (2023), the authors conclude that ChatGPT performs better when more detailed prompts are employed.

## 5. Analysis and results

Each of the 38 sentences that constitute the analysed dataset yielded eight different translations: one by each MT system (Google Translate and DeepL) and three by each AI system (ChatGPT and Google Bard). Three different prompts were used (ranging from simple ones with no information on GNL to customised and detailed prompts), each prompt generating three different outcomes.

The first translations obtained from the AI systems (identified as GPT1 and BD1 in the dataset) were comparable to those generated by DeepL and Google Translate, as no information about GNL was provided. The chatbots were only required to translate the sentences from English into European Portuguese. Providing such limited information in the first prompt was crucial to determine if the customised prompts would, as hypothesised, generate better results.

The results of the second translation produced by the AI systems (GPT2 and BD2) were obtained from an improved prompt deemed fitting, for instance, for human translators who were to produce a gender-neutral translation of the sentences.

The first attempt at generating a gender-neutral translation using a customised prompt was the following: *“Translate the sentences below from English into European Portuguese considering the following instructions. Please translate the following sentences from English into European Portuguese using a gender-neutral language. You can use rewriting and reformulation strategies to avoid masculine and feminine nouns, pronouns, adjectives, and verbs, for instance, passive voice, imperative forms, and impersonal forms among others. You can also replace gendered nouns for epicene synonyms or collective nouns. You cannot use neopronouns or neomorphemes, such as “Elu”, “Delu”, “amigue”, “bonite”, nor solutions like “Car@”, “Carxs” or solutions that include the male/female like “Caro/as”, “bonito/a”.*

Before devising the final customised prompt, which was expected to produce the best results for the purposes of this study, several attempts were made with two of the sentences<sup>18</sup>. These tests to fine-tune the customised prompt were done on ChatGPT<sup>19</sup>, and the same final prompt was then used on Google Bard. It can be argued that the same process of prompts testing in order to find the most effective one might have produced a different result if run on Google Bard. However, for the purpose of this study, it was necessary to use the same prompt on both systems to achieve the third and final version of the translations.

As shown in the previously provided hyperlink, the level of detail ChatGPT was provided with was different, and additional interactions with the chatbot were necessary so as to provide the machine information regarding why some of the translations presented did not comply with the instructions given. The tests clearly demonstrated that it was quite a challenge to establish the appropriate amount of information the prompts should contain; providing either too much or too little information can create “noise” and affect the performance of the AI system. The third set of AI-produced translations of the sentences (see GPT3 and BD3 in the dataset) were obtained from the prompt that seemed to have rendered better results than the earlier tests.

The final customised prompt used was: *Translate the following sentence from English into European Portuguese. The target sentence must be totally gender-neutral. The nouns, pronouns, adjectives and verbs used in the translation CANNOT be gender-specific, so the sentences MUST be 100% neutral as the client forbids the use of gender marks for professions, job titles, groups of people, etc. Example: instead of “fotógrafo” / “fotógrafa” you*

---

<sup>18</sup> The interactions with the AI chatbots, as well as the three prompts used, can be accessed via the previously provided hyperlink.

<sup>19</sup> The tests were done on ChatGPT because it was launched before Google Bard.

could use *"profissional de fotografia"*. You can use several rephrasing, paraphrasing strategies such as passive voice, imperative, collective nouns and other strategies. Example: instead of *"Se estiver interessado..."* / *"Se estiver interessada..."* you could use *"Em caso de interesse..."* You cannot use dual gender approaches such as *"O/A candidato/a"* in any circumstance. You also cannot use neo systems such as neopronouns *"elu/delu"*.

This prompt was considered the most balanced version of all three since it provides some grammatical information but avoids complex metalanguage. It also provides some examples, suggests neutralisation strategies, and defines which expressions and solutions align with the pseudo-client's instructions.

### **5.1. Assessment criteria**

In order to analyse the translations produced by Google Translate, DeepL, ChatGPT and Google Bard, we set up a group of assessors composed of three professional translators: T1 has 1 year of experience, T2 has 4 years of experience, and T3 has over 10 years of experience.

Each of the assessors was given the 38 original sentences in English and the eight different automatic translations for each sentence, totalling 304 translations per assessor. They were instructed that the main criterion of analysis was not the overall quality of the translation. As professional translators accustomed to post-editing, they should understand that the output of MT is not intended to serve as a final translation. They were instructed to assess each of the translation versions with the understanding that they could classify sentences based on more than one criterion, when necessary, and according to the following criteria: 1) *Adequate gender-neutral translation*; 2) *Uses generic masculine*; 3) *Uses a dual gender approach (e.g. Ele/Ela; bonito/bonita)*; 4) *Uses a non-grammatical approach (e.g. todxs/tod@s/todes)*; 5) *Uses a biased-gender approach (e.g. translate nurse as feminine)*; 6) *Brazilian Portuguese*; 7) *Mixed approaches*. The three elements of the assessment group worked separately on the task and did not discuss results with each other.

## 5.2. Quantitative analysis

**Gender-neutral translations**

	GT	DL	GPT 1	GPT 2	GPT 3	BD 1	BD 2	BD 3
T1	12	15	11	13	15	10	9	9
T2	10	12	9	11	11	8	7	9
T3	11	13	11	12	13	10	9	9

Table 2 - Number of translations considered gender-neutral.

**Non gender-neutral translations**

	GT	DL	GPT 1	GPT 2	GPT 3	BD 1	BD 2	BD 3
T1	26	23	27	25	23	28	29	29
T2	28	26	29	27	27	30	31	29
T3	27	25	27	26	25	28	29	29

Table 3 - Number of translations considered Non gender-neutral.

As for the quantitative results, there was very little difference in the total number of sentences that each of the assessors considered to be written in accordance with the required criteria of gender neutrality. Above, Table 2 shows the total number of sentences considered neutral by the three assessors, while Table 3 shows the total number of sentences considered non-neutral by the three assessors. The non-neutral translation results correspond only to criteria "Uses generic masculine" (2) and "Uses a gender-biased approach" (5), as none of the systems produced translations that met criteria (3), (4) and (7). The data in Tables 2 and 3 above encompass the European and the Brazilian varieties of Portuguese. Nevertheless, variety-specific outputs were not identified since this variable did not affect gender neutrality. Yet, the number of sentences translated into Brazilian Portuguese (PT-BR), but considered correct by T1, T2 and T3 regarding the gender neutrality criterion were as follows: two results for GT; two for GPT3; six for BD1; six for BD2 and one for BD3. It should be noted that all three prompts explicitly stated that the translations must be in PT-PT.

## 5.3. Qualitative analysis

For a more detailed analysis of the qualitative data, only the cases in which the translations were unanimously rated as gender-neutral by the three assessors are analysed. To present the relevant information concisely, the following tables display only the translations,

the ratings from each assessor and the reference to the original sentence. The most relevant cases are also commented on; however, it is possible to analyse and explore all the various responses by accessing the dataset made available in the provided hyperlink, along with the 38 original source sentences identified by their corresponding reference numbers.

### 5.3.1. DeepL and Google Translate

Ref.	Translation	T1 .r	T2 .r	T3 .r
1	Cada Representante Permanente deve apresentar as suas credenciais ao Protocolo	1	1	1
4	Os planos de externalização de alguns serviços não avançaram ao ritmo previsto, uma vez que se regista uma escassez significativa de mão de obra.	1	1	1
5	Os planos de externalização de alguns serviços não avançaram ao ritmo previsto, uma vez que existe uma escassez significativa de pessoal	1	1	1
15	Serão responsáveis pela gestão dos procedimentos de contratação pública.	1	1	1
16	Uma pessoa pode não saber o seu número de contribuinte.	1	1	1
19	Cada visitante deve apresentar o seu bilhete de identidade para obter um temporário de visitantes.	1	1	1
21	Cada participante deve certificar-se de que não deixa para trás os seus documentos classificados quando sai da sala.	1	1	1
22	Cada participante deve certificar-se de que não deixa os seus documentos classificados ao sair da sala.	1	1	1
25	Para requerer a residência permanente, é necessário ter residido continuamente no país durante 5 anos.	1	1	1
27	Cada agente de ação deve enviar o seu documento à secretaria para distribuição.	1	1	1
28	Os documentos devem ser enviados ao registo para distribuição.	1	1	1

Table 4 - DeepL translations assessed as GNT by T1, T2 and T3.

Contrary to our hypothesis, that is, that the GNT results generated by AI chatbots would be more accurate due to the use of customised prompts, DeepL was the system that presented the highest number of sentences translated into a completely gender-neutral style (11 sentences). The results may be attributed to the fact that the source sentences were written in a neutral and "institutional" style. It would be worthwhile to conduct a study to verify if different types of texts, not originally written with GNL, would generate different outcomes, potentially with a higher use of the generic masculine. By looking at the sentences generated in a gender-neutral style, it may be safe to assume that the strategies used by DeepL essentially involve using impersonal or collective nouns. The phrases “Cada representante/visitante/participante” and “Uma pessoa”, for example, avoid gender markings and make sentences impersonal.

Ref.	Translation	T1 .Y	T2 .Y	T3 .Y
4	Os planos para terceirizar cerca de 19 serviços não prosseguiram no ritmo previsto, pois há escassez significativa de mão de obra.	1	1	1
5	Os planos para terceirizar cerca de 19 serviços não prosseguiram no ritmo previsto, pois há uma escassez significativa de pessoal.	1	1	1
15	Serão responsáveis pela gestão dos procedimentos de contratação pública.	1	1	1
16	Alguém pode não saber seu número fiscal.	1; 6	1; 6	1; 6
19	Cada visitante deverá apresentar o seu documento de identificação para obter um passe de visitante temporário.	1	1	1
21	Cada participante deverá certificar-se de não deixar os seus documentos confidenciais ao sair da sala.	1	1	1
22	Cada participante deverá certificar-se de não deixar os seus documentos confidenciais ao sair da sala.	1	1	1
25	Uma pessoa deve ter residido continuamente no país durante 5 anos antes de solicitar residência permanente.	1	1	1
27	Cada oficial de ação deve enviar seu documento ao cartório para distribuição.	1; 6	1; 6	1; 6
28	Os documentos devem ser enviados ao cartório para distribuição.	1	1	1

Table 5 - Google Translate translations assessed as GNT by T1, T2 and T3.

Similar to the results obtained with DeepL, the number of sentences with GNT did not vary significantly. GT yielded 10 outcomes in GNL. As for the strategies used, no significant differences were observed, and it often resorted to impersonal and collective nouns (e.g. ‘Uma pessoa’), indefinite pronouns (e.g. ‘Alguém’), and a few occurrences of the passive voice.

### 5.3.2. ChatGPT and Google Bard

Ref.	Source	Translation	T1	T2	T3
1	GPT 1	Cada Representante Permanente deve apresentar as suas credenciais ao Protocolo.	1	1	1
1	GPT 2	Cada Representante Permanente deve submeter os seus credenciais ao Protocolo.	1	1	1
1	GPT 3	Cada Representante Permanente deve submeter suas credenciais ao Protocolo.	1; 6	1; 6	1; 6
4	GPT 1	Os planos para terceirizar cerca de 19 serviços não avançaram no ritmo previsto, devido a significativas escassezes de mão de obra.	1	1	1
4	GPT 2	Os planos de externalizar cerca de 19 serviços não avançaram no ritmo previsto, devido a significativas carências de recursos humanos.	1	1	1
4	GPT 3	Os planos de externalizar cerca de 19 serviços não avançaram no ritmo esperado, uma vez que existem importantes carências de recursos humanos.	1	1	1
5	GPT 1	Os planos para terceirizar cerca de 19 serviços não avançaram no ritmo previsto, devido a significativas carências de pessoal.	1	1	1
5	GPT 2	Os planos de externalizar cerca de 19 serviços não avançaram no ritmo previsto, devido a significativas carências de pessoal.	1	1	1
5	GPT 3	Os planos de externalizar cerca de 19 serviços não avançaram no ritmo esperado, uma vez que existem importantes carências de pessoal.	1	1	1
14	GPT 3	Em caso de interesse, por favor, envie um e-mail.	1	1	1
16	GPT 1	Alguém pode não saber o seu número de contribuinte.	1	1	1
16	GPT 2	Alguém pode não saber o seu número de contribuinte.	1	1	1
16	GPT 3	Alguém pode não saber o seu número de contribuinte.	1	1	1
19	GPT 2	Cada visitante deve apresentar o seu documento de identificação para obter um passe temporário de visitante.	1	1	1
21	GPT 1	Cada participante deve garantir que não deixe os seus documentos classificados ao sair da sala.	1	1	1
21	GPT 2	Cada participante deve garantir que não deixa para trás os seus documentos classificados ao sair da sala.	1	1	1
21	GPT 3	Cada participante deve garantir que não deixa os seus documentos classificados para trás ao sair da sala.	1	1	1
22	GPT 1	Cada participante deve garantir que não deixa os seus documentos classificados ao sair da sala.	1	1	1
22	GPT 2	Cada participante deve garantir que não deixa para trás os seus documentos classificados ao sair da sala.	1	1	1
22	GPT 3	Cada participante deve garantir que não deixa os seus documentos classificados para trás ao sair da sala.	1	1	1
25	GPT 1	Uma pessoa deve ter residido continuamente no país por 5 anos antes de solicitar residência permanente.	1	1	1
25	GPT 2	Uma pessoa deve ter residido continuamente no país durante 5 anos antes de solicitar residência permanente.	1	1	1
25	GPT 3	Uma pessoa deve ter residido continuamente no país por 5 anos antes de solicitar a residência permanente.	1	1	1; 6
27	GPT 1	Cada responsável pela ação deve enviar o seu documento ao arquivo para distribuição.	1	1	1
27	GPT 2	Cada responsável pela ação deve enviar o seu documento para o registo para distribuição.	1	1	1
27	GPT 3	Cada responsável pela ação deve enviar o seu documento para o arquivo para distribuição.	1	1	1
28	GPT 1	Os documentos devem ser enviados ao arquivo para distribuição.	1	1	1
28	GPT 2	Os documentos devem ser enviados para o registo para distribuição.	1	1	1
28	GPT 3	Os documentos devem ser enviados para o arquivo para distribuição.	1	1	1

Table 6 - ChatGPT translations assessed as GNT by T1, T2 and T3.

Between the ChatGPT translation version GPT1 and versions GPT2 & 3, only one more sentence was translated following the gender neutrality criterion: GPT1 produced 9 results, while GPT2 & 3 yielded 10 GNL outcomes. However, the sentence that showed improvement between versions GPT2 & 3 varied because gender neutrality was inconsistent across versions. Thus, in GPT1 & 2, the translation of sentence 19 changed from "Todo visitante" (the generic masculine) in GPT1 to "Cada visitante" (neutral) in GPT2; one would expect the customised prompt in GPT3 to replicate or produce a corresponding result. Yet, GPT3 mistakenly retrieved the generic masculine form "Todo o visitante". Sentence 14 shows a significant improvement in GTP3 compared to the earlier versions. While GPT1 & 2 resorted to the generic masculine ("Se estiver interessado"), GPT3 yielded an effectively gender-neutral translation by using the impersonal expression "Em caso de interesse".

From Table 6, we can observe that, in the three sets of translations provided by ChatGPT, the system tended to apply impersonal and collective nouns and phrases (e.g.

“pessoa”, “cada visitante”) as well as to the use of impersonal and/or passive structures in order to produce gender-neutral outcomes<sup>20</sup>.

Ref.	Source	Translation	T1	T2	T3
4	BD 1	Os planos de terceirização de alguns 19 serviços não avançaram no ritmo previsto, pois há deficiências significativas de recursos humanos.	1	1	1
4	BD 2	Os planos de terceirização de alguns 19 serviços não seguiram o ritmo previsto, pois há uma escassez significativa de pessoal.	1	1	1
4	BD 3	Os planos de terceirização de alguns 19 serviços não têm avançado no ritmo esperado, devido a uma escassez significativa de mão de obra.	1	1	1
5	BD 1	Os planos de terceirização de alguns 19 serviços não avançaram no ritmo previsto, pois há deficiências significativas de pessoal.	1	1	1
5	BD 3	Os planos de terceirização de alguns 19 serviços não avançaram como esperado devido à escassez de pessoal.	1	1	1
6	BD 3	Antes de submeter o seu documento, envie-o ao ponto focal para revisão. O ponto focal irá devolver-lhe o documento com comentários.	1	1	1
9	BD 3	Uma pessoa que apresentar uma queixa e não estiver satisfeita com a decisão do conselho pode solicitar uma nova audiência.	1	1	1
15	BD 3	Serão responsáveis pela gestão de procedimentos de contratação pública.	1	1	1
16	BD 1	Alguém pode não saber seu número de contribuinte.	1; 6	1; 6	1; 6
16	BD 2	Alguém pode não saber seu número de contribuinte.	1; 6	1; 6	1; 6
16	BD 3	Poderá haver pessoas que não conheçam o seu número de identificação fiscal.	1	1	1
21	BD 1	Cada participante deve se certificar de não deixar seus documentos classificados para trás ao sair da sala.	1; 6	1; 6	1; 6
21	BD 2	Cada participante deve garantir que não deixe seus documentos classificados para trás ao sair da sala.	1	1; 6	1; 6
22	BD 1	Cada participante deve se certificar de não deixar seus documentos classificados para trás ao sair da sala.	1; 6	1; 6	1; 6
22	BD 2	Cada participante deve garantir que não deixe seus documentos confidenciais para trás ao sair da sala.	1; 6	1; 6	1; 6
25	BD 1	Uma pessoa deve ter residido continuamente no país por 5 anos antes de solicitar a residência permanente.	1	1	1; 6
25	BD 2	Uma pessoa deve residir continuamente no país por 5 anos antes de solicitar a residência permanente.	1	1	1; 6
25	BD 3	Para se candidatar à residência permanente, é necessário residir continuamente no país por 5 anos.	1	1	1; 6
27	BD 1	Cada oficial de ação deve enviar seu documento ao registro para distribuição.	1; 6	1; 6	1; 6
27	BD 2	Cada oficial de ação deve enviar seu documento ao registro para distribuição.	1; 6	1; 6	1; 6
28	BD 1	Documentos devem ser enviados ao registro para distribuição.	1; 6	1	1; 6
28	BD 2	Os documentos devem ser enviados ao registro para distribuição.	1; 6	1	1; 6
28	BD 3	Os documentos devem ser enviados ao registo para distribuição.	1	1	1

Table 7 - Google Bard translations assessed as GNT by T1, T2 and T3.

Google Bard usually presents a range of alternative responses to prompts, including comments and explanations. For the purposes of this study, we always considered the first proposal. We observed a very small difference in the total number of neutrally translated results among the translation outcomes resulting from the three Google Bard prompt versions: BD1 produced eight neutral results; BD2, only seven; and BD3, also eight. However, although they differ by only one sentence overall, the use of GNL from one version to the other was not consistent: GNL outcomes produced by one version were no longer gender-neutral in the next. For example, in BD1 & 3, the word "pessoal" is used as a neutral solution for the English word "staffing", while BD2 translated the same word with the generic masculine "funcionários".

<sup>20</sup> Although the 3 model prompts clearly indicated that the sentences were to be translated into PT-PT, this criterion was followed in GPT1 & 2, whereas, in GPT3, 3 of the gender-neutral translations were rendered in PT-BR.

Sentences 21, 22 and 23 are noteworthy because they complied with the gender neutrality criterion in BD1 & 2, but, in BD3, which was expected to produce better results for being the most detailed prompt, the three sentences were no longer gender-neutral. The English phrases “Each participant” (21 and 22) and “Each action officer” (23) were, at first, correctly neutralised in Portuguese as “Cada participante” and “Cada oficial de ação”, respectively. However, BD3 generated the male-biased translations “todos os participantes” and “todos os agentes de ação”. Sentences 6, 9 and 15, in fact, showed improvement in BD3.

The prompt version in BD3 seems to have efficiently improved the sentences above. Be that as it may, as in the case of ChatGPT, we could not identify a pattern or provide any justification to explain why BD1 & 2 prompt versions generated gender-neutral outcomes but not BD3, which produced generic masculine structures despite being the most customised prompt, thus expected to yield better results.

It should also be noted that, out of the eight translation proposals, BD3 was the only version that provided a gender-neutral solution for sentence 9 (“A complainant who is not satisfied with the board's decision can ask for a rehearing”). This sentence proved to be a challenge for all the other versions, which used the generic masculine, whereas BD3 achieved the desired gender neutrality with the expression “uma pessoa que apresentar uma queixa”.

In fact, a closer look at all the gender-neutral solutions yielded by Google Bard reveals that, in the vast majority of cases, the system resorted to words in Portuguese which carry no gender marking, such as “pessoal”, “mão-de-obra”, “alguém” or “cada pessoa/oficial/participante”. One aspect of the results produced by Google Bard merits particular attention: the number of results in PT-BR dropped substantially from six (in BD1 & 2) to just one (in BD3).

#### **5.4. Summary**

A close analysis of the sets of sentences considered to have been translated in a gender-neutral style by the assessment team revealed that, out of the 38 sentences, only four met the GNT criteria in all eight translation versions (see Table 8 below).

Ref.	Source	Translation
4	E1	<i>Plans to outsource some 19 services have not proceeded at the anticipated pace, as there are significant manpower shortages.</i>
4	GT	Os planos para terceirizar cerca de 19 serviços não prosseguiram no ritmo previsto, pois há escassez significativa de mão de obra.
4	DL	Os planos de externalização de alguns serviços não avançaram ao ritmo previsto, uma vez que se regista uma escassez significativa de mão de obra.
4	GPT 1	Os planos para terceirizar cerca de 19 serviços não avançaram no ritmo previsto, devido a significativas escassezes de mão de obra.
4	GPT 2	Os planos de externalizar cerca de 19 serviços não avançaram no ritmo previsto, devido a significativas carências de recursos humanos.
4	GPT 3	Os planos de externalizar cerca de 19 serviços não avançaram no ritmo esperado, uma vez que existem importantes carências de recursos humanos.
4	BD 1	Os planos de terceirização de alguns 19 serviços não avançaram no ritmo previsto, pois há deficiências significativas de recursos humanos.
4	BD 2	Os planos de terceirização de alguns 19 serviços não seguiram o ritmo previsto, pois há uma escassez significativa de pessoal.
4	BD 3	Os planos de terceirização de alguns 19 serviços não têm avançado no ritmo esperado, devido a uma escassez significativa de mão de obra.
16	E3	<i>Someone may not know their tax number.</i>
16	GT	Alguém pode não saber seu número fiscal.
16	DL	Uma pessoa pode não saber o seu número de contribuinte.
16	GPT 1	Alguém pode não saber o seu número de contribuinte.
16	GPT 2	Alguém pode não saber o seu número de contribuinte.
16	GPT 3	Alguém pode não saber o seu número de contribuinte.
16	BD 1	Alguém pode não saber seu número de contribuinte.
16	BD 2	Alguém pode não saber seu número de contribuinte.
16	BD 3	Poderá haver pessoas que não conheçam o seu número de identificação fiscal.
25	E4	<i>A person must have resided continuously in the country for 5 years before applying for permanent residence.</i>
25	GT	Uma pessoa deve ter residido continuamente no país durante 5 anos antes de solicitar residência permanente.
25	DL	Para requerer a residência permanente, é necessário ter residido continuamente no país durante 5 anos.
25	GPT 1	Uma pessoa deve ter residido continuamente no país por 5 anos antes de solicitar residência permanente.
25	GPT 2	Uma pessoa deve ter residido continuamente no país durante 5 anos antes de solicitar residência permanente.
25	GPT 3	Uma pessoa deve ter residido continuamente no país por 5 anos antes de solicitar a residência permanente.
25	BD 1	Uma pessoa deve ter residido continuamente no país por 5 anos antes de solicitar a residência permanente.
25	BD 2	Uma pessoa deve residir continuamente no país por 5 anos antes de solicitar a residência permanente.
25	BD 3	Para se candidatar à residência permanente, é necessário residir continuamente no país por 5 anos.
28	E4	<i>Documents must be sent to the registry for distribution.</i>
28	GT	Os documentos devem ser enviados ao cartório para distribuição.
28	DL	Os documentos devem ser enviados ao registo para distribuição.
28	GPT 1	Os documentos devem ser enviados ao arquivo para distribuição.
28	GPT 2	Os documentos devem ser enviados para o registo para distribuição.
28	GPT 3	Os documentos devem ser enviados para o arquivo para distribuição.
28	BD 1	Documentos devem ser enviados ao registo para distribuição.
28	BD 2	Os documentos devem ser enviados ao registo para distribuição.
28	BD 3	Os documentos devem ser enviados ao registo para distribuição.

Table 8 - Sentences in GNT in the eight versions.

By closely examining the strategies (see Table 1<sup>21</sup>) used by the MT and AI systems to produce GNT (Table 9 below), we can see that a rather limited number of strategies were employed. As a matter of fact, except for one instance where the strategy used was repetition (F), and another in which the imperative form (H) was the structure of choice, for the most part, both MT systems and AI platforms resorted to strategies C, D and E (see Table 1). Strategy C clearly emerges as the most common one.

<sup>21</sup> Strategies identified in Table 1: A - Epicene synonyms; B - Pluralisation/Singularisation (towards generic or epicene forms); C - Relative and indefinite pronouns; D - Collective and role nouns; E - Omission; F - Repetition; G - Passive voice; H - Imperative forms; I - Impersonal forms.

	STRATEGY								
	A	B	C	D	E	F	G	H	I
<b>DeepL</b>	--	--	6	2	2	--	1	--	--
<b>GT</b>	--	--	6	2	1	--	1	--	--
<b>GPT</b>	--	--	19	6	--	--	3	1	--
<b>BARD</b>	--	--	12	6	1	1	3	--	--

Table 9 - Strategies used by MT and AI systems according to neutralisation strategies defined in Table 1.

Sentence 4 stands out as it is not gender-neutral in English due to the choice of the word "manpower", yet each of the eight results produced a GNT. On the other hand, for sentence 5, which is practically identical to 4, except for being indeed gender-neutral, the BD2 prompt version generated a generic masculine word ("funcionário"). This result is curious, considering the prompt used was exactly the same as for sentence 4.

Understanding the sentences that all systems failed to translate in a gender-neutral style is just as crucial as examining the successful cases. The list below shows the 16 sentences that none of the eight translation outcomes managed to translate in GNL. Since gender markers are widely present in various grammatical categories in PT-PT, translating these sentences with GNL would require substantial changes and more “creative” solutions, which the MT and AI systems seemingly could not carry out.

Ref.	Source	Sentence
2	E1	Professor Smith and Professor Jones will attend the luncheon.
3	E1	Guests are cordially invited to attend with their partners
10	E1	Substitute judges must certify that they have familiarized themselves with the record of the proceedings.
11	E1	The author of a communication must have direct and reliable evidence of the situation being described.
12	E2	An applicant may submit their complaint.
13	E2	A manager may apply for a transfer, provided that permission has been granted by that manager's institution
17	E3	Officials shall carry out their duties.
18	E3	An official's salary is dependent on length of service.
20	E4	All visitors must present their IDs/passes to obtain a temporary visitor's pass.
29	E4	The arrival sheet for newcomers must be completed by every new staff member by the end of the month.
33	E4	I will ask my assistant to make the necessary arrangements.
34	E4	The family allowance shall be granted to staff members whose spouse, within the meaning of the Personnel Regulations, has an overall income (gross income less compulsory social and/or pension contributions) lower than 50% of the basic monthly salary of the scale of the duty station of a Z1/1 grade staff member, plus the basic amount of the allowance.
35	E4	Eligibility shall commence when the staff member and their spouse have established a family unit at the duty station. It shall end when the family unit is dissolved or when the spouse ceases to actually and habitually live with the staff member at the duty station.
36	E4	The COMMITTEE considered that 1 September was not the best time to hold elections because a) it had been noted over the years that the Committee Secretary always had a very heavy workload at that time of year and would be unable to prepare the documents in time and b) many parents had to take leave that day to take their children to school on the first day of the school year;
37	E4	The current Committee Secretary confirmed that she would prefer to hold the elections at a later date that month;
38	E4	All judges must certify that they have familiarized themselves with the process.

Table 10 - Sentences translated with non-gender-neutral language in all translation versions.

## 6. Conclusion, limitations and further research

First of all, it should be noted that the limited number of sample sentences used in this case study reveals inconsistencies in the use of GNL across the outputs generated by the AI systems despite the use of customised prompts. What seems to stand out from the data is that the strategies used by the systems varied minimally and that they cannot handle structures which may require more elaborate reformulations to produce texts with GNL. Besides, we ought to consider that the outcomes emerged exclusively from source sentences already with GNL.

Even though some recent studies have shown that strong prompts can unlock greater potential from ChatGPT, this did not seem to apply when using gender-neutral language in translations from English into Portuguese.

Without customised prompts, ChatGPT and Google Bard demonstrated a level of quality similar to that of DeepL and Google Translate regarding the use of GNL; likewise, the results yielded by the customised prompts were not substantially different. The attempt at fine-tuning the prompts proved to be a challenge, and further investigation is required into how to best balance the amount of information, its specificity, and the examples provided. However, crafting entirely gender-neutral sentences in Portuguese often requires resorting to rather unnatural structures, which would not be a native speaker's first choice if they were not

asked to use a totally gender-neutral language. Ultimately, it is up to content producers, copywriters, translation clients and other stakeholders to decide when complete neutrality is necessary or appropriate or whether it is more adequate to resort to structures with the traditional use of the generic masculine. This case study does not advocate for either approach but rather aims at contributing to analysing the performance of MT and AI systems when, for whatever reason, a totally GNT is a mandatory requirement.

The findings from this case study have led to two additional research projects, which will be reported on in the near future. The first of these upcoming studies aims to understand how human translators approach gender-neutral language in the translations of the sentences which the MT and AI systems in this study failed to translate using GNL. Furthermore, the change from Google Bard to Gemini makes it relevant to research whether the results achieved by Gemini would differ.

To delve deeper into the subject of GNT, further investigations should focus on understanding whether post-editing of MT or whether ‘translating from scratch’ would produce better outcomes. In the future, it would be extremely useful to conduct similar studies testing the use of AI chatbots with customised prompts for translating different text types and genres. Additionally, it would be worthwhile to verify if different types of texts not originally written with GNL, such as those using the generic masculine, would result in different outcomes.

## References

- Bentivogli, L.; Savoldi, B.; Negri, M.; Di Gangi, M. A.; Cattoni, R.; Turchi, M. 2020. Gender in Danger? Evaluating Speech Translation Technology on the MuST-SHE Corpus. Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. Online: Association for Computational Linguistics, 6923–6933. <https://doi.org/10.18653/v1/2020.acl-main.619>.
- Cho, W. I.; Kim, J. W.; Kim, M.; Kim, N. S. 2019a. On Measuring Gender Bias in Translation of Gender-neutral Pronouns. <https://github.com/nolongerprejudice/tgbi>.
- Cho, W. I.; Kim, J. W.; Kim, S. M.; Kim, N. S. 2019b. On Measuring Gender Bias in Translation of Gender-neutral Pronouns. Proceedings of the First Workshop on Gender Bias in Natural Language Processing. Florence: Association for Computational Linguistics, 173–181. <https://doi.org/10.18653/v1/W19-3824>.
- Comunicação e Linguagem Inclusiva. 2018. IGEN - Fórum Organizações Para a Igualdade. [https://forumigen.cite.gov.pt/documentos/oferta-formativa-igen/t4-comunicacao-e-linguagem/t4\\_sara-falcao-casaca.pdf](https://forumigen.cite.gov.pt/documentos/oferta-formativa-igen/t4-comunicacao-e-linguagem/t4_sara-falcao-casaca.pdf).

- Dinu, G.; Mathur, P.; Federico, M.; Al-Onaizan, Y. 2019. Training Neural Machine Translation to Apply Terminology Constraints. Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. Florence: Association for Computational Linguistics, 3063–3068. <https://doi.org/10.18653/v1/P19-1294>.
- Friedman, B.; Nissenbaum, H. 1996. Bias in computer systems. *ACM Transactions on Information Systems*, 14(3): 330–347. <https://doi.org/10.1145/230538.230561>.
- Gao, Y.; Wang, R.; Hou, F. 2023. How to Design Translation Prompts for ChatGPT: An Empirical Study. <http://arxiv.org/abs/2304.02182>.
- Guidelines on gender-inclusive language. 2020. North Atlantic Treaty Organisation. [https://www.nato.int/nato\\_static\\_fl2014/assets/pictures/images\\_mfu/2021/5/pdf/210514-GIL-Manual\\_en.pdf](https://www.nato.int/nato_static_fl2014/assets/pictures/images_mfu/2021/5/pdf/210514-GIL-Manual_en.pdf).
- Guidelines on gender-inclusive language in English. 2018. United Nations. <https://www.un.org/en/gender-inclusive-language/guidelines.shtml>.
- Inclusive communication in the GSC. 2018. General Secretariat, Council of the European Union. [https://www.consilium.europa.eu/media/35446/en\\_brochure-inclusive-communication-in-the-gsc.pdf](https://www.consilium.europa.eu/media/35446/en_brochure-inclusive-communication-in-the-gsc.pdf).
- Jiao, W.; Wang, W.; Huang, J.; Wang, X.; Shi, S.; Tu, Z. 2023. Is ChatGPT a good translator? Yes, with GPT-4 as the engine. arXiv. <https://doi.org/10.48550/arXiv.2301.08745>
- Kocmi, T.; Limisiewicz, T.; Stanovsky, G. 2020. Gender Coreference and Bias Evaluation at WMT 2020. Proceedings of the Fifth Conference on Machine Translation. Online: Association for Computational Linguistics, 357–364.
- Kostikova, A.; Daems, J.; Lazarov, T. 2023. How adaptive is adaptive machine translation, really? A gender-neutral language use case. Proceedings of the First Workshop on Gender-Inclusive Translation Technologies. Tilburg: Open Press TiU, 95-97.
- Manual de linguagem inclusiva. 2021. Comissão para a Cidadania e a Igualdade de Género (CIG). Portugal. Conselho Económico e Social. <https://www.cig.gov.pt/wp-content/uploads/2021/08/12-Manual-de-Linguagem-Inclusiva-CES.pdf>.
- Nord, C. 1997. Defining translation functions. The translation brief as a guideline for the trainee translation. *Ilha Do Desterro* 33: 41–55.
- Nord, C. 2018. *Translating as a Purposeful Activity* (2nd Edition). Routledge. <https://doi.org/10.4324/9781351189354>.
- Parlamento Europeu. 2018. Linguagem neutra do ponto de vista do género no Parlamento Europeu. Retrieved from [https://www.europarl.europa.eu/cmsdata/288151/GNL\\_Guidelines\\_PT-original.pdf](https://www.europarl.europa.eu/cmsdata/288151/GNL_Guidelines_PT-original.pdf).
- Piergentili, A.; Fucci, D.; Savoldi, B.; Bentivogli, L.; Negri, M. 2023. Gender Neutralization for an Inclusive Machine Translation: from Theoretical Foundations to Open Challenges. Proceedings of the First Workshop on Gender-Inclusive Translation Technologies. Tampere: European Association for Machine Translation, 71–83. <http://arxiv.org/abs/2301.10075>.
- Prates, M. O. R.; Avelar, P. H.; Lamb, L. C. 2020. Assessing gender bias in machine translation: a case study with Google Translate. *Neural Computing and Applications* 32(10): 6363–6381. <https://doi.org/10.1007/s00521-019-04144-6>.
- Pym, A. 1993. Christiane Nord. *Text Analysis in Translation. Theory, Method, and Didactic Application of a Model for Translation-Oriented Text Analysis*. Translated from the German by Christiane Nord and Penelope Sparrow. Amsterdam/Atlanta GA, Rodopi, 1991, 250 p. ISBN: 90-5183-311-3. *TTR* 6(2): 184-190. <https://doi.org/10.7202/037160ar>.

- Reiss, K.; Vermeer, H. J. 2014. *Towards a General Theory of Translational Action*. Routledge. <https://doi.org/10.4324/9781315759715>.
- Rescigno, A.; Vanmassenhove, E.; Monti, J.; Way, A. 2020. A Case Study of Natural Gender Phenomena in Translation. A Comparison of Google Translate, Bing Microsoft Translator and DeepL for English to Italian, French and Spanish. In *Proceedings of the Seventh Italian Conference on Computational Linguistics CLiC-it 2020* (359–364). Accademia University Press. <https://doi.org/10.4000/books.aaccademia.8844>.
- Saunders, D.; Byrne, B. 2020. Reducing Gender Bias in Neural Machine Translation as a Domain Adaptation Problem. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Online: Florence: Association for Computational Linguistics, 7724–7736. <https://doi.org/10.18653/v1/2020.acl-main.690>.
- Savoldi, B.; Gaido, M.; Bentivogli, L.; Negri, M.; Turchi, M. 2021. Gender Bias in Machine Translation. *Transactions of the Association for Computational Linguistics* 9: 845–874. [https://doi.org/10.1162/tacl\\_a\\_00401](https://doi.org/10.1162/tacl_a_00401).
- Stanovsky, G.; Smith, N. A.; Zettlemoyer, L. 2019. Evaluating Gender Bias in Machine Translation. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. Florence: Association for Computational Linguistics, 1679–1684. <https://doi.org/10.18653/v1/P19-1164>.
- Sun, T.; Gaut, A.; Tang, S.; Huang, Y.; ElSherief, M.; Zhao, J.; Mirza, D.; Belding, E.; Chang, K.-W.; Wang, W. Y. 2019. Mitigating Gender Bias in Natural Language Processing: Literature Review. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. Florence: Association for Computational Linguistics, 1630–1640. <https://doi.org/10.18653/v1/P19-1159>.
- Sun, T.; Webster, K.; Shah, A.; Wang, W. Y.; Johnson, M. 2021. They, Them, Theirs: Rewriting with Gender-Neutral English. <http://arxiv.org/abs/2102.06788>.
- Vanmassenhove, E.; Hardmeier, C.; Way, A. 2019. Getting Gender Right in Neural Machine Translation. *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*. Brussels: Association for Computational Linguistics, 3003–3008. <https://doi.org/10.18653/v1/D18-1334>.
- Vanmassenhove, E.; Emmerly, C.; Shterionov, D. 2021. NeuTral Rewriter: A Rule-Based and Neural Approach to Automatic Rewriting into Gender Neutral Alternatives. *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*. Online and Punta Cana: Association for Computational Linguistics, 8940–8948. <https://doi.org/10.18653/v1/2021.emnlp-main.704>.
- Yamada, M. 2023. Optimizing Machine Translation through Prompt Engineering: An Investigation into ChatGPT’s Customizability. *Proceedings of Machine Translation Summit XIX, Vol. 2: Users Track*. Macau: Asia-Pacific Association for Machine Translation, 195–204. <http://arxiv.org/abs/2308.01391>.

## LANGUAGE ASSESSMENT IN THE NORMALIZING SOCIETY

Diego Fernando de Oliveira<sup>22</sup>

[diego.fernando@unesp.br](mailto:diego.fernando@unesp.br)

SÃO PAULO STATE UNIVERSITY (BRAZIL)

**ABSTRACT.** The aim of this study is to discuss linguistic norms present in the *Common European Framework*, particularly its criteria for assessing phonetic and phonological features in oral performances. For this purpose, we employed Performance Decision Trees to demonstrate that its operational definition for oral proficiency punishes the presence of foreignness in linguistic performances, even when segmental features of the speech do not affect candidates' performances in linguistic tasks. As a result, we observed that the *Common European Framework* progressively penalizes the presence of phonetic and phonological elements associated with other languages spoken by candidates.

**KEYWORDS.** Language Assessment, Common European Framework, Pronunciation, Phonological Control.

**RESUMO.** Este estudo tem como objetivo discutir as normas linguísticas presentes no *Quadro Comum Europeu*, atribuindo especial atenção aos seus critérios para a avaliação de aspectos fonéticos e fonológicos em desempenhos orais. Para isso, empregamos Árvores de Decisão de Desempenho a fim de demonstrar que a sua definição operacional de proficiência prejudica a presença da estrangeiridade nos desempenhos orais, mesmo quando esta não afeta o desempenho proficiente dos sujeitos em tarefas linguísticas. Como resultado, observamos que o *Quadro Comum Europeu* pune progressivamente a presença de elementos de natureza fonética e fonológica associados a outras línguas faladas por candidatos.

**PALAVRAS-CHAVE.** Avaliação de Línguas, Quadro Comum Europeu, Pronúncia, Controle Fonológico.

---

<sup>22</sup> Second-year doctoral student in the Postgraduate Program in Language Studies at São Paulo State University (Unesp), Brazil.

## 1. Introduction

*We are entering the age of the infinite examination and of compulsory objectification.*

(Foucault 1995: 189)

Applied Linguistics' (AL) literature traditionally defines language assessment as a technical process that aims to collect data and measure evidence about a subject's ability to perform language tasks that simulate real-life contexts (Segat & Sarmiento 2022). Another common aspect in specialized literature is the long-standing taxonomy that classifies assessment as formative and summative. Formative assessment focuses on local contexts and pedagogical purposes, while summative assessment has as its main objective the certification of skills and knowledge at the end of a given school term. There is also standardized assessment, which specialists define as a process that seeks to gauge the level of proficiency of candidates, usually associated with international mobility settings (Fulcher 2010).

Despite its conservative perspective, specialized literature has raised debates about the social scope of assessment practices, paying special attention to issues related to the fairness and validity of assessment processes, as well as their impact on promoting social justice (Fulcher & Davidson 2007; Fulcher 2010, 2015; Fulcher & Harding 2021). Thus, assessment studies raise debates that challenge the long-standing taxonomies that confine assessment processes to institutional practices, demonstrating the social and political scope of contemporary assessment practices.

In this regard, McNamara (2011) and Harding & Mcnamara (2018) acknowledge the fact that language assessment praxis in the 21st century demands reflection. According to McNamara (2011: 500),

Language assessment, as we are increasingly realizing, is first and foremost a social and political activity [...]. It is a site where social values are expressed and contested. As the values of a globalized economy have come to occupy and preoccupy many spheres of social and political life, so have they become central to current developments and challenges in language testing and assessment.

Regardless of raising important questions about language assessment processes, literature is often limited in its sociological perspective related to the social impacts of assessment. Assessing languages is also assessing identities and cultures, since these are inseparable elements (Darvin & Norton 2015, 2017). In a world of increasingly liquid boundaries but progressively more controlled spaces, linguists must comprehend language

assessment as a social dynamic. In this respect, this paper aims to demonstrate that institutions can potentially employ language assessment as an instrument of normalization: not only normalization of linguistic performances, but also of linguistic identities.

In view of the context in which this study is set, as well as the precedents in AL that are concerned with discussing the social impacts of language assessment, this study aims to discuss linguistic norms present in the *Common European Framework* (CEFR) companion volume (Council of Europe 2020), specifically to its phonetic and phonological criteria. For this purpose, we employed Performance Decision Trees (PDTs) to demonstrate that its operational definition of language proficiency punishes the presence of foreignness in oral performances, even when it does not affect candidates' proficient performances in linguistic tasks.

As far as the structure of this study is concerned, we have structured our discussions into five main sections: the first section contextualizes our object of analysis, the assessment criteria for oral performances in the CEFR. The second section briefly discusses the intersection between phonetics, phonology and assessment. The third section introduces Foucault's sociological ideas about normalization. The fourth section deals with the research methodology adopted, the PDTs. Finally, the fifth section presents the results, discussing them in the light of the discussions raised in the theoretical sections.

## 2. The CEFR as a standard for language assessment

The CEFR (Council of Europe 2020) is one of the most successful policies of the Council of Europe, setting standards for language assessment in different parts of the world. According to the Council of Europe, the CEFR

[...] is intended to promote quality plurilingual education, facilitate greater social mobility and stimulate reflection and exchange between language professionals for curriculum development and in teacher education. Furthermore the CEFR provides a metalanguage for discussing the complexity of language proficiency for all citizens in a multilingual and intercultural Europe, and for education policy makers to reflect on learning objectives and outcomes that should be coherent and transparent (Council of Europe 2020: 11)

As can be seen, the Council of Europe emphasizes the role of the CEFR as a language policy open to plurilingualism, multilingualism, and interculturalism. There are several passages in the referred document that portray the CEFR as an instrument that reflects the

plurality of languages, “[...] including the promotion of reflective learning and learner autonomy” (Council of Europe 2020: 13). As such, the CEFR is presented in this document as an assessment tool that is open to diversity, a political statement that meets the demands of decolonial studies in AL.

Taking into account this concern with progressive debates in the field of foreign language teaching and learning, there is also an attempt to construct a transparent image of the linguistic and educational policies described in the document. According to the Council of Europe (2020: 27)

The CEFR was developed as a continuation of the Council of Europe’s work in language education during the 1970s and 1980s. The CEFR “action-oriented approach” builds on and goes beyond the communicative approach proposed in the mid-1970s in the publication “The Threshold Level”, the first functional/notional specification of language needs. The CEFR and the related European Language Portfolio (ELP) that accompanied it were recommended by an intergovernmental symposium held in Switzerland in 1991. As its subtitle suggests, the CEFR is concerned principally with learning and teaching. It aims to facilitate transparency and coherence between the curriculum, teaching and assessment within an institution and transparency and coherence between institutions, educational sectors, regions and countries.

Furthermore, the Council of Europe (2020: 28) states that its main aim is “[...] facilitating quality in language education and promoting a Europe of open-minded plurilingual citizens” and it “[...] brings a new, empowering vision of the learner”. Although statements of this nature proliferate throughout the companion volume, several authors present questions regarding different aspects of the CEFR. The absence of clarity in their descriptors, the presence of culture-specific bias and lack of attention to contexts and purposes of language use and assessment are examples of demands addressed by the CEFR companion volume (Council of Europe 2020).

In addition, various studies highlight challenges in the application of the CEFR, particularly in social contexts that diverge significantly from the European linguistic milieu (Çagatay & Gurocak 2016; Mohammed, Raof & Yusof 2021; Zaki & Darmi 2021, among others). We attribute this to the fact that the CEFR not only serves as an operational framework for the language construct but also embodies the language policy of European social institutions (Mcnamara 2011), which causes dissonance with the realities of South American, African and Asian societies, for example.

Regarding the implications of using the language policies of the CEFR, McNamara (2011: 506-507) states that

The imposition of a single set of cultural meanings and social and political values for language education, for each setting in which the CEFR is adopted, eviscerates the traditions of language teaching which are incompatible with the CEFR. In cultural and historical terms, learning English is simply not the same for a Singaporean, an Indonesian, a Vietnamese, a French person, a Dutch person, or a Hungarian. And different languages are indeed different, and vary in socio-political role, range and purpose of use, and carry very different histories of contact, often violent, between speakers of the speech communities involved. The CEFR puts all foreign languages into one and the same category, thereby erasing the fact that they are ‘foreign’ in very different ways. Learners from those backgrounds may be aware of the cultural and political significance of the act of learning the language of a cultural group with a particular historical relationship to their own, unless they are invited, in an act of collective historical amnesia, to wipe the slate clean and re-identify themselves as citizens of the new globalized world.

As can be seen, despite the position of the CEFR as an instrument developed from a perspective that approaches the diversity of languages in its theoretical scope, its implementation entails a number of problems, especially in societies that are culturally distant from Europe. Language assessment policies such as the CEFR affect different dimensions of people’s lives, which has consequences for access to education, work and social mobility itself (Spolsky 1995; Davies 2008; Mcnamara 2011).

Last, but certainly not least, in introducing the scale, the Council of Europe (2020) points to improvements in the clarity of criteria compared to the 2001 original version of their framework. To this end, regarding pronunciation assessment criteria, the Council of Europe (2020) states that it consulted AL literature and language experts to operationalize language proficiency with greater scientific validation, contemplating the following empirically verifiable aspects of its framework:

Articulation, including pronunciation of sounds/phonemes; prosody, including intonation, rhythm and stress – both word stress and sentence stress – and speech rate/chunking; accentedness, accent and deviation from a “norm”; intelligibility, accessibility of meaning for interlocutors, covering also the interlocutors’ perceived difficulty in understanding (normally referred to as “comprehensibility” (Council of Europe 2020: 133).

In this sense, there is a concern to find scientific evidence to support the standardizing role of the CEFR's language policies. The Council of Europe (2020) further emphasizes the way in which it conceives the elements operationalized on its framework:

The focus is on familiarity and confidence with the target language sounds (the range of sounds a speaker can articulate and with what degree of precision). The key concept operationalised in the scale is the degree of clarity and precision in the articulation of sounds. The focus is on the ability to effectively use prosodic features to convey meaning in an increasingly precise manner (Council of Europe 2020: 133).

What becomes evident is the presence (almost imperceptible) of a norm that supports all the criteria mentioned by the CEFR. In addition, the norm for phonological control refers precisely to the assessors' socially constructed perspective towards the construct of language proficiency. Concerning the assessor's influence over the assessment process, the literature points out that rater training would be necessary to provide more reliability in the results of the assessment process (Fulcher 2010), which does not mean that this strategy is truly effective. However, behind the scientificism of language assessment dynamics, there is the noise of human subjectivity that assessment institutions and instruments cannot erase.

Considering the political, educational and linguistic aspects described by the Council of Europe (2020) in relation to the CEFR, the next section explores the criteria for assessing phonological aspects in the AL literature.

### **3. Phonology and language assessment and the problem of the nativeness**

In order to assessment take place, assessment institutions operationalize language according to models developed by specialists (Fulcher 2010). Such models of language proficiency are constantly being expanded in order to translate proficiency accurately into observable elements in oral performance. Once materialized in criteria, these observable aspects compose descriptors in proficiency scales. In this manner, assessment contexts and purposes determine the operationalization of proficiency, since the results usually guide assessors and institutions to make decisions based on the candidates' performances.

Phonological control is an assessment category that composes a broader field, more commonly referred to in the literature as pronunciation (Fulcher 2010). In order to explain what we are referring to, we have adopted Fulcher's definition of pronunciation, according to which “[...] the outer manifestation of speech is sound. The speaker must first decide what to say, be able to articulate the words, and create the physical sounds that carry meaning” (Fulcher 2010: 25). However, regarding pronunciation assessment practices, Fulcher (2010: 25) states that:

When we design speaking tests we must decide whether assessing pronunciation at this level is relevant at all to the situation. This depends upon the test purpose. Pronouncing words in the way they would be pronounced in the standard variety of a language may be important to a newsreader, but for most learners testing pronunciation may only be a matter of general intelligibility.

As a result, in cases where pronunciation is a relevant language feature for decision-making, the assessment of oral performances takes into account segmental and suprasegmental dimensions of speech (Fulcher 2010). Segmental elements of oral performance refer to vowels, consonants, and their respective singularities. The suprasegmental elements refer to word stress, intonation, tone, rhythm, among other aspects (O'Brien 2021). Such operationalization can be observable in the CEFR criteria for assessing phonological control - and, in this paper, we focus on its hierarchical organization.

Another relevant aspect for assessing pronunciation is intelligibility. According to Kennedy & Trofimovich (2008) and Thomson (2018), intelligibility is the extent to which the listener understands a sentence. Julkowska & Cebrian (2015) conceive intelligibility as the degree to which a speaker's output is actually understood by the listener. Thomson (2018) also points out that in recent works related to language assessment, authors tend to measure intelligibility from segmental aspects rather than at the utterance level. According to Bundgaard-Nielsen, Best, Kroos & Tyler (2012) and Gooch, Saito & Lyster (2016), a segment within a word is intelligible if it is identified by the listener as belonging to the target language system.

The theoretical criteria which references segmental and suprasegmental aspects of speech recognizable to the interlocutor as a part of the "phonological system" of a given foreign language evidence a formalist perspective of language, still considerably dominant in different fields of study. Such a perspective is based on the linguistic fallacy of the existence of a language as a homogeneous system spoken by an idealized native speaker (Camargo 2016), which includes phonological features in oral performances.

In this respect, according to Pennycook (2017), the perspective of a language as a people's heritage owned by native speakers relates to the formation of European states in the Renaissance period. Language, in this sense, is homogeneous, invariable, governed by rules to be followed and not questioned, which must be learned in order to unify a national identity. In this context, Machado, Delfino & Rodrigues (2021: 117, our translation) argue that

conceiving language as a plastic, malleable structure, dependent on actual speakers and real speakers and socio-cultural interactions, which presents regional, social, age, gender, etc. linguistic variations and gender, even though there are many lexical, phonetic and grammatical elements that are stable and common to the varieties of a language, implies questioning colonialist epistemology, valuing plurality and different ways of being, know and produce knowledge of the subjects.

In the case of the CEFR criteria for phonological control, assessing pronunciation in terms of intelligibility puts the performance of the assessed at the mercy of an idealized social norm, often not well described or defined. Assessors, based on their perspective of language and pronunciation, classify the assessed performances as satisfactory or unsatisfactory: if satisfactory, the assessor crowns the norm that they define as the standard variety of the foreign language assessed; if unsatisfactory, the assessor penalizes candidates based on particular views related to ideal speakers.

In view of the discussions raised in this section about the assessment of phonological aspects and the problem of nativeness, the next section explores concepts from sociology that are important for understanding our perspective on assessment as a normalizing dynamic.

#### **4. Norms, normalization and the examination**

In this section, we explore sociological concepts that are important for understanding the discussions we are proposing in relation to our conception of assessment as a normalizing process. Firstly, it is important to establish notions of norms and normalization, which guide the objective of this study. Despite not providing a single definition for his concept of norms, we base ourselves on the ideas developed by Foucault (1987) interpretations of the norm as an ideal social model, taking as a reference the study by Kelly (2019), a specialist in Foucauldian studies.

In this way, we consider norms and normalization to be “[...] a model of perfection that operates as a guide to action in any particular sphere of human activity, and normalization correlatively as the movement by which people are brought under these norms” (Kelly 2019: 2). Based on this, the construct of language proficiency operationalized by policies and assessment instruments represents the norm and normalization, in turn, deals with the process in which candidates subject themselves to language assessment criteria.

In addition, Foucault states that normalization consists

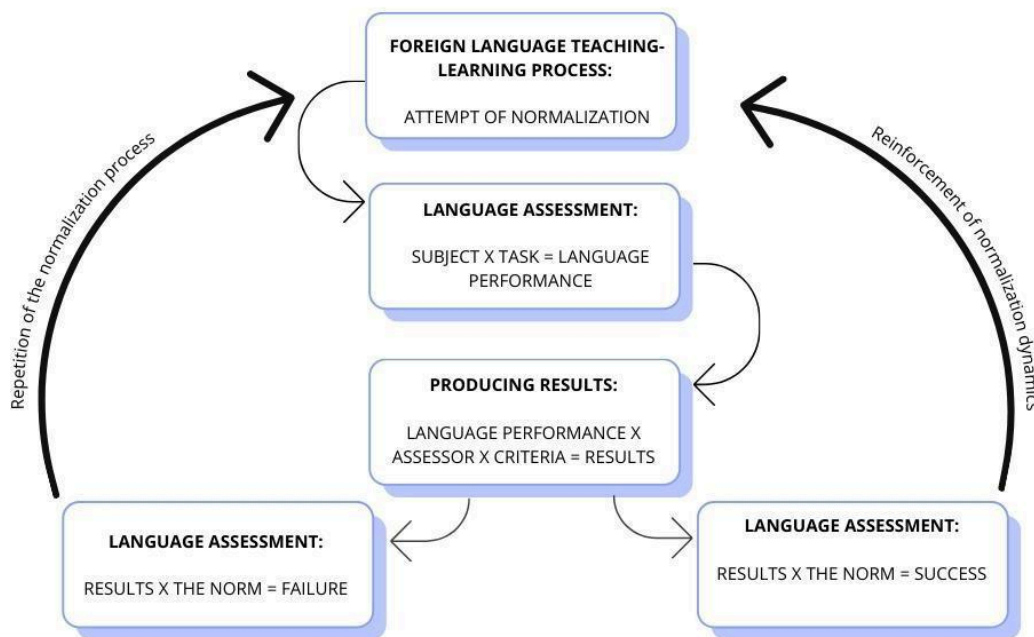
first of all in positing a model, an optimal model that is constructed in terms of a certain result, and the operation of disciplinary normalization consists in trying to get people, movements, and actions to conform to this model, the normal being precisely that which can conform to this norm, and the abnormal that which is incapable of conforming to the norm. In other words, it is not the normal and the abnormal that is fundamental and primary in disciplinary normalization, it is the norm. That is, there is an originally prescriptive character of the norm and the determination and the identification of the normal and the abnormal becomes possible in relation to this posited norm (Foucault 2007: 58).

In his writings, Foucault (1987) further discusses the normalization of subjects through school pedagogical action, referring to what he calls the examination. According to Foucault,

The examination combines the techniques of an observing hierarchy and those of a normalizing judgment. It is a normalizing gaze, a surveillance that makes it possible to qualify, to classify and to punish. It establishes over individuals a visibility through which one differentiates them and judges them. That is why, in all the mechanisms of discipline, the examination is highly ritualized. In it are combined the ceremony of power and the form of the experiment, the deployment of force and the establishment of truth. At the heart of the procedures of discipline, it manifests the subjection of those who are perceived as objects and the objectification of those who are subjected. The superimposition of the power relations and knowledge relations assumes in the examination all its visible brilliance (Foucault 1987: 184).

Based on Foucault (1987, 2007) and Kelly (2019), we interpret the CEFR's linguistic and assessment policies as a series of norms aimed at describing the ideals of linguistic performance to be achieved by candidates who undergo assessment processes. Assessment processes, in turn, categorize, classify and hierarchize subjects according to criteria that punish the presence of linguistic elements considered abnormal within certain idealized standards of foreign languages. We therefore illustrate our conception of assessment as a normalization process in Figure 1 below.

Figure 1 - Language assessment process.



Source: Elaborated by the author.

As can be observed, exam assessment processes based on frameworks such as the CEFR criteria propose linguistic tasks to their candidates in order to collect evidence of their linguistic proficiency — or what we call their linguistic normalization. As a result, subjects are classified as more or less proficient hierarchically: those who succeed in this process have access to socioeconomic rights and privileges and those who fail are condemned to a continuous process of normalization, left vulnerable to arbitrary domestic power.

In short, assessment as a normalizing process makes clear one of the biggest incongruities in AL literature: if the candidate performs a task proficiently and achieves its objectives, why should institutions punish their performance due to the presence of other spoken languages in their performances? The very technical concept of proficiency, as explained above, brings to light assessment as a ritual for punishing identities and, consequently, maintaining social privileges.

Taking into account Foucault's concepts of norms and normalization, the next section presents the methodological path that guided the analysis of the criteria for assessing phonological control in the CEFR.

## 5. Methodological procedures

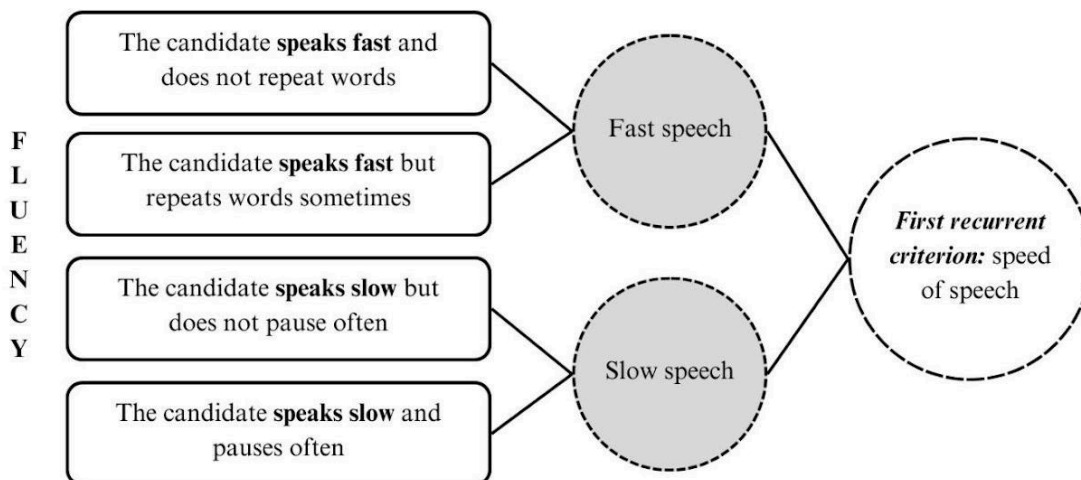
In order to analyze the assessment criteria present in the CEFR scale for phonological control, this paper presents an original adaptation of Fulcher, Davidson & Kemp's (2011) methodology for selecting assessment criteria and developing rating scales referred as PDTs or Empirically derived, Binary Choice and Boundary Definition scales (EBBs). According to Fulcher, Davidson & Kemp (2011: 9), their methodology

[...] is set forth as a series of repeated and branching binary decisions. EBBs are constructed by rank ordering performances on test tasks and then identifying key features that judges use to separate the performances into adjacent levels. EBBs represent an innovation in the logic of how raters judge performance with reference to performance data in specific contexts of language use. EBBs [...] are relatively easy to use in real-time rating, and do not place a heavy burden on the memory of the raters.

As can be seen, Fulcher, Davidson & Kemp (2011) define their methodology as qualitative-interpretative, empirically based, since exam developers consider oral performance data and the expertise of assessors and/or experts for the selection and ranking of assessment criteria. Although PDTs are empirical methodologies for developing criteria and constructing rating scales, they can also be useful in deconstructing the same scales by isolating the constitutive elements of the rating criteria present in their descriptors.

In this context, we began the analysis process with the proficiency scale of interest. The first procedure is to look at the recurring criteria between descriptors in order to identify the criteria that limit the classifications of satisfactory and unsatisfactory candidate performances. Figure 2 below illustrates the first analytical step, which we have adopted in this study, exemplified by the analysis of a hypothetical scale for assessing fluency.

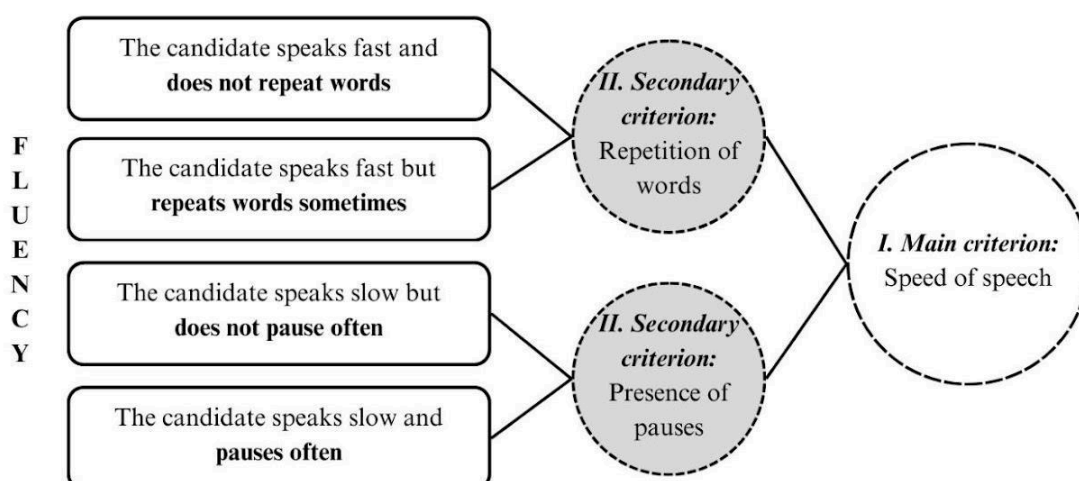
Figure 2 – Step 1 for isolating criteria.



Source: Elaborated by the author.

As illustrated in Figure 2, considering a four-level proficiency scale for assessing fluency, we isolate the criterion that limits satisfactory and unsatisfactory performances - the speed of speech. Likewise, we continued the analysis by isolating the criteria that limit the two highest and lowest ranges of the scale, in order to identify the criterion that determines the limits between them. Figure 3 below illustrates this process.

Figure 3 – Step 2 for isolating criteria.



Source: Elaborated by the author.

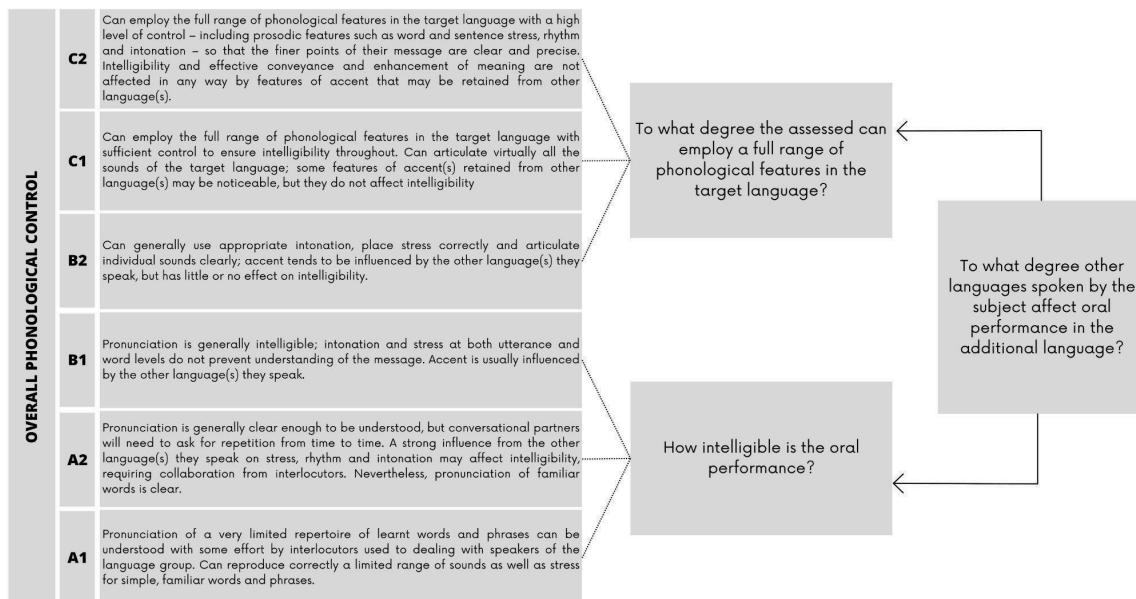
As illustrated above, by analyzing the recurrence and hierarchization of criteria, it is possible to reconstruct the operationalization of the fluency construct on this hypothetical scale. In this instrument, the speed of speech is what determines the success or failure of the candidate in the assessment process, while the repetition of words limits the more proficient levels and pauses limit the less proficient performances. The number of branches of the PDTs thus depends on the number of levels presented in the proficiency scale to be analyzed and researchers' interpretation and expertise analyzing descriptors are fundamental for establishing hierarchical relationships between criteria.

Given the methodological procedures employed in this paper in order to analyze rating scales, the next section presents results of the application of PDTs in the CEFR's phonological control assessment scale, as well as fosters discussions about the normalizing potential of its operational definition.

## **6. Analysis and discussion**

The CEFR's phonological control assessment scale comprehends six bands, which describe levels comprising A1 as the least proficient performance and C2 as the most proficient performance. As described in the methodology section of this paper, initially, recurrent criteria that divide performances according to their proficiency level were isolated. Then, we conducted the same process to identify the criteria that limit levels A1, A2, B1 and B2, C1 and C2, as Figure 4 illustrates below.

Figure 4 - Criteria for overall phonological control from CEFR.



Source: Elaborated by the author.

As illustrated in Figure 4, the main criterion for rating oral performances is the intrusion of “language features” retained from other languages into the pronunciation of the target language. Next, at the three highest levels of the scale, segmental and suprasegmental aspects of the target language's phonological system is what distinguishes higher proficiency levels. Regarding the three lowest levels, intelligibility of oral performances sets limits that differentiate the three most basic ranges. With regard to the criteria extracted from the scale, we observed that both the CEFR and the specialized literature validate the PDT with regard to the operational definition of phonological control and the hierarchization of its assessment criteria. Interestingly, the Council of Europe (2020) provides its own definition of intelligibility, centering it on the figure of the interlocutor who, in this case, is represented by the assessor: “intelligibility, how much effort is required from the interlocutor to decode the speaker’s message; the extent of influence from other languages spoken; control of sounds [...]” (Council of Europe 2020: 133).

Additionally, intelligibility is associated with the production of segmental elements - that is, the assessor's perception of what characterizes an accent that is easy or difficult to understand, what is referred to in the literature as the listener-centered intelligibility issue (Fulcher 2010; Thomson 2018; among others). The role of phonetic segments in performance

classification is distant from the progressive proposal of the Council of Europe (2020) itself on language and assessment policies.

In this way, penalizing variation in the pronunciation of segments based on the perception of intelligibility centered on the assessor, without providing a clear definition of standards to be achieved or reproduced by the candidates, creates conditions for the potential promotion of social injustices. Based on this, the presence of elements retained from other languages in oral performances can arbitrarily determine candidates' access to job and education opportunities, as well as international mobility, which for us is a clear parallel between Foucault's (1987, 2007) concept of norms as unattainable imposed ideals for social control.

In view of this scenario, table 1 below presents important issues we raise in relation to the hierarchization of the CEFR's assessment criteria for phonological control.

Table 1 - Operationalization issues in the CEFR scale for overall phonological control.

Band	Excerpts of descriptors	Operationalization issue
C2	Intelligibility and effective conveyance and enhancement of meaning <b>are not affected in any way by features of accent that may be retained from other language(s).</b>	Segmental and suprasegmental retained from other languages are not apparent to the assessor in oral performances. The absence of elements that evidences the candidate's foreign origin guarantees their higher ranking, which depicts a view of language as a homogeneous phonological structure.
C1, B2	Can articulate virtually all the sounds of the target language; <b>some features of accent(s) retained from other language(s) may be noticeable, but they do not affect intelligibility.</b>  <b>Accent tends to be influenced by the other language(s) they speak, but has little or no effect on intelligibility.</b>	Although intelligibility is not affected, the subjective perception of an "accent" possibly retained from other languages downgrades the candidate's performance to a lower satisfactory position. The "intrusion" of other languages as a descriptor of the second lowest level of the scale illustrates that candidates classified as B2 demonstrate proficiency performing a given language task, but their foreign background is still noticeable.
B1	<b>Accent is usually influenced by the other language(s) they speak.</b>	Unlike the previous cases, intrusion of elements retained from other languages spoken by candidates is evident. This aspect thus defines a boundary between satisfactory and unsatisfactory performances. Norms, in this case, marginalize the acceptable performance to the lower level of proficiency bands.
A2, A1	<b>A strong influence from the other language(s) they speak on stress, rhythm and intonation may affect intelligibility, requiring collaboration from interlocutors.</b> Nevertheless, <b>pronunciation of familiar words is clear.</b>  <b>Can reproduce correctly a limited range of sounds as well as stress for simple, familiar words and phrases.</b>	In the lower bands of the scale, candidates' performance has a clear influence of other languages and it becomes evident to the assessor (who can even "correct" the assessed as a form of "collaboration"). Candidates' ability to reproduce sounds considered to belong to ideal norms still distinguish lower proficiency bands, which evidences that the greater the presence of foreignness in linguistic performances, the greater the potential penalty suffered by candidates.

Source: Elaborated by the author.

As can be observed, despite Council of Europe propaganda (2020), what is observable in the CEFR descriptors is a very distinct reality: the "intrusion" of elements retained from other languages, even if they do not affect the candidate's proficiency in performing language tasks, is enough reason for penalties. In this respect, what is evident is the operational definition of language proficiency based on formalist, homogenous, purist and ideal models (Oliveira 2014, Camargo 2016). Language assessment, in this sense, can potentially assume normalizing traits and, in broader political scenarios, can hinder the promotion of fair conditions for social, economic and educational opportunities.

In conclusion, the PDTs and the hierarchization of segmental and suprasegmental aspects in the description of the different levels of proficiency demonstrates that the CEFR and the Council of Europe's (2020) language policies can be employed as normalizing dynamics, possibly determining the arbitrary penalization of linguistic varieties and their speakers. In addition, descriptors that define the most and least proficient performances can serve as arguments justifying the promotion of social injustices and restrict foreigners' access to basic human rights, such as international mobility.

## 7. Conclusions

In this paper, we examined assessment criteria contained in the CEFR, developed by the Council of Europe (2020), for assessing phonological control in oral performances in foreign languages. We observed the presence of two main linguistic aspects for classifying candidates: the production of segmental elements, like vowels and consonants, as well as the employment of suprasegmental elements, like word stress, intonation, tone, and rhythm. From the development of the PDTs, we noticed that the CEFR hierarchically organizes segmental criteria related to the presence of linguistic aspects retained from other languages spoken by the candidate in such a way as to progressively punish the speaker's foreign accent.

The progressive punishment of the presence of segmental aspects related to other languages spoken by the candidate highlights the potential normalizing dynamic of assessment processes, since it promotes an ideal of linguistic homogeneity and arbitrarily punishes varieties that do not belong to ideal norms. Furthermore, the establishment of ideals of nativeness for the classification of candidates enables assessors and educational institutions to promote social injustices, thus hindering foreigners' access to international mobility.

We hope to stimulate the interest of applied linguists who investigate language assessment about its social impact, since assessment results guide decision-making that not only affects candidates' daily lives, but also their access to job opportunities, education and, in some cases, even survival. In addition, we expect the discussions about assessment as a normalization process to set precedents for the development of studies that are concerned with other assessment instruments and policies with global impact.

Finally, this study is limited in its scope, since we dedicated our analysis to identify potential operational issues in the phonological control rating scale developed by the Council

of Europe (2020), thus not reflecting the breadth and complexity of the CEFR operationalization of language proficiency entirely.

## References

- Bundgaard-Nielsen, R. L.; Best, C. T.; Kroos, C.; Tyler, M. D. 2012. Second language learners' vocabulary expansion is associated with improved second language vowel intelligibility. *Applied Psycholinguistics* 33(3): 643-664.
- Çagatay, S.; Gürocak, F. U. 2016. Is CEFR Really Over There? *Procedia - Social and Behavioral Sciences* 232: 705-712.
- Camargo, H. R. E. 2016. Uma Língua Inglesa para chamar de minha: equívocos sobre o bom falante de inglês. *Revista Estudos Linguísticos* 45(2): 651-665.
- Council of Europe. 2020. *Common European Framework of Reference for Languages: Learning, teaching and assessment*. Strasbourg, France: Council of Europe Publishing.
- Darvin, R.; Norton, B. 2015. *Identity and a model of investment in Applied Linguistics*. Cambridge: Cambridge University Press.
- Darvin, R.; Norton, B. 2017. Language, Identity, and Investment in the Twenty-first Century. In *Language Policy and Political Issues in Education (227-240)*. *Encyclopedia of Language and Education*. Springer International Publishing.
- Davies, A. 2008. *Assessing Academic English: Testing English proficiency 1950–89: The IELTS solution*. Cambridge: Cambridge University Press.
- Foucault, M. 1987. *Vigiar e Punir: história da violência nas prisões*. Petrópolis: Editora Vozes.
- Foucault, M. 1995. *Discipline and Punish: The Birth of the Prison*. Second edition. New York: Vintage Books.
- Foucault, M. 2007. *Security, Territory, Population*. Basingstoke: Palgrave Macmillan.
- Fulcher, G. 2010. *Practical Language Testing*. London: Hodder Education/Routledge.
- Fulcher, G. 2015. *Re-examining Language Testing: A Philosophical and Social Inquiry*. London & New York: Routledge.
- Fulcher, G.; Davidson, F. 2007. *Language Testing and Assessment: An Advanced Resource Book*. London & New York: Routledge.
- Fulcher, G.; Davidson, F.; Kemp, J. 2011. Effective rating scale development for speaking tests: Performance decision trees. *Language Testing* 28(1): 5-29.
- Fulcher, G.; Harding, L. (Eds.). 2021. *The Routledge Handbook of Language Testing*. London: Routledge. <https://doi.org/10.4324/9781003220756>.
- Gooch, R.; Saito, K.; Lyster, R. 2016. Effects of recasts and prompts on L2 pronunciation development: Teaching English /r/ to Korean adult learners. *System Journal* 60(1): 117-127.
- Harding, L.; Mcnamara, T. F. 2018. Language assessment: The challenge of ELF. In *Routledge Handbook of English as a Lingua Franca (1st ed., 570-582)*. Abingdon & New York: Routledge.
- Julkowska, I. A.; Cebrian, J. 2015. Effects of listener factors and stimulus properties on the intelligibility, comprehensibility and accentedness of L2 speech. *Journal of Second Language Pronunciation* 1(2): 211-237.
- Kelly, M. 2019. What's In a Norm? Foucault's Conceptualisation and Genealogy of the Norm. *Foucault Studies* 27: 1-22. <https://doi.org/10.22439/fs.v27i27.5889>.

- Kennedy, S.; Trofimovich, P. 2008. Intelligibility, comprehensibility and accentedness of L2 speech: The role of listener experience and semantic context. *Canadian Modern Language Review* 64(3): 459-489.
- Machado, F. F.; Delfino, M. S. L.; Rodrigues, C. O. 2021. Olhares decoloniais para o ensino de inglês. *Revista Primeira Escrita* 8(1): 115-129.
- Mcnamara, T. 2011. Managing learning: Authority and language assessment. *Language Teaching* 44(1): 500-515.
- Mohammed, N. A.; Raof, A. H. A.; Yusof, M. A. 2021. Challenges in Implementing the CEFR in English Language Classrooms. *LSP International Journal* 8(2): 115-128. <https://doi.org/10.11113/lspi.v8.17977>.
- O'Brien, M. G. 2021. Ease and Difficulty in L2 Pronunciation Teaching: A Mini-Review. *Front. Commun* 5: 626985. <https://doi.org/10.3389/fcomm.2020.626985>.
- Oliveira, L. A. 2014. Métodos de ensino de inglês: teorias, práticas, ideologias. São Paulo: Parábola Editorial.
- Pennycook, A. 2017. The cultural politics of English as an international language. London and New York: Routledge.
- Segat, G. L.; Sarmiento, S. 2022. Avaliação linguística em periódicos de Linguística Aplicada no Brasil: mapeamento das publicações. *Revista Horizontes de Linguística Aplicada* 21(2): 1-20.
- Spolsky, B. 1995. Measured words. Oxford: Oxford University Press.
- Thomson, R. 2018. Measurement of accentedness, intelligibility, and comprehensibility. In *Assessment in Second Language Pronunciation* (1st ed., 11-29). New York: Routledge.
- Zaki, A. W.; Darmi, R. 2021. The implementation of CEFR in ESL Learning: Why does it matter to the Malaysian Education System? *Asian Journal of Assessment in Teaching and Learning* 11(2): 1-13.