

A LÍNGUA PORTUGUESA E A TRADUÇÃO NO MUNDO DIGITAL

António Branco
Universidade de Lisboa

1. Introdução

O presente texto corresponde às notas que preparei para a minha participação no **painel de debate "Português Plural"**, em que participaram também Fernando Frutuoso de Melo, Chefe da Casa Civil do Presidente da República e antigo Diretor-Geral na Comissão Europeia, Ana Paula Laborinho, Presidente do Camões - Instituto da Cooperação e da Língua I.P., José Ribeiro e Castro, advogado e ex-deputado na Assembleia da República, e Célio Conceição, professor e investigador na Universidade do Algarve, painel que foi conduzido pelo jornalista Nicolau Santos, Diretor do semanário "Expresso".

Este painel de debate fez parte do programa da **conferência "30 anos de Português na UE"**, que teve por objetivo celebrar a língua portuguesa e assinalar a presença do português nas instituições europeias.

Esta conferência foi organizada conjuntamente pela Comissão Europeia, pelo Parlamento Europeu e pelo Camões - Instituto da Cooperação e da Língua I.P., e teve lugar no Museu do Oriente, em Lisboa, por ocasião do **Dia Europeu das Línguas, no dia 26 de setembro de 2016**.

Na sua essência, estas notas retomam análises mais alongadas e devidamente circunstanciadas que se encontram no *Livro Branco sobre a Língua Portuguesa na Era Digital* (Branco *et al.*, 2012) e no artigo (Branco, 2013), publicado na revista do IILP-Instituto Internacional da Língua Portuguesa, organismo da CPLP-Comunidade de Países de Língua Portuguesa.

2. Linguagem e tecnologia

Tal como em outras áreas da existência humana, a evolução científica e tecnológica tem alterado as condições de utilização, e da própria existência, das línguas naturais ao longo da história. Em alguns casos, estas novas condições de utilização das linguagens resultaram daquilo a que se pode chamar de choques tecnológicos, que estiveram na origem de profundas revoluções civilizacionais.

Um dos primeiros choques tecnológicos envolvendo a linguagem humana de que há registo consistiu no advento da escrita, há cerca de seis mil anos. Esta inovação passou a permitir que os interlocutores comunicassem em linguagem natural de forma assíncrona, sem terem de estar na presença um do outro em simultâneo.

Com a escrita, pode dizer-se que se quebrou a barreira do tempo na utilização da linguagem. Por sua vez, com o advento da imprensa mecânica, há cinco séculos atrás, quebrou-se a barreira social no acesso à informação escrita. As publicações generalizaram-se e deixaram de estar acessíveis apenas para um pequeno grupo de leitores.

Há algumas décadas atrás um outro choque tecnológico para a linguagem natural teve lugar com o advento das telecomunicações. A barreira do espaço na utilização da linguagem foi quebrada, passando então a ser possível aos interlocutores comunicarem de forma síncrona apesar de não se encontrarem presentes no mesmo local.

Todos estes choques tecnológicos na utilização da linguagem natural tiveram impactos civilizacionais enormes, sobejamente assinalados pelos historiadores. O que tem sido porém menos assinalado são os seus impactos nas condições de existência das próprias línguas naturais. Estas mudanças tecnológicas têm proporcionado uma evolução tremenda nas condições de comunicação entre indivíduos: possibilitaram alargar as virtualidades da linguagem humana muito para além do que esta permite se restringida apenas a situações de conversa oral face a face. Mas a par dessa evolução, induziram também fortes mecanismos de involução ao nível das línguas.

A cada choque tecnológico, novas condições de utilização da linguagem têm colocado novas e mais estritas condições de existência para os diferentes idiomas. Um exemplo bem conhecido é o das línguas sem sistema de escrita, que foram desaparecendo, e continuam a desaparecer, perante a vantagem de se usar outras línguas, com escrita, que permitam tirar partido dos benefícios dessa superioridade tecnológica.

O progresso tecnológico parece assim exibir uma natureza mefistofélica no que toca à linguagem humana. Ao mesmo tempo que permite alargar a capacidade de comunicação entre indivíduos, cria também fatores de redução da diversidade linguística e do multilinguismo, e dessa forma de redução do património cultural e da pluralidade de mundividências que as diferentes línguas sustentam. Dos cerca de 6 000 idiomas existentes hoje em dia, estima-se que cerca de 2 500 correm o risco de se extinguirem nas próximas décadas, como acontece a cada ano que passa com as várias línguas que morrem com o desaparecimento do seu último falante.

Neste enquadramento, cabe atentarmos no mais recente choque tecnológico para a utilização da linguagem humana, que se encontra em curso nos dias de hoje. Este choque é provocado pela utilização das novas tecnologias digitais e resulta na expansão das condições de utilização das línguas para um novo patamar sem precedentes.

3. Tecnologia da linguagem

A aplicação das novas tecnologias à linguagem natural, e em concreto o processamento computacional das línguas naturais, está a dar origem a uma nova área de investigação, desenvolvimento e inovação conhecida por tecnologia da linguagem. De um ponto vista genérico, esta tecnologia pode ser caracterizada como ocupando-se com a obtenção da representação do significado a partir do processamento computacional de expressões linguísticas, e vice-versa, da obtenção de expressões linguísticas a partir do processamento da representação do seu significado.

A tecnologia da linguagem engloba duas subáreas, conhecidas por processamento da linguagem e por processamento da fala. O processamento da fala ocupa-se com a obtenção de uma representação discreta a partir de um sinal analógico correspondente a enunciados orais, e vice-versa. O processamento da linguagem, por sua vez, ocupa-se em mapear entre uma sequência discreta de símbolos linguísticos e a representação do seu significado.

A tecnologia da linguagem está a abrir o leque inaudito de novas condições e de novas oportunidades para a utilização da linguagem natural: Ajudará as pessoas a comunicar entre si ainda que não falem uma língua comum; Apoiará uma nova geração de interfaces naturais e intuitivas, baseadas em linguagem natural, com todo o tipo de dispositivos, desde eletrodomésticos até robôs pessoais.

As novas oportunidades propiciadas pela tecnologia da linguagem resultam da sua exploração para o desenvolvimento de diversas aplicações computacionais, que podem afetar de modo diverso as condições de utilização da linguagem natural, e que se encontram atualmente em diferentes estados de maturidade tecnológica e de comercialização.

Algumas destas aplicações são usadas em contextos de utilização profissional especializados, como é o caso, por exemplo, dos detetores de plágio, para o trabalho de avaliação dos estudantes por parte dos professores, ou dos ambientes de apoio automatizado à tradução, para o trabalho dos tradutores.

Algumas outras aplicações são utilizadas de forma mais generalizada, fazendo parte do nosso quotidiano, como por exemplo, os corretores ortográficos.

Certas aplicações da tecnologia da linguagem, ainda que cruciais para o desempenho do sistema global em que se inserem, não são apercebidas pelos utilizadores, como é o caso, por exemplo, dos detetores de linguagem nos motores de busca.

Algumas delas, sendo de interesse para todo o tipo de utilizadores, são de especial relevo para pessoas portadoras de deficiências, como é o caso dos sintetizadores de voz no apoio à leitura por invisuais.

Há aplicações que servem propósitos circunscritos de utilização, como no caso dos sistemas de produção de documentação com base em linguagem controlada, usados na produção de manuais técnicos de manutenção de aeronaves. Outras aplicações servem um leque mais amplo de situações de utilização, como é o caso dos sistemas de reconhecimento de fala, que podem ser integrados em dispositivos de controlo em automóveis ou apoiar sistemas de legendagem automática de vídeos, entre inúmeras outras utilizações.

Algumas destas aplicações já encontraram não só a sua maturidade tecnológica mas também os modelos de negócio apropriados que permitem a sua ampla difusão e utilização, como acontece com os populares motores de busca de documentos.

Outras aplicações encontram-se ainda em fase de investigação científica ou de protótipo, como é o caso, por exemplo, dos sumarizadores.

Muitas outras aplicações encontrar-se-ão certamente ainda por ser desenhadas e imaginadas consoante a tecnologia da linguagem, que lhes servirá de base, vier a ser desenvolvida, amadurecida e explorada em novas soluções para os utilizadores.

No quadro do presente painel de debate, e da conferência em que se integra, a tradução automática, quer de texto para texto, como de voz para voz, é uma das aplicações da tecnologia da linguagem do maior interesse e importância.

4. Língua portuguesa e tradução

Depois das barreiras do tempo e do espaço e da barreira social terem sido quebradas em anteriores choques tecnológicos, é a própria barreira da linguagem, e da diversidade linguística, que se encontra agora a ser ultrapassada com a revolução resultante da tecnologia da linguagem.

Tal como nos choques tecnológicos anteriores, o impacto civilizacional será tremendo e de uma dimensão ainda difícil de antever na sua totalidade. De modo análogo, a par das

novas oportunidades abertas, também novos desafios e riscos se colocam agora aos idiomas e à sua existência.

No passado, línguas que em consequência do seu contexto económico e histórico particular, não foram estudadas e, por exemplo, para as quais não foi desenvolvido um sistema de escrita viram-se votadas ao beco sem saída da extinção. Hoje em dia, esse desafio joga-se ao nível da ciência e da tecnologia da linguagem. Línguas que, em consequência do seu contexto económico e histórico, não venham a ser alvo de estudo científico e para as quais não sejam desenvolvidas as soluções tecnológicas apropriadas, são línguas que a prazo arriscam a sua progressiva irrelevância, em favor de outras melhor preparadas, e eventualmente a sua própria extinção.

Estas observações aplicam-se à tecnologia da linguagem, em geral, e por maioria de razão, à tradução automática em particular. Uma língua que não esteja equipada com tradução automática de nível competitivo com o estado da arte será uma linguagem que na era digital será gradualmente preterida a favor de outras que, estando devidamente equipadas com esta tecnologia, permitem comunicar com o maior número possível de falantes de outros idiomas e assim explorar as vantagens económicas, individuais e coletivas, que daí advêm.

No seu atual estado de desenvolvimento tecnológico, a tradução automática está a abrir caminho para ganhos de produtividade substanciais no sector da tradução. Está igualmente a abrir caminho para a ampliação da exploração comercial deste sector ao viabilizar novas respostas às imensas necessidades de tradução ainda por satisfazer, nas mais diversas esferas de atividade a nível mundial, a custos que deixaram de ser impeditivos.

Na atividade como investigador na área da ciência e tecnologia da linguagem, tive a oportunidade de participar e dirigir projetos de investigação na área da tecnologia da linguagem, e da tradução automática em particular, com colegas e equipas de vários países da UE. É de assinalar que, nos meios da investigação científica, o Departamento de Português da Diretoria Geral da Tradução da Comissão Europeia é consistentemente assinalado e reconhecido por, de entre todos, se ter destacado como o Departamento que é pioneiro na aplicação da tradução automática aos serviços de tradução e é exemplar na colaboração sempre disponível e diligente com a investigação e o desenvolvimento desta tecnologia.

Referências

Branco, António, Amália Mendes, Sílvia Pereira, Paulo Henriques, Thomas Pellegrini, Hugo Meinedo, Isabel Trancoso, Paulo Quaresma, Vera Lúcia Strube de Lima e Fernanda Bacelar. 2012. *Livro Branco sobre a Língua Portuguesa na Era Digital / White Paper on the Portuguese Language in the Digital Age*, Coleção Livros Brancos, Berlim, Springer, ISBN 978-3-642-29592-8.

Branco, António. 2013. "A Língua Portuguesa face ao Choque Tecnológico Digital", *Revista do Instituto Internacional da Língua Portuguesa*, Praia, IILP-Instituto Internacional da Língua Portuguesa, vol. 2, nº 3, pp.28-36.