O computador e a tradução

Diana Santos
Processamento computacional do português
SINTEF Telecom and Informatics
Box 124 Blindern, N-0314 Oslo, Noruega
http://www.portugues.mct.pt/
Diana.Santos@informatics.sintef.no

Nesta apresentação tentarei comunicar três ideias principais:

1. O computador é uma ferramenta multifacetada em relação à actividade humana da tradução

2. A tradução técnica e/ou científica não é mais simples do que a tradução literária

3. A prática (= implementação) é essencial

Por esta última razão, depois de tecer algumas considerações gerais sobre o assunto, e apelar à criação de um banco de traduções público, passo a palavra à representante da Logos e minha antiga colaboradora e amiga Anabela Barreiro Colasuonno.

Que papéis pode ter o computador?

Ao debruçarmo-nos sobre a actividade humana da tradução (insisto no modificador "humano" visto que a tradução entre linguagens naturais (=línguas) só faz sentido para a comunicação entre pessoas com diferentes línguas maternas), existem pelo menos quatro formas diferentes (mas não mutuamente exclusivas) de encarar o computador:

1. O computador como **facilitador da tradução** (permitindo a edição do texto, a organização da informação, a criação de um repositório de antigas traduções, um observatório da(s) (duas) língua(s) em uso)

2. O computador como **auxiliar da tradução** (permitindo o acesso a terminologia bilingue, a bases de dados de traduções, como verificador de uma linguagem controlada, como crítico de repetições ou sugerindo traduções já efectuadas)

3. O computador como **auxiliar do estudo e do ensino da tradução** (descobrindo regularidades e problemas, automatizando a construção de um léxico bilingue, procurando automaticamente sugestões, descobrindo dificuldades sobre as quais é preciso insistir na formação dos tradutores, criando problemas para serem usados como auxiliares para ensino; veja-se Frankenberg-Garcia (1999))

4. O computador como **tradutor** (funcionando como tutor, consultor, como tradutor automático, como "gist translator" – dando uma ideia do que poderá ser o conteúdo do que estamos a ver numa língua totalmente desconhecida; cf. Resnik (1997))

Neste último caso, repare-se que um tradutor não faz apenas a tradução completa de textos. De facto, também pode ser usado como fonte de informação (consultor), ou como avaliador de traduções. De forma análoga, um computador como tradutor não tem de se limitar à situação "sistema automático que executa a tradução de textos".

Por outro lado, ainda que a responsabilidade do computador em relação ao produto final (uma tradução a ser efectuada) esteja numa escala crescente na lista anterior, essa ordenação não reflecte necessariamente uma medida de dificuldade (seja da tarefa da tradução, seja da programação dessa mesma tarefa para ser executada pelo computador). É muito mais simples pôr um computador a funcionar como tradutor (por exemplo no caso restrito de palavras ou

termos técnico-científicos), atingindo mesmo facilmente um desempenho superior ao de um ser humano, do que conseguir que um computador auxilie efectivamente a concepção de cursos de formação específica de tradutores para um dado par de línguas, através da identificação automática de casos problemáticos.

De qualquer forma, parece-me que esta identificação dos vários papéis que o computador pode representar no objectivo final de melhorar o trabalho do tradutor (tanto as condições de trabalho como o produto final) é útil para posicionar as questões do desenvolvimento de ferramentas computacionais na área, da formação (contínua) de recursos humanos e do planeamento de recursos partilháveis. Para uma panorâmica actual dos usos do computador na tradução – processo e resultado – sugiro a consulta às actas da conferência "Translating and the Computer 21" (TC21), há dez dias em Londres.

A tradução literária é diferente; não mais difícil

Na minha opinião, é completamente falsa a convicção de que a tradução técnica e científica é mais fácil do que a tradução literária. (Tal está ligado a duas outras hipóteses que também carecem de justificação: a de que um texto técnico é mais simples do que um literário[1]; e a de que a tradução automática começou a ser aplicada a textos técnicos porque estes eram mais simples[2].)

Não querendo evidentemente diminuir a dificuldade e o interesse da tradução de textos literários, gostava de chamar a atenção para que, **quando bem feita**, a tradução científica pode implicar processos de grande complexidade linguística e não só. Com efeito,

1. A tradução científica exige maior rigor

2. A tradução técnico-científica dá origem a maior reestruturação e reescrita do que a tradução das obras literárias, visto que não existe o "respeito pel(a linguagem[3] d)o autor", como explica por exemplo Doherty (1992)

3. A tradução científica tem dificuldades linguísticas comparáveis às da tradução literária, embora noutros eixos (Snell-Hornby, 1995): metáforas, linguagem criativa, elevado recurso a conhecimento implícito na área; para alguns exemplos nesse campo remeto para Santos (1999b) ou Fabricius-Hansen (1998)

4. A tradução técnica não é culturalmente neutra (porque a língua nunca o é), e isso pode ter consequências na estruturação do texto, na forma de apresentação, no tipo de linguagem

O mesmo, aliás, se pode dizer da "localização" – neologismo (de sentido) para indicar "tornar um dado produto ou serviço apropriado para um dado local=cultura", e que eu sugiro passar a "portuguesificação" no que nos diz respeito.[4] Ora a localização – ou melhor, a portuguesificação – é uma tarefa complexa, se for bem feita, que tem de entrar em conta com aquilo que é funcionalmente relevante (veja-se Catford, 1967) no texto/produto/serviço fonte e vertê-lo para a cultura portuguesa.

---

[1] Evidentemente que pode haver várias dimensões de simplicidade, que darão resultados opostos: tamanho das frases; grau de polissemia das unidades lexicais; complexidade sintáctica e argumentativa; necessidade de compreensão do mundo ou de um domínio especializado; audiência esperada; etc.

[2] As empresas que financiaram a tradução automática investiram na tradução dos textos com que mais facilmente poderiam experimentar, e para os quais fosse mais fácil medir as vantagens económicas: os textos produzidos pelas próprias empresas. Foram, portanto, considerações económicas, não científicas ou linguísticas, que levaram os textos técnicos a serem a matéria-prima inicial na tradução automática.

[3] No sentido da forma como o autor se exprime.

[4] Por oposição a "aportuguesamento", ou seja, mascarar de português uma coisa estrangeira e que é, portanto, uma operação superficial e, além disso, conotada negativamente.

Nunca é demais insistir, a este propósito, que conhecer (mesmo dominar perfeitamente) o aspecto formal de uma língua não é compreender (nem por conseguinte funcionar em) essa língua. O conjunto das estratégias comunicativas, dos códigos implícitos, dos sistemas de valores associados a uma língua são essenciais para a tradução, para a localização, para a comunicação, como o demonstram por exemplo os estudos reunidos na conferência "Cultural Attitudes Towards Communication and Technology", CATaC'98 (Ess, 1998).

Note-se que enunciar as considerações anteriores não significa afirmar a impossibilidade de automatização da tradução, mas apenas e só identificar estes factores como necessários ao seu sucesso. De facto, a tradução científica e técnica é uma arte e uma técnica que exigem um conhecimento profundo das culturas e subculturas de cada língua envolvida, de forma a que o texto produzido desempenhe a sua função no sistema da língua alvo (Toury, 1995). Esse conhecimento e a sua manipulação podem ser extremamente facilitados pelo computador, tanto por a tradução técnica envolver um tamanho muito superior de dados em relação à tradução literária, como pelo facto de esses dados serem cada vez mais assistidos pelo computador também na sua criação.[5]

Na prática é que surge a teoria

Como engenheira, custa-me ver, em áreas variadas da ciência, que a teoria toma de tal modo a dianteira que perde totalmente o contacto com a prática, de forma que já só a nível do nome as duas se encontram, com prejuízos evidentes para ambas. Também na engenharia isto acontece, com sistemas de que apenas existem protótipos ou mesmo que nunca saíram do papel, mas que são "teoricamente correctos".

No campo da tradução, não vou aqui debruçar-me sobre o que isto pode significar em termos de recursos humanos, visto que tal já terá, porventura, sido discutido no painel sobre formação do presente encontro. Vou simplesmente referir-me à necessidade absoluta de criar sistemas que funcionem (e sejam usados) nas quatro perspectivas que mencionei anteriormente, para que a nossa língua não seja vítima da sociedade da informação (Santos, 1999).

A primeira propriedade que se deve exigir de um sistema é que exista, e que funcione, ou seja, que desempenhe uma dada função com uma margem de erro aceitável. Por outro lado, para ser possível criar um sistema informático que desempenhe uma tarefa útil é preciso conhecer a fundo as necessidades do utilizador alvo desse sistema e, depois, fazer muitas concessões.

Em particular, é preciso

a) considerar o sistema que se está a desenvolver como um **auxiliar** na resolução de um problema complexo – não a solução derradeira desse problema;

b) saber avaliar os resultados: quais as vantagens da sua utilização em comparação com a situação anterior, em que o sistema não existia;

c) saber medir e defender as opções tomadas ao longo da criação e desenvolvimento do sistema.

Para levar a bom termo uma actividade tão complexa como o é a tradução, é essencial a interacção utilizadores – engenheiros, ou seja, é preciso que os tradutores (os utilizadores do computador como ferramenta de tradução) colaborem, em todas as fases, com os engenheiros da linguagem/linguistas computacionais/especialistas em processamento de linguagem natural que implementam os sistemas.

[5] A maior parte da documentação técnica é já presentemente criada com o apoio de programas de ajuda à redacção cada vez mais inteligentes.

Gostava, pois, de sugerir a criação de um banco de traduções e de problemas de tradução que pudesse ser partilhado

- por tradutores profissionais na sua actividade quotidiana,
- por alunos e professores para satisfazer as necessidades do ensino da tradução,
- e por investigadores e engenheiros que a ele pudessem recorrer para tornar os seus sistemas mais adequados.

Tal projecto deveria contemplar a criação de

- corpora paralelos
- memórias de tradução
- exemplos comentados de problemas de tradução

de forma a permitir um conjunto de repositórios, a aumentar pelos utilizadores, passível de constituir um manancial de conhecimento que avançasse significativamente a tradução e a construção de ferramentas computacionais para ela vocacionadas, no que toca à nossa língua.

Referências

Catford, J.C. 1967. *A Linguistic Theory of Translation: An Essay in Applied Linguistics*, Oxford: Oxford University Press.

Doherty, Monika. 1992. "Informationelle Holzwege: Ein Problem der Übersetzungswissenschaft", *Zeitschrift für Literaturwissenschaft und Linguistik* **84** (1992), pp.30-49.

Ess, Charles. 1998. "First Looks:CATaC'98", in C.Ess & F. Sudweeks (eds.), *Proceedings of Cultural Attitudes Towards Communication and Technology '98* (University of Sydney, Australia), pp.1-17, http://www.it.murdoch.edu.au/~sudweeks/catac98/01_ess.html.

Fabricius-Hansen, Cathrine. 1998. "Information density and translation, with special reference to German-Norwegian-English", in Stig Johansson & Signe Oksefjell (eds.), *Corpora and Cross-linguistic Research: Theory, Method, and Case Studies*, Amesterdão: Rodopi, pp.197-234.

Frankenberg-Garcia, Ana. 1999. "Crosslinguistic influence as a key to extracting second language teaching materials for monolingual classes from translation corpora", in Sylviane Granger (ed.), *Proceedings of the Workshop on Contrastive Linguistics and Translation Studies: Empirical Approaches* (Louvain-la-Neuve, 5-6 Fevereiro de 1999).

Resnik, Philip. 1997. "Evaluating Multilingual Gisting of Web Pages", in *Natural Language Processing for the World Wide Web, Papers from the 1997 AAAI Spring Symposium* (Stanford, March 24-26, 1997), Menlo Park, California: AAAI Press, pp.129-135.

Santos, Diana. 1999. "Processamento computacional da língua portuguesa: Documento de trabalho", http://www.portugues.mct.pt/branco/, Versão base de 9 de Fevereiro de 1999, última alteração efectuada a 13 de Abril de 1999.

Santos, Diana. 1999b. "Um olhar computacional sobre a tradução", *Terminología y Traducción* **2/99**.

Snell-Hornby, Mary. 1995. *Translation Studies: An integrated approach*. Amesterdão/Filadélfia: John Benjamins, Edição revista [1ª edição 1988].

TC21. Translating and the Computer 21, http://www.aslib.co.uk/conferences/tc21.html.

Toury, Gideon. 1995. *Descriptive translation studies and beyond*, Amesterdão: Benjamins.