

EDITORA HUMANITAS

Presidente

Mario Miguel González

Vice-Presidente

Marco Aurélio Werle

CONSELHO EDITORIAL

Titulares

*Antonio Dimas
Beth Brait
Beatriz Perrone-Moisés
Berta Waldman
José Jeremias de Oliveira Filho
Sueli Angelo Furlan
Valéria de Marco
Vera Lúcia Amaral Ferlini*

Suplentes

*Gildo Marçal Brandão
Margarida Maria Taddoni Petter
Maria Luíza Tucci Carneiro
Oswaldo Humberto Leonardi Caschin
Vera da Silva Telles
Véronique Dablet*

UNIVERSIDADE DE SÃO PAULO
FACULDADE DE FILOSOFIA, LETRAS E CIÊNCIAS HUMANAS

Diretor

Gabriel Cohn

HUMANITAS

Av. Prof. Luciano Gualberto, 315, sala 11
Cid. Universitária
05508-010 – São Paulo – SP – Brasil
Tel.: 3034-2733
e-mail: editorahumanitas@usp.br
<http://www.editorahumanitas.com.br>

Foi feito o depósito legal
Impresso no Brasil / *Printed in Brazil*
Setembro 2008

**AVANÇOS DA
LINGUÍSTICA DE CORPUS
NO BRASIL**

*Organizado por
Stella E. O. Tagnin
Oto Araújo Vale*

HUMANITAS

São Paulo, 2008

Text and technology: in honour of John Sinclair. Amsterdam: John Benjamins, 1993. p.233-250.

_____. *Corpora in translation studies: an overview and some suggestions for future research*. *Target*, Amsterdam, v. 7, n. 2, p. 223-243, 1995.

_____. *Corpus-based translation studies: the challenges that lie ahead*. In: SOMERS, H. (Ed). *Terminology, LSP and translation studies in language engineering in honour of Juan C. Sager*. Amsterdam: John Benjamins, 1996. p. 177-186.

BERBER SARDINHA, A. P. *Linguística de corpus: histórico e problemática*. *DELTA: documentação de estudos em linguística teórica e aplicada*, São Paulo, v. 16, n. 2, p. 323-367, 2000.

_____. *Uso de corpora na formação de tradutores*. *DELTA: documentação de estudos em linguística teórica e aplicada*, São Paulo, v. 9, n. especial, p. 323-367, 2003.

_____. *Linguística de Corpus*. Barueri, SP: Manole, 2004.

_____. *Pesquisa em linguística de corpus com WordSmith Tools*, no prelo.

CAMARGO, D. C. de *Padrões de estilo de tradutores: um estudo de semelhanças e diferenças em corpora de traduções literárias, especializadas e juramentadas*. 2005. 512 f. Tese (Livre-Docência em Tradução) – Instituto de Biociências, Letras e Ciências Exatas – Unesp, Universidade Estadual Paulista, São José do Rio Preto, 2005.

_____. *Metodologia de pesquisa em tradução e linguística de corpus*. São Paulo: Cultura Acadêmica; São José do Rio Preto, SP: Laboratório Editorial do IBILCE, UNESP, 2007.

PAIVA, P. T. P. *Estudo baseado em corpora de traduções e três glossários bilíngües nas subáreas de anestesiologia, cardiologia e ortopedia*. 2006. 223f. Dissertação (Mestrado) Instituto de Biociências, Letras e Ciências Exatas – IBILCE, Universidade Estadual Paulista, São José do Rio Preto, 2006.

SCOTT, M. *WordSmith Tools: version 3.0*. Oxford: Oxford University Press, 1999.

Estudo da competência tradutória e seu desenvolvimento com uso de *corpus* de traduções

Heloisa Pezza Cintrão*

Resumo: Neste trabalho, (1) consideram-se algumas convergências e divergências entre a Linguística de *Corpus* e as abordagens cognitivas da tradução; (2) descreve-se um estudo sobre o desenvolvimento da competência tradutória baseado em *corpus*.

Palavras-chave: abordagens cognitivas da tradução, *corpus*, desenvolvimento da competência tradutória

Abstract: This paper presents (1) a discussion of some convergences and divergences between *Corpus* Linguistics and cognitive approaches to translation; (2) a description of a *corpus*-based study on the development of translation competence.

Keywords: cognitive approaches to translation, *corpus*, development of translation competence

* Docente nas Letras – Universidade de São Paulo - Área de Espanhol. E-mail: helocint@usp.br

1. As abordagens cognitivas da tradução e a Linguística de Corpus

Em 1972, James Holmes fazia a primeira proposta de visão geral dos Estudos de Tradução, distinguindo os possíveis ramos dentro dessa disciplina acadêmica. A proposta de Holmes tornou-se clássica: seu mapeamento é hoje “amplamente aceito como um quadro sólido para situar as atividades acadêmicas dentro desse domínio”¹ (Baker 1998: 277).

Foi também Holmes o responsável pela adoção do nome ‘Estudos de Tradução’ (*Translation Studies*) para referir-se à disciplina como um todo, propondo que esse termo fosse entendido conforme a seguinte formulação de Werner Koller (1971): “uma designação de conjunto e abrangente para todas as atividades de pesquisa que tomam o fenômeno do ‘traduzir’ e da ‘tradução’ como base ou foco” (*apud* Holmes 2000 [1972]: 176). Porque se baseiam na observação de certos objetos – o traduzir e a tradução –, Holmes afirma que os Estudos de Tradução são uma **disciplina empírica**, e aponta os dois principais objetivos desse tipo de disciplina, citando Carl G. Hempel: “descrever fenômenos particulares do mundo de nossa experiência e estabelecer princípios gerais que permitam explicar e prever tais fenômenos” (*apud* Holmes, *idem*).

De acordo com isso, em seu **ramo de pesquisa ‘pura’** (não aplicada), os Estudos de Tradução teriam, na visão de Holmes, dois objetivos principais: (1) descrever os fenômenos do traduzir e da tradução, tal como se manifestam no mundo de nossa experiência; (2) estabelecer princípios capazes de explicar e prever tais fenômenos. Poder-se-ia, então, considerar a existência de dois ramos de estudos ‘puros’: (i) os estudos descritivos da tradução (EDT); (ii) os estudos teóricos da tradução. Os **estudos descritivos** seriam o ramo da disciplina com contato sistematicamente mais direto com os fenômenos

¹ A tradução das citações será sempre minha.

empíricos estudados. Para Holmes, pareceria haver três tipos de pesquisa possíveis no ramo descritivo: (a) estudos do produto (*product-oriented*); (b) estudos da função (*function-oriented*); (c) estudos do processo (*process-oriented*).

Nos EDT orientados ao produto, descrevem-se as traduções existentes. Segundo Holmes, esse ramo, em sua fase inicial, ocupou-se da descrição de traduções individuais e, numa segunda fase, de análises comparativas de várias traduções do mesmo texto, para uma mesma língua ou para várias. As descrições individuais ou comparadas fornecem material para compor *corpora* de traduções, observou Holmes, representativos de um período determinado, de certa língua, e/ou de certo tipo textual ou discursivo.

Os EDT orientados à função não descrevem as traduções em si, mas sim suas funções na situação sociocultural receptora. Propõem-se questões sobre os textos que foram (ou não) traduzidos em certo período e lugar, as influências que exerceram etc.

Por fim, os EDT orientados ao processo caracterizam-se pelo foco no processo ou ato tradutório em si. Naquele ano de 1972, Holmes comentava que os fenômenos na “pequena caixa-preta” da “mente” do tradutor, no processo de criação de um texto novo e mais ou menos coincidente (*matching*) numa outra língua, haviam sido objeto de muita especulação, mas de pouca pesquisa sistemática em situações experimentais. Reconhecendo a complexidade do fenômeno, Holmes considerava que tal complexidade não inviabilizaria o desenvolvimento do sub-ramo processual, posto que a psicologia vinha tendo sucesso em desenvolver métodos para estudar fenômenos complexos.

Tendo em mente o mapeamento de Holmes [esquema no *Quadro 1*], os sub-ramos dos EDT anteriormente referidos são aqueles cujos pontos de contato com a Linguística de Corpus (LC) serão considerados a seguir.

ESTUDOS de TRADUÇÃO								
Puros					Aplicados			
TEÓRICOS		DESCRITIVOS			Formação do Tradutor	Ferramentas de apoio para a Tradução	Crítica de Trad.	Políticas de Trad.
Gerais	Parciais	Orientados ao Produto	Orientados ao Processo	Orientados à Função				
Restritos ao Meio	Restr. à Área	Restritos ao Nível	Restritos ao Tipo de Texto	Restritos no Tempo	Restritos ao Tipo de Problema			

Quadro 1: Esquema do mapa dos Estudos de Tradução proposto por Holmes

Laviosa (2004: 32) considera que os fortes vínculos estabelecidos entre os EDT e a LC, a partir do início dos anos 90, se deveram a um “conjunto de preocupações comuns provenientes de uma **perspectiva empírica**” (grifo meu).

Desde o trabalho seminal de Baker (1993), os **Estudos de Tradução Baseados em Corpora** têm-se construído a partir do diálogo entre os EDT e a LC. No entanto, na descrição dos Estudos de Tradução Baseados em *Corpora* por Laviosa (2004), nota-se que a junção entre EDT e LC se deu nas subdivisões dos EDT que Holmes chamava de “orientados ao produto” e “orientados à função”, ou então dentro do ramo aplicado dos Estudos de Tradução. Quanto aos estudos “orientados ao processo”, as possibilidades de parceria com a LC não parecem ser tão diretas nem evidentes.

Nos EDT, o sub-ramo orientado ao processo teve seu impulso inicial em fins dos anos 80 (quando se realizaram as primeiras investigações sobre os processos mentais subjacentes ao ato tradutório), e ganhou força na década de 90 (Venuti 2000: 339). De acordo com as previsões-sugestões de Holmes, essa subdivisão dos Estudos de Tradução começou valendo-se justamente de uma técnica proveniente da psicologia experimental, os *Think-Aloud Protocols* (TAPs), ou protocolos verbais, em sua forma ‘concomitante’, que consiste em (1) solicitar

a um sujeito que verbalize o máximo possível de seus pensamentos e ações durante uma tradução que estiver realizando, (2) gravar essas verbalizações e depois (3) transcrevê-las para estudo. Com essa técnica de coleta de dados, as pesquisas ‘orientadas ao processo’ nasceram almejando valer-se de métodos indutivos para confirmar hipóteses, identificar componentes ou desenhar modelos do processo tradutório com base em dados.

Embora esse tipo de estudos também compartilhe com a LC a orientação empírica, os *corpora* processuais coletados mediante protocolos verbais, pela própria natureza ‘artesanal’ da coleta, dificilmente poderiam constituir os *corpora* de grandes ou mesmo de médias dimensões que têm caracterizado a LC. É difícil imaginar *corpora* para os estudos processuais em outra forma que não seja a de um conjunto relativamente pequeno de dados. Além disso, os dados dos *corpora* compostos pelos TAPs, na verdade, não constituem o objeto de estudo, mas considera-se que podem fornecer vislumbres de um objeto não observável diretamente, ou seja, os processamentos mentais e as estratégias do tradutor. Esse tipo de *corpus* processual vem demandando a criação de procedimentos próprios de análise, muitas vezes a partir de categorias cognitivas ou psicolinguísticas. Os parâmetros de observação dos dados que começaram a ser levantados e testados, caracterizaram-se como métodos de análise mais qualitativos que quantitativos, difíceis de serem automatizados para processamento computadorizado.

A presença das ferramentas computacionais nesse ramo dos Estudos da Tradução se fez sentir com mais força recentemente, com a criação de um programa dirigido à coleta de dados processuais de tradução, o *Translog*, que, ao registrar todos os movimentos de teclado em tempo real, passou a permitir o estudo mais preciso das pausas no processo tradutório, assim como a delimitação de fases processuais que Jakobsen (2002) distinguiu como fases de ‘orientação’, ‘produção’ e ‘revisão’. Outras ferramentas computacionais têm sido exploradas, como os programas que gravam, em forma de vídeo, as ações de um tradutor,

tais como visualizadas na tela do computador, como o *Proxy* e o *Camtasia*. Essas gravações passam a permitir a observação de, por exemplo, consultas que o tradutor faz a dicionários eletrônicos ou via *Internet* (valendo-se de buscadores etc.).

Assim, os EDT processuais têm compartilhado com a LC tanto a orientação empírica quanto o uso de computadores e de ferramentas tecnológicas, embora, neste último caso, as novas tecnologias empregadas sejam usadas mais acentuadamente para a coleta de dados e formação dos *corpora* do que para a análise automática ou semi-automática de um grande volume de dados. Mesmo assim, há claras semelhanças entre o papel que os computadores vêm tendo nos EDT processuais e seu papel 'revolucionário' na linguística, via LC, tal qual formulado, por Sardinha (2004), que faz uma comparação com "o papel transformador que o telescópio teve na física e nas ciências exatas. Passamos da idealização para a sistematização da observação da evidência" (p. xvii).

O processo tradutório constituiu o objeto inicial das ditas 'abordagens cognitivas' da tradução (Hurtado 2001: 128 e 311-408). A **competência tradutória** (CT) logo despontou como o segundo construto central dessas abordagens. A CT mantém estreitas relações com o desenrolar dos processos mentais ao traduzir e, mais recentemente, aparecem propostas de cruzamento de dados de produto e processo para seu estudo empírico. No caso dos dados de produto, cada vez mais se tem chegado a um consenso de que as traduções que podem ser relevantes na constituição de um *corpus* para o estudo da CT, devem ser traduções de profissionais experientes, quando se entenda a CT como aquele conjunto de habilidades e conhecimentos que o tradutor experiente e competente possui, que o habilitam a realizar traduções de boa qualidade, segundo exigências profissionais de desempenho.

No desenvolvimento dos estudos sobre o processo tradutório e sobre a CT, os *corpora* são peculiares em vários aspectos. Os dados coletados são inevitavelmente textos produzidos especificamente para constituir *corpora* de estudo, e estes *corpora* tendem a ser de pequenas dimensões, como já apontamos, o que poderia pôr em questão sua

representatividade, do ponto de vista dos postulados da LC. Na própria via de estudo empírico-descritivo que as abordagens cognitivas foram encontrando, predomina uma observação dos dados em grande medida qualitativa, ainda que buscando regularidades (ou padrões 'quantitativos') capazes de lançar luzes sobre o processo tradutório e a CT.

Por fim, o terceiro construto que hoje constitui um objeto das abordagens cognitivas da tradução é o 'desenvolvimento da competência tradutória' ou 'aquisição da competência tradutória', que supõe a existência de diferentes níveis percorridos pelos aprendizes na direção da competência do tradutor profissional, ou seja, da CT propriamente dita. Nas palavras de Toury (1995), esse processo seria aquele pelo qual "um bilíngüe se torna um tradutor". Do ponto de vista de Shreve (1997), tratar-se-ia de um processo sem ponto final, já que não haveria exatamente algo como uma CT geral (para qualquer tipo de texto ou tarefa tradutória), universal (igual para todos os sujeitos) e estabilizada (que não chegasse mais a sofrer modificações a partir de determinado estágio de desenvolvimento).

Um estudo empírico do desenvolvimento da CT pede, em princípio, a composição de *corpora* de aprendizes de tradução, a ser comparados com *corpora* de referência de tradutores profissionais (e, portanto, 'comparáveis' com estes). No momento atual, a complexidade para se chegar a um desenho satisfatório desse tipo de *corpus* mal começou a ser enfrentada e resolvida. Supõe, entre outras coisas, um grau de criteriosidade que leve a diferenciar traduções de aprendizes em diferentes fases do processo de desenvolvimento; estabelecer um critério objetivo e uniforme para os *corpora* de tradutores 'profissionais' que servirão como referência para estimar os indicadores do 'desenvolvimento' dos aprendizes; diferenciar traduções da língua estrangeira para a língua materna daquelas feitas no sentido inverso; resolver a questão de se devem ser considerados como iniciantes no processo de desenvolvimento apenas sujeitos com elevada proficiência bilíngüe e/ou encontrar critérios objetivos para diferenciar graus de competência linguística entre os aprendizes observados; diferenciar *corpora* coletados entre

aprendizes que estão vivenciando o processo de desenvolvimento a partir de diferentes metodologias de ensino ou diferentes experiências de tradução (no caso de tradutores que aprendem 'na prática'); considerar os diferentes pares lingüísticos, buscando possíveis 'universais de aprendizagem', bem como divergências de aprendizagem que possa haver segundo as especificidades dos pares lingüísticos; fazer uma diferenciação similar entre as várias modalidades de tradução (literária, jurídica, legendação etc).

Em cada um desses casos, pode-se prever novamente a dificuldade de trabalhar com *corpora* de grandes dimensões e de conseguir métodos de análise com alto ou mesmo médio grau de automatização. A necessária criteriosidade que a LC coloca como requisito para o desenho dos *corpora* (Sardinha 2004: 19) pareceria poder ser seguida no desenho de um *corpus* para o estudo do desenvolvimento da CT no que se refere ao refinamento de sua **adequação** ao objeto de estudo, mas bem mais dificilmente no que se refere à **representatividade**, sempre que esta seja vinculada ao acúmulo em larga escala de dados digitalizados. Cabe perguntar, no contexto do "VI Encontro de Linguística de *Corpus*" (em que o presente trabalho foi preparado), quais seriam, então, os principais pontos de contato desse tipo de Estudo Descritivo da Tradução com a LC, e se há um futuro estreitamento possível de suas relações com a LC, como houve no casos dos EDT orientados ao produto e à função.

De início poderíamos responder que os pontos de contato com a LC estariam principalmente nas seguintes características das abordagens cognitivas da tradução, tal como estas se apresentam em seu estado atual:

- * adoção de um paradigma empírico-descritivo nas investigações, incluindo preocupações de validade estatística dos dados coletados e dos resultados das pesquisas;
- * composição de *corpora* em formato digital (de diferentes tipos: dados de produto e dados de processo);

- * uso de ferramentas computacionais, sendo que, nas abordagens cognitivas, essas ferramentas predominam na coleta de dados (*Translog*, *Camtasia*), mas também têm seu papel na análise dos dados dos *corpora* (ex.: uso de programas estatísticos e do *Excel*);
- * possibilidade de procedimentos comuns na organização dos *corpora*, especialmente no que se refere aos dados de produto (traduções dos sujeitos investigados), como, por exemplo, coleta de dados de perfil dos sujeitos, cabeçalhos e etiquetagem.

Dada a incipiência das investigações empíricas da CT e de seu desenvolvimento, o primeiro desafio tem sido estabelecer critérios e ferramentas para configurar *corpora* adequados aos objetivos de pesquisa e para analisá-los. O que se relata nas próximas linhas deste trabalho é um estudo que consistiu numa tentativa nesse sentido. Falaremos dos procedimentos de constituição de um *corpus* para a observação empírica do desenvolvimento da CT no par português-espanhol; das características atuais que esse *corpus* apresenta (já que, por enquanto, o temos ainda como um *corpus* aberto a ampliações e complementações); das primeiras análises e daqueles procedimentos de análise que vislumbramos poderem ser aplicados ao *corpus* a partir do instrumental típico da LC.

2. Um *corpus* para observar o desenvolvimento da CT e uma metodologia de análise

Entre outubro de 2004 e maio de 2005, compilei um *corpus* com o **objetivo geral** de reunir dados para um estudo sobre o desenvolvimento da CT. O desenho peculiar desse *corpus* respondeu, mais afuniladamente, ao **objetivo específico** de testar algumas hipóteses metodológicas relativas a uma disciplina-piloto de tradução direta geral, em forma de um curso-oficina, oferecida a estudantes universitários em fases iniciais de aprendizagem de espanhol como língua estrangeira. Tratou-se de um *corpus* de pequenas dimensões, composto de produ-

ções de estudantes que participaram da disciplina-piloto, e de sujeitos de dois grupos de controle. O núcleo desse *corpus* foram traduções para a língua materna, realizadas pelos sujeitos no *Translog*. As traduções realizadas no *Translog* fornecem um texto digitalizado em arquivo TXT. Esse programa também registra todos os movimentos do teclado em tempo real, gerando um arquivo LOG que (1) permite reproduzir na tela toda a digitação do texto pelo sujeito, como se estivesse sendo feita naquele momento (função *replay*); (2) fornece dados precisos do tempo total da tradução, das pausas feitas, da duração e lugar de cada pausa, das correções na tradução durante a escrita e durante a revisão, enfim, facilita a extração de uma série de dados processuais e garante sua precisão.

As traduções em formato digital foram complementadas com dados de protocolos verbais retrospectivos, da seguinte maneira: imediatamente após cada tradução realizada, gravava-se uma entrevista com o sujeito, feita enquanto ele visualizava no computador a tradução que havia feito, estimulando que explicasse os motivos das pausas, lacunas e correções registradas pelo *Translog* durante o processo de tradução, sendo que também se pedia que justificasse as soluções que havia dado para alguns trechos potencialmente problemáticos, cujo levantamento havia sido feito por mim, previamente, para cada texto traduzido.

O grupo principal de estudantes (NP = Novatos Principal), cujo processo de desenvolvimento da CT durante a disciplina-piloto se pretendia observar, foi constituído de 8 voluntários, todos estudantes que tinham como língua materna o português do Brasil e que haviam iniciado seus estudos de espanhol como língua estrangeira na Letras-USP. Todos eles, no momento do início da coleta de dados, estavam terminando a disciplina de 'Língua Espanhola 2'. Outros 7 voluntários, exatamente com o mesmo perfil, compuseram o grupo de controle de estudantes (NC = Novatos Controle).

O *corpus* longitudinal de traduções dos dois grupos de estudantes ficou assim configurado: para cada um dos 15 sujeitos estudantes, reuniram-se 4 traduções, cuja realização foi sincronizada com o início,

o meio e o fim da disciplina-piloto de introdução à tradução. Todas essas traduções foram feitas da língua estrangeira (espanhol) para a língua materna (português), todos os textos-fonte das traduções eram contos infantis da escritora argentina María Elena Walsh, e as traduções foram feitas segundo o seguinte esquema:

T1 – conto “Historia de una princesa” – tradução realizada antes do início da disciplina²

T2 – conto “El patio” – tradução realizada entre os dois módulos da disciplina

T3 – conto “La Luna y la Vaca” – tradução realizada após o término da disciplina

T4 – conto “Historia de una princesa” – retradução realizada após o término da disciplina, imediatamente depois de T3 (poucos dias depois), entre 3 e 4 meses depois da primeira tradução do conto (T1), sem saber previamente que voltariam a traduzir um conto já traduzido, e sem ter acesso àquela primeira tradução realizada.

Além desses dois *subcorpora* longitudinais de aprendizes, foi coletado um *subcorpus* composto por traduções do mesmo conto traduzido pelos estudantes em T1 e T4, “Historia de una princesa”, por 6 outros voluntários, todos eles com formação superior em Letras-Espanhol concluída e com português como língua materna, cinco deles com mestrado completo em Letras-Espanhol, e todos com boa experiência

² A partir de T1, os 15 estudantes foram divididos entre grupo principal e de controle, e os 8 sujeitos do grupo principal iniciaram a disciplina-piloto, enquanto que os 7 sujeitos do grupo de controle apenas continuaram com as aulas de espanhol na graduação em Letras, sem receber nenhuma instrução específica em tradução, e vieram fazer as demais traduções nos mesmos períodos em que os 8 sujeitos do grupo principal as realizavam.

no ensino de espanhol como língua estrangeira. A maioria deles já havia atuado na docência de nível universitário. Dado seu perfil de formação em nível de excelência na área de Letras e dada a qualidade de sua experiência profissional nessa área, foram considerados bilíngües altamente proficientes. Foi nesse sentido que esse grupo foi chamado de 'grupo de profissionais'. É importante ressaltar, no entanto, que, embora a maioria tivesse experiência em tradução, com intensidade variável entre eles, não seriam considerados profissionais experientes em tradução segundo critérios mais estritos de estudos empíricos realizados sobre a CT. Esses sujeitos realizaram, como se disse, apenas uma tradução. Fizeram também tarefas adicionais, cumpridas por todos os grupos, que consistiram em: (a) preenchimento de um perfil; (b) resposta a um questionário para sondagem de idéias sobre tradução; realização de (c) uma tarefa de escrita em língua espanhola (estrangeira) e (d) uma tarefa de leitura em língua portuguesa (materna). Todos autorizaram por escrito a utilização dos dados do *corpus* para pesquisas. Os sujeitos de NP também responderam um questionário sobre a disciplina cursada, após terminá-la. Um dos sujeitos do grupo de profissionais não completou toda a coleta de dados.

As tarefas de leitura e de escrita também foram realizadas no *Translog*, de modo que, assim como para as traduções, esse *subcorpus* coletado disponibiliza dados de processo e de produto. O conjunto dos dados do *corpus* descrito pode ser visualizado na *Tabela 1*.

O contraste entre o grupo de controle de profissionais de Letras (PC=Profissionais Controle) e os grupos de estudantes determinou, como uma importante variável entre os sujeitos, o 'nível de competência na língua estrangeira' ou, mais amplamente, o 'nível da subcompetência bilíngüe', que diferia entre os dois grupos de estudantes, por um lado, e o grupo de profissionais, por outro. Essa variável permitiu observações interessantes nas primeiras análises feitas, com respeito a certos aspectos da independência que a competência tradutória tem da competência bilíngüe.

O elemento diferenciador entre os dois grupos de estudantes foi a participação na disciplina-piloto em tradução. De resto, procuramos controlar uma série de outras variáveis. As condições de realização das tarefas de tradução e de leitura e escrita foram controladas: todos os sujeitos as realizaram nas mesmas salas, em sessões individuais de 4h para cada tarefa, no mesmo computador, sempre no programa *Translog*, e tendo disponíveis os mesmos materiais de consulta (principalmente dicionários em papel) e recursos de pesquisa (acesso irrestrito à *Internet*). Como vimos, os textos traduzidos também foram variáveis controladas do experimento.

	NP (8 sujeitos)	NC (7 sujeitos)	PC (6 sujeitos)	TOTAL
Folha de perfil	8	7	6	21
Autorização para uso dos dados	8	7	6	21
Questionário de idéia sobre tradução	8	7	5	21
Arquivo TXT do questionário de Leitura em LM	8	7	5	20
Arquivo LOG do questionário de Leitura em LM	8	7	5	20
Arquivo TXT da redação em LE	8	7	5	20
Arquivo LOG da redação em LE	8	7	5	20
Pequeno questionário sobre a redação em LE	8	7	5	20
DADOS DE PRODUTO E PROCESSO EM TRADUÇÃO				
Arquivo TXT de T1	8	7	6	21
Arquivo LOG de T1	8	7	6	21
Transcrição do TAP retrospectivo e da entrevista de T1	8	7	6	21
Arquivo TXT de T2	8	7	-	15
Arquivo LOG de T2	8	7	-	15
Transcrição do TAP retrospectivo e da entrevista de T2	8	7	-	15
Arquivo TXT de T3	8	7	-	15
Arquivo LOG de T3	8	7	-	15
Transcrição do TAP retrospectivo e da entrevista de T3	8	7	-	15
Arquivo TXT de T4	8	7	-	15
Arquivo LOG de T4	8	7	-	15
Transcrição do TAP retrospectivo e da entrevista de T4	8	7	-	15

DADOS DE PRODUTO E PROCESSO EM TRADUÇÃO (POR TIPO)	T1	T2	T3	T4	TOTAL
Arquivos TXT de traduções (dados de produto)	21	15	15	15	66
Arquivos LOG de traduções (dados de processo)	21	15	15	15	66
Transcrições de TAPs retrospectivos (dados de processo)	21	15	15	15	66

Tabela 1: Quantificação das partes do *corpus*

A coleta do *corpus* longitudinal foi, como dissemos, sincronizada com uma disciplina-piloto de introdução à tradução direta escrita, cujo desenho foi inspirado na definição de Hurtado (1996) para um curso de introdução geral à tradução como um “espaço didático no qual o estudante deverá captar os **princípios fundamentais**, começando a assumir um método de trabalho adequado”. Foram feitas várias leituras sobre características de processamento tradutório de novatos em contraste com o de sujeitos mais competentes em tradução, bem como sobre o desenvolvimento da CT. A partir das informações obtidas nessa revisão da literatura e a partir da consideração de casos problemáticos de tradução, selecionamos aqueles ‘princípios tradutórios’ que pareceriam ser os mais decisivos na detecção de problemas e nas tomadas de decisão no processo de tradução. Esses foram os ‘princípios fundamentais’ que foram especialmente focalizados no trabalho realizado na disciplina-piloto. Receberam destaque certos princípios funcionais, notadamente o do princípio de finalidade e a noção da tradução como operação intercultural, e não apenas interlinguística. O princípio de finalidade, enunciado em 1984 por Reiss & Vermeer, num livro clássico das abordagens funcionais da tradução, propõe um modo de entender a tradução diferente daquele baseado na equivalência que o texto-meta deveria estabelecer com o texto-fonte. No trabalho daqueles autores, defende-se que a tradução não é apenas uma operação determinada pelo foco no texto-fonte, mas tanto ou mais pela adequação do texto-meta às necessidades peculiares que determinam sua produção. Assim, os parâmetros para os métodos e estratégias utilizados ao traduzir, em primeiro lugar, se definem em função da finalidade que o texto traduzido deverá cumprir na cultura-meta. O chamado ‘princípio de finalidade’ é enunciado por Reiss & Vermeer como nos exemplos a seguir: “o princípio dominante de toda tradução é sua finalidade” (p. 80); “é mais importante que um *translatum* (uma translação) atinja um objetivo dado do que o modo determinado em que se realiza” (p. 84). De acordo com esse princípio, Vermeer considera o seguinte:

[...] suponhamos que se queira traduzir o Gênesis com a função de texto ritual [...]. É importante reproduzir o texto o mais literalmente possível; seu sentido é secundário. Suponhamos que se queira traduzir a Bíblia com uma função estética. Será mais importante atingir um valor estético, de acordo com as expectativas da cultura-meta (!), do que reproduzir o texto literalmente. Suponhamos que se queira traduzir a Bíblia com a função de texto informativo. O importante será que fique claro o sentido do texto (na medida do possível) [...] Portanto, não existe A (única forma de realizar uma) tradução de um texto; os textos-meta variam dependendo da finalidade que se pretenda atingir. (p. 84)

As implicações do princípio de finalidade sobre as decisões ao longo do processo de tradução fica especialmente evidenciada, por um lado, em certos momentos em que há distâncias culturais que impossibilitam pensar em ‘equivalência’ entre certas unidades de tradução (ex.: caso dos termos culturalmente marcados, como *feijoada*), e, por outro lado, em certos tipos de tradução em que um código não verbal (icônico ou acústico) convive em estreita simbiose com o código verbal, como no caso das histórias em quadrinhos, das palavras cruzadas, das falas de personagens de um filme, de cartazes publicitários, da canção... Ao longo da disciplina-piloto, o uso de tarefas que envolviam esse tipo de tradução foi frequente, para favorecer *insights* sobre a importância do princípio de finalidade em tradução, com a intenção de ir deslocando o foco dos sujeitos participantes, de modo que se afastassem da tradução puramente baseada na equivalência, em direção à tradução voltada para a finalidade do texto-meta, o que, em nosso entender, significaria uma mudança-chave na ‘subcompetência estratégica’, a subcompetência central da CT, segundo o modelo componencial de CT do grupo PACTE (2001). A subcompetência estratégica é definida pelo PACTE como sendo aquela que controla o processo tradutório e que tem entre suas funções: (1) traçar o projeto e percorrer o processo tradutório (**escolher o método mais adequado**); (2) avaliar o processo e os resultados parciais obtidos **com relação ao propósito final**; (3) ativar as diversas

subcompetências e compensar deficiências nelas; (4) **identificar problemas de tradução e aplicar procedimentos para resolvê-los.**

Devido ao grande volume de dados do *corpus* total coletado, as primeiras análises feitas centraram-se num recorte de características de adequação funcional, em nível de produto (sem levar em conta os *sub-corpora* de dados puramente processuais), mais precisamente, em certos pontos das traduções T1, T3 e T4, pontos esses que permitiam observar a captação do princípio de finalidade ao longo da disciplina introdutória. Para isso, utilizamos a seguinte metodologia, desenvolvida para este estudo: (a) foram selecionados 18 fragmentos, nos textos da primeira e das duas últimas traduções, como indicadores de componentes funcionais da subcompetência estratégica, fragmentos que foram numerados para referência numa tabela do *Excel*, de modo a facilitar cálculos a partir da tabulação de valores atribuídos manualmente à tradução de cada fragmento por cada sujeito; (b) foram estabelecidos critérios para analisar o desempenho de cada sujeito na tradução dos fragmentos selecionados; (c) a partir das quantificações em tabelas do *Excel*, foi analisado o desempenho funcional dos grupos (ao longo do tempo, no caso dos estudantes), atribuindo a cada grupo um valor percentual de 'adequação funcional' atingida em cada uma das tarefas de tradução (ver tabelas no *Anexo*).

Seguem algumas amostras dos 18 fragmentos observados. Nelas, o público-alvo (crianças brasileiras) proposto para os sujeitos implicaria privilegiar certas características contextuais ou textuais, prevalecendo sobre os critérios de equivalência semântica ao traduzir, como nos seguintes casos: (1) fragmentos rimados nos quais a manutenção da rima deveria prevalecer sobre a tradução do significado preciso das palavras em relação de rima: "*Los dos llegaron al templo en monopatín y luego dieron una fiesta en el jardín, una fiesta que duró diez días y un enorme chupetín*"; (2) fragmentos com elementos culturalmente marcados: "... *bailaron la rancherita y el pericón*" (danças folclóricas argentinas, que, devido ao público, pediriam adaptação e 'domesticação').

A identificação dos problemas funcionais nesses trechos e a apresentação de soluções adequadas ao público-meta, de acordo com a finalidade da tarefa de tradução proposta, implicariam que a subcompetência estratégica havia funcionado de modo a procurar equivalências conferindo alto valor hierárquico a componentes (con)textuais e funcionais.

A metodologia desenvolvida para avaliação quantitativa dessa capacidade de adequação funcional envolveu dois tipos de análises. No primeiro tipo, examinamos se, nas traduções (dados de produto), os sujeitos mostravam (valor 1) ou não mostravam (valor 0) evidências de detecção da necessidade de adequação funcional em cada um dos 18 fragmentos, tentando ou não manter as rimas nos fragmentos rimados e procurando ou não adaptar os elementos culturalmente marcados, de modo a poderem fazer sentido no universo da criança brasileira (*Tabela 2* no anexo, gráfico resultante na *Figura 1*). O segundo tipo de análise envolveu uma atribuição de valores de 0 a 3, segundo a qualidade das soluções oferecidas aos trechos selecionados (*Tabela 3* no anexo, gráfico resultante na *Figura 2*). Os dois tipos de análise mostraram que, entre o início e o final da disciplina, os sujeitos do grupo principal (que participaram da disciplina) tiveram um incremento claro na capacidade de detectar problemas funcionais-contextuais, assim como melhorou visivelmente sua competência para oferecer soluções de qualidade para esse tipo de problema. Nesse aspecto da habilidade tradutória, nas tarefas de tradução finais (T3 e T4), o grupo principal (NP nos gráficos) chegou a igualar-se em desempenho aos profissionais (PC), conforme se vê nos gráficos gerados a partir das tabelas do *Excel* (*Figuras 1 e 2*).

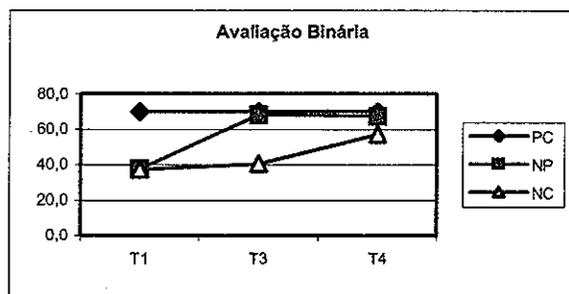


Figura 1: Gráfico binário

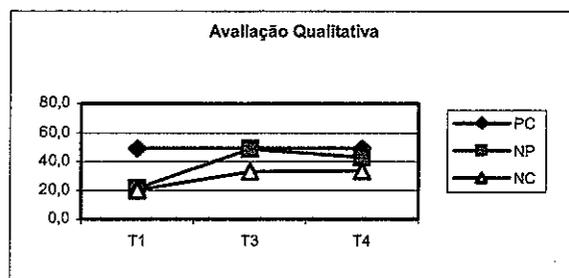


Figura 2: Gráfico qualitativo

As conclusões dessas análises foram que os dados: (1) oferecem evidências empíricas de que a CT se diferencia qualitativamente da competência bilíngüe, e indicam que a capacidade de detecção e resolução de problemas funcionais tem importante papel nessa assimetria; (2) evidenciam que componentes funcionais da subcompetência estratégica respondem rápida e eficazmente às propostas metodológicas da disciplina-piloto, mesmo entre sujeitos ainda não proficientes na L2.

Para encerrar, acreditamos que este estudo possa contribuir com uma proposta metodológica de análise para *corpora* de pequenas dimensões, longitudinais, de aprendizes de tradução. Essa metodologia parece ter-se mostrado sensível a importantes mudanças na subcompetência da CT tida como central no modelo do PACTE, sendo,

portanto, de interesse para estudos empírico-experimentais da CT e de seu desenvolvimento, com auxílio de *corpus*. As características quantitativas dessa proposta metodológica talvez possam, ainda, ser adaptadas para a análise de outros tipos de problemas de tradução, cobrindo outros aspectos que se queira observar no desenvolvimento da CT.

Referências bibliográficas

- BAKER, M. (ed.) *Routledge Encyclopedia of Translation Studies*. London: Routledge, 1998. p. 277-280.
- BAKER, M. *Corpus linguistics and translation studies: implications and applications*. In: BAKER, M.; FRANCIS, G.; TOGNINI-BONELLI, E. (ed.). *Text and technology. In honour of John Sinclair*. Amsterdam / Philadelphia: John Benjamins, 1993. [apud LAVIOSA 2004]
- CINTRÃO, H. P. *Colocar lupas, transcriar mapas. Iniciando o desenvolvimento da competência tradutória em nível básico de espanhol como língua estrangeira*. 2006. Tese (Doutorado em Letras) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2006. Disponível em <<http://www.dominiopublico.gov.br>>
- HOLMES, J. S. The name and nature of translation studies. In: VENUTI, Lawrence (ed.). *The Translation Studies Reader*. London: Routledge, 2000. p. 172-185.
- HURTADO ALBIR, A. *Traducción y traductología*. Madrid: Cátedra, 2001.
- _____. La enseñanza de la traducción directa 'general'. In: *La enseñanza de la traducción*. Castelló de la Plana: Universitat Jaume I, 1996. p. 31-55.
- JAKOBSEN, A. L. Orientation, segmentation, and revision in translation. In: HANSEN, G. (ed.) *Empirical Translation Studies: process and product*. Copenhagen: Samfunds, 2002. p. 191-204.
- LAVIOSA, S. *Corpus-based translation studies: where does it come from? Where is it going?* *TradTerm*, São Paulo, v. 10, p. 29-57, 2004.
- PACTE (Proceso de Adquisición de la Competencia Traductora y su Evaluación). La competencia traductora y su adquisición. *Quaderns. Revista de Traduc-*

	1 1 1 1 1 1 2 3 4 5	2 2 2 2 2 1 2 3 4 5	1 1 1 1 1 1 2 3 4 5	2 2 2 2 2 1 2 3 4 5	3 3 3 3 1 2 3 4	3 1	4 4 4 1 2 3						
PO1	0 2 3 3 2	0 7	1 1 1 0 1	2 7									
PO2	0 0 1 0 2	20	1 0 0 0 0	0 7									
PO3	0 1 0 1 2	27	1 2 2 1 3	0 0									
PO4	2 3 3 1 3	0 0	3 3 3 3 3	1 0 0									
PO5	0 2 1 1 0	27	3 3 3 3 3	1 0 0									
PO6	0 0 0 0 2	1 9	0 2 2 2 3	0 0									
%	11 44 44 33 01 39		00 01 01 50 72	0 9									
NO1	0 0 0 0 0	0	3 2 2 2 3	0 0	2 1 0 0 2	3 3	0 2 1 3 3	0 0	1 3 3 2	7 5	3	2 1 3	0 7
NO4	0 2 1 1 1	3 3	0 0 0 0 0	0	0 1 2 3 2	5 3	1 1 2 2 2	5 3	2 2 0 3	5 8	2	3 1 3	7 8
NO6	0 0 0 0 2	1 3	0 2 2 0 0	2 7	0 1 2 3 2	5 3	3 2 1 3 3	5 0	1 1 0 0	1 7	1	2 2 3	7 8
NO7	0 0 0 0 1	0 7	0 0 0 0 0	0	0 0 0 0 2	1 3	0 0 0 0 0	0	1 0 0 0	0 3	1	2 3 3	0 0
NO8	0 0 0 0 2	1 3	1 0 1 2 0	2 7	0 2 0 1 2	3 3	3 1 1 2 0	4 7	0 0 0 0	0	1	2 2 3	7 8
NO11	0 0 1 0 2	2 0	1 2 1 2 3	0 0	0 1 2 2 0	3 3	1 1 2 2 3	0 0	0 1 3 0	3 5	2	0 1 2	3 3
NO12	0 0 0 0 0	0	1 2 1 2 1	4 7	0 0 0 3 2	3 3	3 1 2 2 1	0 0	0 0 0 0	0	2	1 3 3	7 8
NO13	0 0 0 0 2	1 3	0 0 0 0 0	0	0 0 0 0 2	1 3	3 1 2 2 1	0 0	0 0 0 0	0	2	0 1 0	1 1
	0 0 3 0 3 4 2 4 2	1 5	2 5 3 2 0 3 2 0	3 0	0 3 2 5 5 0 5 6 3 3	3 0	3 0 4 0 7 5 4	3 3	2 1 2 0 2 1 2 4	5 6			5 0 5 0 3 0 4
NO2	0 0 0 0 2	1 3	0 2 1 3 3	0 0	3 0 2 0 1	4 0	1 1 1 3 3	0 0	0 1 0 3	3 3	2	0 1 3	4 4
NO5	0 0 0 0 2	1 3	0 0 0 0 0	0	0 0 0 0 0	0	0 0 0 0 0	0	0 0 0 0	0	0	3 3 3	1 0 0
NO9	0 2 2 2 1	4 7	0 1 1 0 0	1 3	0 1 2 3 2	5 3	1 1 1 1 1	3 3	3 1 0 0	3 3	3	3 3 1	7 8
NO10	0 2 0 3 1	4 0	1 1 1 1 2	4 0	0 2 0 3 1	4 0	3 2 2 2 3	0 0	0 1 0 0	0 0	2	0 0 0	0
NO14	0 0 0 0 2	1 3	0 0 0 0 0	0	0 0 0 0 1	0 7	3 1 1 2 1	5 3	0 0 0 0	0	0	0 0 1	1 1
NO15	0 0 0 0 0	0	1 0 0 0 0	0 7	0 0 0 0 1	0 7	1 0 0 0 0	0 7	0 0 3 0	2 5	0	0 0 0	0
NO16	0 0 0 0 0	0	1 1 1 1 1	3 3	0 0 2 0 1	2 0	1 2 2 2 3	0 7	0 0 0 0	0	3	0 0 2	2 2
	0 1 0 0 5 2 4 3 0 1 8	1 4	2 4 1 0 2 4 2 0	2 7	1 4 1 4 2 0 3 3 2 4 4 0 3 3 4 5 3 2 4 3	1 4	1 4 1 4 1 4 1 4 4 5		2 0 3 0 4 0 3 7				

Tabela II: Avaliação qualitativa – Soluções para os problemas funcionais

Linguística de *corpus* e análise literária: o que revelam as palavras-chave

Lourdes Bernardes Gonçalves*

Resumo: A Linguística de *Corpus*, por ser uma abordagem que permite a manipulação de um grande número de textos, oferece um panorama diferente do que os pesquisadores geralmente se deparam ao analisar um texto literário por meios não-computacionais. Assim, a Linguística de *Corpus* se evidencia imediatamente como uma metodologia extremamente facilitadora do trabalho de pesquisadores e críticos literários, pois vai além de apenas listar palavras, fornecendo, entre outras possibilidades, palavras-chave dos textos estudados e linhas de concordâncias. Neste trabalho mostraremos alguns exemplos das informações que podemos obter observando as palavras-chave obtidas numa comparação entre um *Corpus* de Estudo – no caso a obra *Dubliners*, uma coletânea de 15 contos de James Joyce (1914) – e um *Corpus* de Referência, composto de contos dos autores Katherine Mansfield, D. H. Lawrence e Virginia Woolf, três autores contemporâneos a Joyce, de reconhecida estatura literária. Para tanto, utilizaremos o programa *WordSmith Tools*, de Mike Scott (versão 3, 1998) e as ferramentas *KeyWord* e *Concord*. Uma análise cuidadosa da lista de palavras-chave será realizada e mostrará, algumas vezes de maneira inesperada, que essas palavras-chave podem revelar peculiaridades curiosas sobre o texto que

* Universidade Federal do Ceará