



Universidade do Estado do Rio de Janeiro

Centro de Educação e Humanidades

Instituto de Letras

Barbara Cristina Marques Pereira Ramos

**Corpus de Aprendizes de Tradução: possíveis aplicações na sala de aula de
uma disciplina de tradução**

Rio de Janeiro

2016

Barbara Cristina Marques Pereira Ramos

**Corpus de Aprendizizes de Tradução: possíveis aplicações na sala de aula de uma
disciplina de tradução**



Dissertação apresentada, como requisito parcial para obtenção do título de Mestre, ao Programa de Pós-Graduação em Letras, da Universidade do Estado do Rio de Janeiro. Área de concentração: Linguística.

Orientadora: Profª. Dra. Maria Alice Gonçalves Antunes

Rio de Janeiro

2016

CATALOGAÇÃO NA FONTE
UERJ/REDE SIRIUS/BIBLIOTECA CEH/B

R389

Ramos, Barbara Cristina Marques Pereira.

Corpus de aprendizes de tradução: possíveis aplicações na sala de aula de uma disciplina de tradução / Barbara Cristina Marques Pereira Ramos. – 2016.
93 f.

Orientadora: Maria Alice Gonçalves Antunes.
Dissertação (mestrado) – Universidade do Estado do Rio de Janeiro, Instituto de Letras.

1. Tradução e interpretação – Estudo e ensino– Teses. 2. Serviços de tradução – Teses. 3. Tradutores – Teses. 4. Linguística de corpus – Teses. I. Antunes, M. Alice (Maria Alice), 1964-. II. Universidade do Estado do Rio de Janeiro. Instituto de Letras. III. Título.

CDU 82.035(07)

Autorizo, apenas para fins acadêmicos e científicos, a reprodução total ou parcial desta dissertação desde que citada a fonte.

Assinatura

Data

Barbara Cristina Marques Pereira Ramos

Corpus de Aprendizizes de Tradução: possíveis aplicações na sala de aula de uma disciplina de tradução

Dissertação apresentada, como requisito parcial para obtenção do título de Mestre, ao Programa de Pós-Graduação em Letras, da Universidade do Estado do Rio de Janeiro. Área de concentração: Linguística.

Aprovada em 31 de março de 2016.

Banca Examinadora:

Prof.^a Dra. Maria Alice Gonçalves Antunes (Orientadora)
Instituto de Letras – UERJ

Prof.^a Dra. Tânia Mara Gastão Saliés
Instituto de Letras – UERJ

Prof.^a Dra. Adriana Nogueira Accioly Nóbrega
Pontifícia Universidade Católica do Rio de Janeiro

Rio de Janeiro

2016

DEDICATÓRIA

À minha família, que a cada dia me mostra que vale a pena ser melhor. Sempre.

AGRADECIMENTOS

A Deus, por sua grandiosidade e abrigo, e a meus queridos e amados Mestres, por estarem sempre comigo.

A meus pais, Márcia e Robson, pelos valores, inspiração e todo o incentivo. A meus irmãos, Brian, pelo exemplo silencioso de disciplina e dedicação, e Renata, pelo exemplo de responsabilidade e compromisso.

A meu avô Hélio, in memoriam, e a minha avó Adélia, pelo amor que sempre me deram. A meus padrinhos, Solange e Paulo, e a Karla, por estarem presentes em minha vida me apoiando.

A meus amigos: por cada conselho, palavra de afeto e motivação, mesmo os que estão longe. Em especial a meu amigo João Pedro, por ter aceitado a árdua tarefa de revisar esta dissertação e por tê-la feito com toda atenção e carinho. A Bruno, por ter estado ao meu lado durante este período. Não posso deixar de agradecer a minhas amigas e companheiras de mestrado, em especial a Danielle Marques, Kamilla Loivos, Marcela Lanus, Mariana Cerino e Tainara Belusso, pela disposição e parceria, “na alegria e na tristeza”, e por toda motivação e pelos exemplos, que me davam forças para continuar.

A minha querida orientadora, Maria Alice, pela sutileza, amizade e força; pelos conselhos, pelo maravilhoso encaminhamento desde a graduação e pelos puxões de orelha, claro.

Às professoras que compuseram minha banca examinadora, por terem aceitado o convite para participar da minha defesa e pelos comentários atenciosos, e à PUC-Rio, pela oportunidade de cursar as disciplinas de pós-graduação das professoras Helena Martins e Márcia Martins como aluna extraordinária.

À CAPES, pelo apoio financeiro que muito me auxiliou para que concluísse esta etapa com uma caminhada repleta de apresentações em seminários.

Eu me esforço para ser melhor a cada dia, pois bondade também se aprende.

Cora Coralina

RESUMO

RAMOS, Barbara Cristina Marques Pereira. *Corpus de Aprendizes de Tradução: possíveis aplicações na sala de aula de uma disciplina de tradução*. 2016. 93f. Dissertação (Mestrado em Linguística) – Instituto de Letras, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2016.

Os estudos sobre Corpus de Aprendizes de Tradução (CAT) ainda estão em fases preliminares no Brasil, fazendo com que ainda haja um longo caminho a ser percorrido nessa área. Valendo-se dessa informação, esta dissertação visa a contribuir para a área de ensino e formação de tradutores através do desenvolvimento de um CAT na Universidade do Estado do Rio de Janeiro, na tentativa de preencher uma pequena parcela dessa lacuna. As teorias e pesquisas existentes com os Corpora de Aprendizes de Tradução foram elencadas e comentadas nesta dissertação. Para a construção desse CAT, chamado CATUERJ, foram compilados sete textos em inglês e oito traduções para o português feitas por tradutores aprendizes em formação no Escritório Modelo de Tradução da universidade. Esses textos foram divididos em dois subcorpora, denominados CATUERJ-Ing e CATUERJ-Tra, respectivamente. Eles foram alinhados em nível de frases com auxílio de uma ferramenta para alinhamento gratuita, tornando o CATUERJ um corpus paralelo. Em seguida, as teorias e pesquisas existentes na área foram contrapostas aos exemplos presentes no CATUERJ, com o objetivo de criar possibilidades de aplicação na sala de aula de uma disciplina eletiva sobre tradução (Walsh, 2014), a fim de ressaltar a importância de ilustrar e registrar não somente a criação do corpus de aprendizes, mas também sua utilização real (Santos, 2014). Resultados da pesquisa com o CATUERJ apontam para a extensa possibilidade de uso dos exemplos do corpus dentro de diferentes aulas da disciplina, contribuindo para o enriquecimento das atividades e discussão das escolhas feitas pelos tradutores aprendizes.

Palavras-chave: Corpus de Aprendizes de Tradução. Ensino de tradução. Linguística de corpus. Corpus Paralelo. Estudos de tradução com base em corpus.

ABSTRACT

RAMOS, Barbara Cristina Marques Pereira. *Learner Translator Corpus: possible applications in a classroom of introductory translations studies*. 2016. 93f. Dissertação (Mestrado em Linguística) – Instituto de Letras, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2016.

The research on Learner Translator Corpus (LTC) in Brazil is still on early stages. Therefore, the aim of this study is to contribute to translation teaching and translator training areas by designing an LTC in a university in the state of Rio de Janeiro as an attempt to minimize the gap in such areas of research in Brazil. The existing theories and studies related to Learner Translator Corpora have been described and commented here. For the CATUERJ LTC to be designed, seven source texts in English and eight translations to Portuguese have been gathered. These translations were made by trainee translators from the translation office of the university. The texts were, then, divided into two subcorpora called CATUERJ-Ing, covering the source texts, and CATUERJ-Tra, covering the translations to Portuguese. They were aligned in sentence level using an open source aligning tool, turning the CATUERJ into a parallel corpus. Next, the existing theories and studies in the area were compared to examples taken from the CATUERJ, aiming at creating application possibilities in a classroom of introductory translation studies (Walsh, 2014), in order to highlight the importance of illustrating and keeping record not only of the design of the LTC, but also of the real possibilities of use (Santos, 2014). Results deriving from the research with CATUERJ point at the numerous possibilities of using examples from the corpus within different classes, contributing to the enrichment of activities as well as the debate about choices made by the trainee translators.

Keywords: Learner Translator Corpus. Translation teaching. Corpus linguistics.Parallel corpus. Corpus-based translation studies.

LISTA DE ILUSTRAÇÕES

Tabela 01 – Principais características dos CAT pelo mundo.....	21
Tabela 02 – Referências dos textos originais que compõem o CATUERJ.....	39
Tabela 03 – <i>Tokens</i> de textos originais e traduzidos do CATUERJ-Tra.....	40
Tabela 04 – Números de segmentos dos subcorpora CATUERJ-Ing e CATUERJ-Tra	45
Tabela 05 – Exemplos de quebras de frases do inglês para o português no CATUERJ.....	45
Tabela 06 – Ilustração do deslocamento de oito segmentos no par TO05 – TT05	46
Tabela 07 – Condensação de frases do TO05 para o TT05.....	48
Tabela 08 – Aulas selecionadas para a pesquisa e seus respectivos conceitos.....	60
Tabela 09 – Relação do recorte de segmentos de cada par de textos do CATUERJ.....	62
Tabela 10 – Análise dos segmentos do par TO01 – TT01 do CATUERJ.....	63
Tabela 11 – Análise dos segmentos do par TO02 – TT02 do CATUERJ.....	65
Tabela 12 – Análise dos segmentos do par TO03 – TT03 do CATUERJ.....	68
Tabela 13 – Análise dos segmentos do trio TO04/TO05 – TT04 – TT05 do CATUERJ.....	71
Tabela 14 – Análise dos segmentos do par TO06 – TT06 do CATUERJ.....	74
Tabela 15 – Análise dos segmentos do par TO07 – TT07 do CATUERJ.....	77
Tabela 16 – Análise dos segmentos do par TO08 – TT08 do CATUERJ.....	79
Figura 1 – Fluxograma da tipologia de erros do MeLLANGE	91
Figura 2 – Fluxograma da tipologia de erros do CELTraC	92
Figura 3 – Propaganda da <i>American Heart Association</i>	93

LISTA DE ABREVIATURAS E SIGLAS

CAT	Corpus de Aprendizes de Tradução
CATUERJ	Corpus de Aprendizes de Tradução da Universidade do Estado do Rio de Janeiro
CELTRAC	<i>Czech-English Learner Translation Corpus</i>
CLC	<i>Cambridge Learner Corpus</i>
COCA	<i>Corpus of Contemporary American English</i>
ENTRAD	<i>Enseñanza de la Traducción</i>
EscrTrad	Escritório Modelo de Tradução Ana Cristina César
FCE	<i>First Certificate in English</i>
HKUST	<i>Hong Kong University of Science and Technology</i>
ICLE	<i>International Corpus of Learner English</i>
JEFLL	<i>Japanese English as a Foreign Language Learner</i>
JPU	<i>Janus Pannonius University</i>
KOPTE	<i>Korpus studentischer Übersetzungen</i>
L1	Primeira Língua
L2	Segunda Língua
LLC	<i>Longman Learners' Corpus</i>
MELD	<i>Montclair Electronic Language Database</i>
MeLLANGE	<i>Multilingual e-Learning in LANGuage Engineering.</i>
MISTiC	<i>Multiple Italian Student Translation Corpus</i>
NEST	<i>Norwegian-English Student Translation corpus</i>
PELCRA	<i>Polish and English Language Corpora for Research and Applications</i>
RusLTC	<i>Russian Learner Translator Corpus</i>
RuTLC	<i>Russian Translation Learner Corpus</i>
STA	<i>Student Translation Archive</i>
STTS	<i>Student Translation Tracking System</i>
TO	Texto Original
TSLC	<i>Teachers of English Language Education Centre Secondary Learner Corpus</i>
TT	Texto Traduzido
UCL	Universidade de Louvain

UNESP	Universidade Estadual Paulista Júlio de Mesquita Filho
UPF LTC	<i>Universitat Pompeu Fabra Learner Translation Corpus</i>
USE	<i>Uppsala Student English</i>
USP	Universidade de São Paulo

SUMÁRIO

	CONTEXTUALIZAÇÃO: INTRODUÇÃO	13
1	OS CORPORA DE APRENDIZES E CAT	17
1.1	Corpora de aprendizes	17
1.2	Corpora de aprendizes de tradução (CAT)	20
1.2.1	<u>CAT da Universidade de Leipzig</u>	22
1.2.2	<u>PELCRA</u>	23
1.2.3	<u>STA</u>	24
1.2.4	<u>ENTRAD</u>	25
1.2.5	<u>MeLLANGE</u>	26
1.2.6	<u>RusLTC</u>	27
1.2.7	<u>KOPTE</u>	29
1.2.8	<u>UPF LTC</u>	30
1.2.9	<u>MISTiC</u>	31
1.2.10	<u>CELTraC</u>	32
1.3	CAT no Brasil	33
1.4	Comentários finais	34
2	METODOLOGIA DE PESQUISA	37
2.1	O Escritório Modelo de Tradução Ana Cristina César	37
2.2	O CATUERJ e sua construção	38
2.3	A documentação	41
2.4	O alinhamento	42
2.4.1	<u>O software <i>LFAligner</i></u>	44
2.4.2	<u>O alinhamento do CATUERJ</u>	44
2.5	As abordagens com uso de corpus	48
2.6	Comentários finais	50
3	A DISCIPLINA: INTRODUÇÃO AOS ESTUDOS DA TRADUÇÃO I	52
3.1	Aula 03	55
3.2	Aula 04	56
3.3	Aula 06	57
3.4	Aula 07	57

3.5	Aula 08	58
3.6	Comentários finais	59
4	ANÁLISE	60
4.1	TO01 – TT01	63
4.2	TO02 – TT02	65
4.3	TO03 – TT03	68
4.4	TO04/TO05 – TT04 e TT05	71
4.5	TO06 – TT06	74
4.6	TO07 – TT07	77
4.7	TO08 – TT08	79
4.8	Comentários finais	81
	CONSIDERAÇÕES FINAIS	83
	REFERÊNCIAS	87
	ANEXO A – Tipologia de erros do MeLLANGE	91
	ANEXO B – Tipologia de erros do CELTraC	92
	ANEXO C – Propaganda de uma associação de cardiologia	93

CONTEXTUALIZAÇÃO: INTRODUÇÃO

Durante minha vida acadêmica – mais especificamente em minha graduação – tive a oportunidade de estagiar como tradutora do par linguístico inglês-português no Escritório Modelo de Tradução Ana Cristina César (EscrTrad doravante). Nos anos de 2012 a 2014, pude traduzir os mais diversos gêneros textuais, desde capítulos de livros a receitas culinárias. A experiência foi enriquecedora para meu início como pesquisadora na área dos Estudos da Tradução, pois conseguia conectar as pesquisas acadêmicas à atividade tradutória propriamente dita.

Na área dos Estudos da Tradução, a inserção das tecnologias nos contextos de ensino, aprendizagem e prática modificou o panorama da atividade tradutória (WYLER, 2003), criando novas modalidades, como a localização, que é “a prática de tradução de textos em meio eletrônico” (STUPIELLO, 2012, p. 76) por meio de uma compartimentalização de textos de grande volume entre diferentes prestadores de serviço. Tal inserção exigiu que o profissional dominasse novas técnicas como, por exemplo, os programas de memória de tradução. No entanto, mesmo com um aparato tecnológico a meu dispor dentro do escritório, sentia falta de um material para consulta que apoiasse minhas escolhas linguísticas e lexicais durante as traduções.

Já no início do mestrado, com um tema inicialmente voltado para o ensino de tradução em universidades, tive contato com a disciplina de Linguística de Corpus, que me demonstrou ter grande potencial para desenvolvimento de pesquisas na área da tradução. Gradativamente, fui me aprofundando na disciplina, e direcionei minha revisão de literatura para os Estudos da Tradução com Base em Corpora.

Após os primeiros contatos com corpora, optei por mesclar minha nova área de interesse com o que sempre chamou minha atenção – o ensino. Iniciei, então, minha pesquisa que visava à montagem, documentação e anotação¹ de erros de um Corpus de Aprendizes de Tradução (CAT doravante) constituído por traduções de estagiários do EscrTrad onde estagiei na graduação, o CATUERJ. A ideia de fazer a anotação de erros no CATUERJ surgiu da expectativa de que tradutores iniciantes e aprendizes cometem muitos erros² em suas

¹ “Anotar é delimitar um segmento de texto e atribuir-lhe uma etiqueta (ou conjunto de etiquetas)” que contenha informação linguística associada (FREITAS, 2015, p. 36).

² Destaco, nesse ponto, que adoto a visão de erro definida por Maria Paula Frota, sendo ele causado por “ignorância - ou seja, o desconhecimento acerca de alguma estrutura linguística, algum aspecto cultural ou,

traduções. Sendo assim, comecei a observar o corpus de análise, com o objetivo de destacar e anotar os erros dos tradutores aprendizes. O CATUERJ conta com textos originais em língua inglesa e as respectivas traduções para a língua portuguesa. Essas traduções foram feitas por estagiários do par linguístico inglês-português, durante a época em que estiveram ativos no EscrTrad, cuja origem e funcionamento estão explicados na seção 2.1 desta dissertação.

No entanto, ao alinhar o corpus, analisando os textos originais em paralelo às traduções, foi possível perceber que, na verdade, as traduções eram naturais e fluentes, sem uma grande presença de erros que impedissem a comunicação daquele conteúdo da língua-fonte para a língua-alvo – contrariando minha expectativa inicial de que tradutores iniciantes e aprendizes em formação cometeriam muitos erros. Desta forma, optei por enfatizar aspectos positivos das traduções e das escolhas tradutórias dos aprendizes, indo de encontro a meu pensamento anterior. Essa escolha seria uma tentativa de colaborar para a construção de uma imagem mais positiva do tradutor aprendiz e do próprio processo de aprendizagem – que pouco realça o acerto –, salientando não apenas aspectos negativos da tradução, mas enfatizando escolhas tradutórias positivas.

Assim, o trabalho com o CATUERJ se iniciou e amadureceu a partir da pesquisa por exemplos de outros Corpora de Aprendizes de Tradução já existentes no mundo. A nova busca seria analisar o recorte do CATUERJ escolhido para esta dissertação à luz dos usos e aplicações já existentes em pesquisas com outros CAT às quais pude ter acesso através de publicações. Além disso, a análise do corpus também visou relacionar os recortes do CATUERJ com conteúdos trabalhados em uma disciplina eletiva de tradução (WALSH, 2014) buscando ilustrar possíveis aplicações em sala de aula. Então, passei a considerar que a construção e documentação desse corpus poderiam promover a elaboração de materiais de apoio para o ensino de tradução a partir de exemplos retirados do próprio corpus. Esse tratamento do CATUERJ também facilitaria o acesso de outros pesquisadores para que, futuramente, possam fazer uso desse mesmo corpus para outras pesquisas na área de Estudos da Tradução com base em corpora e de Corpus de Aprendizes de Tradução.

Os objetivos desta pesquisa estão divididos em quatro momentos. Em um primeiro momento, enumerarei os CAT que existem pelo mundo com o objetivo de me familiarizar com sua estrutura e pesquisas que já foram feitas com cada um deles. No segundo momento, o objetivo é compilar, organizar e alinhar o CATUERJ, Corpus de Aprendizes de Tradução

ainda, alguma informação relativa ao contexto da tradução, como o seu público alvo, seu veículo etc. - e um certo fenômeno ou funcionamento psíquico inconsciente que é muito comum e que redundava no lapso de língua (...) [este] consiste em uma manifestação do inconsciente e pode ocorrer em qualquer operação que envolva a linguagem verbal” (FROTA, 2006, p. 142).

composto apenas por textos originais e suas respectivas traduções, feitas por tradutores aprendizes da Universidade do Estado do Rio de Janeiro (UERJ). O terceiro momento consiste em selecionar aulas de uma disciplina eletiva sobre tradução (WALSH, 2014) com a finalidade de conhecer os temas trabalhados em cada uma delas. O objetivo do quarto e último momento, que também é o objetivo principal desta pesquisa, parte da análise do CATUERJ para criar possibilidades de aplicação na sala de aula dessa disciplina eletiva sobre tradução (WALSH, 2014), a fim de ressaltar a importância de ilustrar e registrar a criação do corpus de aprendizes e sua utilização real (SANTOS, 2014), procurando ressaltar sobretudo os aspectos positivos das traduções realizadas pelos tradutores aprendizes.

As pesquisas com CAT no Brasil ainda estão começando a ser reconhecidas e, principalmente, a ter seus potenciais de usos e aplicações explorados. Portanto, ao enfatizar as aplicações do CATUERJ em sala de aula, mesmo que em um contexto específico de uma única disciplina, esta pesquisa se mostra inovadora na área de Estudos da Tradução com base em corpora por duas razões principais. A primeira é por ela poder servir como modelo para uma ponte entre a pesquisa com um corpus especializado – que seja construído a partir de textos produzidos por tradutores em treinamento somente na universidade onde a pesquisa é desenvolvida – e aplicações na sala de aula da mesma universidade, mantendo um diálogo direto entre pesquisa e aplicação pedagógica. A segunda, por ter seu foco nos aspectos positivos do processo de tradução dos tradutores aprendizes, buscando colaborar na construção de uma imagem mais positiva do tradutor em treinamento, da tradução de maneira geral, e do próprio processo de aprendizagem. Para isso, procuro enfatizar escolhas tradutórias positivas e tratar erros como dificuldades, que podem ser superadas e consideradas aprendizado no caminho para prática e experiência tradutórias cada vez mais aprimoradas.

Esta dissertação foi organizada em quatro capítulos. No primeiro capítulo estabeleço os pilares teóricos utilizados nesta pesquisa. Apresento e defino corpora de aprendizes e relato brevemente como alguns desses corpora têm sido utilizados em termos de aplicações práticas no ensino e formação de tradutores, baseando-me também em minha experiência como professora particular de língua inglesa. Parto da origem dos corpora de aprendizes para explicar a estrutura dos corpora de aprendizes de tradução, além de discorrer sobre as aplicações e contribuições para o ensino na área de Estudos da Tradução com base em corpus aos quais tive acesso. Por último, demonstro o resultado de minha pesquisa sobre corpora de aprendizes de tradução existentes no Brasil e suas aplicações, além de colocar minhas considerações finais sobre o aporte teórico exposto no capítulo. No segundo capítulo, descrevo os métodos utilizados para o desenvolvimento da construção do CATUERJ, desde a

origem dos textos compilados para o corpus, passando pelo tratamento dado a ele e critérios adotados em seu design e alinhamento, até a abordagem metodológica adotada para a análise dos dados do CATUERJ. O terceiro capítulo detalha a estrutura da disciplina Introdução aos Estudos da Tradução I (WALSH, 2014), descrevendo mais detalhadamente as cinco aulas escolhidas como o recorte de trabalho para esta pesquisa. O quarto capítulo trata da análise dos dados do CATUERJ, relacionando os segmentos escolhidos para serem incluídos nesta dissertação a outras pesquisas já feitas com os CAT existentes no mundo e aos conteúdos das aulas da disciplina Introdução aos Estudos da Tradução I (WALSH, 2014).

Após os quatro capítulos, há uma seção final, a conclusão. Nela, há o resumo desta dissertação, comentários finais e observações sobre esta pesquisa. Em um segundo momento, menciono as limitações enfrentadas e discorro sobre possíveis caminhos a serem trilhados a partir dos achados aqui colocados e da contribuição desta dissertação para a área de Corpora de Aprendizes de Tradução.

1 OS CORPORA DE APRENDIZES E CAT

Para desenvolver esta dissertação, precisei traçar um panorama dos estudos e pesquisas relacionados a Corpora de Aprendizes de Tradução, resenhando os CAT dos quais se têm registro. Inicialmente, apresento os corpora de aprendizes, que são a origem dos CAT. De forma breve, falo sobre os objetivos de pesquisa com corpora de aprendizes e relato como esses corpora têm sido utilizados em pesquisas e aplicados em práticas pedagógicas. Em seguida, descrevo a estrutura de um CAT, seu potencial para pesquisas, aplicações e contribuições para a área de Estudos da Tradução. Resenho cada um dos CAT encontrados ao longo da revisão de literatura, apresentando sua estrutura, principais características e trabalhos que já foram realizados ou ainda estão em andamento. Por fim, discuto a exploração do potencial desses CAT e publicações sobre seus possíveis usos e aplicações e discorro brevemente sobre a análise do CATUERJ, corpus construído e analisado nesta pesquisa. De antemão, enfatizo que a análise do CATUERJ terá relação com o conteúdo do corpus visando possíveis aplicações na sala de aula, e não erros e dificuldades aparentes dos aprendizes – os quais comento de forma mais superficial, justamente por não serem o foco desta pesquisa.

1.1 Corpus de aprendizes

Um Corpus de Aprendizes é uma “coletânea computadorizada de textos autênticos produzidos por aprendizes de uma língua estrangeira”³ (GRANGER, 2002, p. 7, tradução minha). O objetivo principal é reunir material autêntico que habilite o pesquisador a descrever a linguagem de aprendizes, auxiliando pesquisadores a destacar aspectos pedagógicos e educadores, que são as necessidades dos aprendizes (PRAVEC, 2002, p. 81). Mesmo os corpora de aprendizes, que são mais antigos que os CAT, têm um início de publicações relacionadas no final da década de 1990, mostrando quão recentes são os achados nessa área de estudos baseados em corpus.

Logo no início desses estudos, Nadia Nesselhauf (2004, p. 127) afirmou que haveria três possibilidades de se trabalhar com os corpora de aprendizes. A primeira, considerada

³ O texto em língua estrangeira é: “Systematic computerized collections of texts produced by language learners.”

mais direta, seria identificar as dificuldades de um grupo de alunos e trabalhá-las enfaticamente com diferentes materiais. A segunda possibilidade seria mais indireta e problemática, pois trataria de descobrir as implicações para o ensino de segunda língua a partir da descrição e análise do corpus. A terceira consistiria na possível aplicação direta das informações retiradas do corpus de aprendizes na sala de aula, possibilidade considerada rara na época da publicação e ainda não explorada extensivamente hoje em dia (SANTOS, 2014, p. 69).

O corpus de aprendizes pode ser anotado de forma a destacar padrões linguísticos. As anotações podem ser feitas de forma a marcar lema⁴, classe gramatical (*POStagging*, no inglês)⁵, semântica e sintaxe que servem, entre outras razões, para tornar a “troca de informações mais eficiente” (ZANETTIN, 2002, p. 88). A partir de um corpus anotado, o pesquisador pode direcionar sua análise para uma característica específica.

Corpora de aprendizes como o ICLE⁶, LLC⁷ e o CLC⁸ vêm sendo usados em pesquisas baseadas em corpus desde o início da década de 1990. Das possíveis maneiras de se trabalhar com o corpus citadas por Nesselhauf, as duas últimas – mais especificamente descrição e análise do corpus e aplicação das informações do Corpus de Aprendizes em sala de aula – já haviam sido antecipadas por Sylvianne Granger, pois a pesquisadora afirmou dois anos antes que esses Corpora de Aprendizes têm sido utilizados em prol da elaboração de materiais e metodologia (GRANGER, 2002, p. 16). Na área de materiais, os corpora de aprendizes servem para enriquecer informações em dicionários de línguas, gramáticas e livros didáticos.

Para ilustrar a utilização do material didático elaborado a partir de corpora de aprendizes, cito o material que uso em minha prática como professora de língua inglesa. O livro *First Certificate Trainer*, elaborado por Peter May (2012), funciona como uma série de simulados e exercícios preparatórios para o exame internacional de Cambridge chamado

⁴ Lema é a forma da palavra como aparece no dicionário. As palavras “joguei”, “jogou” e “jogo” vêm do lema “jogar”, por exemplo. Disponível em: <http://wmtang.org/corpus-linguistic-glossary/> (último acesso em 03/01/2016).

⁵ Ou “parts-of-speech tagging”. São etiquetas morfogramaticais que apontam a classe das palavras inseridas naquele contexto. Disponível em: <http://wmtang.org/corpus-linguistic-glossary/> (último acesso em 03/01/2016).

⁶ *International Corpus of Learner English*.

⁷ *Longman Learners' Corpus*.

⁸ *Cambridge Learner Corpus*.

FCE⁹. Esse livro utiliza exercícios que foram criados a partir do *Cambridge Learner Corpus*, corpus compilado a partir de respostas de aprendizes de inglês como segunda língua que se submeteram ao exame. Dessa forma, o livro passa exercícios aos novos candidatos baseados nas dificuldades mais comuns de aprendizes que fizeram exames anteriores. Ou seja, as possibilidades de aplicação de uso de corpora de aprendizes em sala de aula, que Nadia Nesselhauf (2004, p. 127) havia considerado raras, atualmente foram incorporadas de forma natural à prática pedagógica de professores de inglês como segunda língua que não necessariamente precisam ser pesquisadores da área de corpus.

Esta pesquisa, que prevê a utilização real de um Corpus de Aprendizes de Tradução, o CATUERJ, não envolve a elaboração de materiais. No entanto, pretendo explorar esse corpus no sentido de possibilitar o desenvolvimento de uma consciência crítica acerca do processo de aprendizagem de tradução e da prática tradutória, em si.

Já para a metodologia, Granger (2002) defende o uso dos dados de corpora de aprendizes em sala de aula, porém com o monitoramento do professor. Os aprendizes produzem textos, os avaliam e trabalham em conjunto, aprimorando suas produções. Esse exercício torna o aluno mais responsável e ativo no processo de aprendizagem. Em sala de aula, aprendizes podem também constatar usos fossilizados da língua estrangeira com base, por exemplo, em comparações de concordâncias problemáticas de textos nativos e não nativos, fazendo com que o aluno perceba as lacunas existentes entre formas em sua língua e a língua-alvo.

Um exemplo de utilização dessa prática pedagógica em sala são atividades desenvolvidas durante as aulas dos cursos de língua estrangeira da UERJ, nas quais os alunos entregam seus textos ao professor, que os redistribui para que os próprios alunos corrijam os textos de seus colegas de turma. Dessa forma, os estudantes estão ajudando os outros a melhorarem sua produção escrita, aprendendo com os erros dos colegas e construindo seu próprio senso crítico em relação a certo e errado na língua que estão aprendendo. Essa dinâmica pode ser relacionada ao uso dos dados de corpora de aprendizes defendido por Granger (2002). O que ocorre durante essa atividade é que os alunos estão informalmente realizando uma análise de erro daquele material que poderia se tornar um corpus de aprendizes para servir como apoio para as turmas seguintes, por exemplo, que executariam atividades semelhantes.

⁹ *First Certificate in English*, que significa Primeiro Certificado em Inglês, em tradução livre. É um exame de qualificação em língua inglesa para falantes não nativos. É de nível intermediário e contempla o nível B2 do Quadro Europeu Comum de Referências para Línguas (CEFR).

Na seção 1.2, a seguir, discorro sobre os Corpora de Aprendizes de Tradução. Primeiro, descrevo características principais de cada CAT, como as línguas envolvidas no projeto; a universidade à qual ele pertence; os pesquisadores responsáveis; a quantidade de textos e/ou o número de caracteres do CAT. Em seguida, apresento os tratamentos que foram dados aos CAT, se houve anotação de erros, elaboração de uma tipologia de erros própria para o CAT, se ele foi documentado e se está disponível on-line. Por último, descrevo, exemplifico e comento pesquisas que tenham sido feitas com esses CAT, caso tenham sido utilizados e a finalidade de aplicações e usos nas pesquisas, apresentando, ao final da seção, o foco da análise desta dissertação, que visa realçar os aspectos positivos do tradutor aprendiz e as possíveis contribuições que sua produção pode ter para a formação de outros tradutores.

1.2 **Corpus de Aprendizes de Tradução (CAT)**

O Corpus de Aprendizes de Tradução (CAT), por sua vez, é uma compilação de traduções de aprendizes, também em formato eletrônico. As traduções podem ser comparadas a textos originais, no caso de corpora paralelos, bem como a traduções dos mesmos textos originais, porém feitas por tradutores profissionais; a textos originais (não traduzidos) da língua nativa dos aprendizes (ZANETTIN, 2014, p. 39); e até mesmo a traduções feitas por outros aprendizes.

O uso dos CAT por tradutores aprendizes durante seu treinamento pode aumentar a consciência dos aprendizes ao aproximá-los do processo da tradução. Por exemplo, ao precisar optar entre diferentes palavras que seriam sinônimas de acordo com o contexto, o aluno que souber manusear um CAT que esteja a sua disposição poderá pesquisar por si próprio qual das opções foi mais popular entre outros aprendizes. Diante de outros exemplos, o aprendiz passa a ponderar outras opções e tende a tornar-se um tradutor com pensamento mais crítico. O CAT, portanto, serviria como uma espécie de ferramenta de apoio para impulsionar uma compreensão mais racional e crítica do processo de tradução e da língua estrangeira. Isto é, em vez de o aluno acertar concordâncias instintivamente durante a prática tradutória, o tradutor aprendiz poderá fazer questionamentos e escolhas conscientes baseadas no conhecimento adquirido através do CAT de forma independente.

Traçando um paralelo entre os Corpora de Aprendizes e os Corpora de Aprendizes de Tradução (CAT), foi possível concluir que os CAT podem ser usados para descrever a

linguagem do tradutor aprendiz, bem como identificar dificuldades que poderiam surgir em suas traduções (BENNISON; BOWKER, 2003; ZANETTIN, 2014, p. 24 e 42). Vale destacar que, exceto nas situações antevistas por Granger (2002) na seção 1.1 desta dissertação, a meu ver, a menos que esses corpora tenham sido submetidos à anotação de erros – por tratar-se de produções de aprendizes –, não poderiam ser utilizados para consulta de cunho gramatical e aspectos linguísticos, visto o grau de desconfiança existente na utilização da língua estrangeira por um aprendiz (GRANGER, 2002, p. 20). Entretanto, uma possibilidade de utilização desses corpora com produções de tradutores aprendizes em sala de aula existiria a partir da interação e do diálogo crítico entre professor e alunos, por exemplo.

Na Tabela 01, abaixo, catalogo os CAT existentes que serão apresentados nesta seção, tendo como base a tabela feita por Kristýna Štěpánková (2014, p. 15). Organizo os CAT em relação a suas principais características: línguas-fonte e alvo do projeto; a condição da língua-alvo para o tradutor aprendiz; a presença de anotação de erro no corpus; e sua disponibilidade on-line.

Tabela 01– Principais características dos CAT pelo mundo

Corpus	Língua-fonte	Língua-alvo	Condição da língua-alvo	Anotação de erro	Disponível on-line
CAT da Universidade de Leipzig	Alemão	Inglês	Estrangeira	Não	Não
PELCRA	Polonês	Inglês	Estrangeira	Sim	Não
STA	Espanhol Francês	Inglês	Materna	Não	Não
ENTRAD	Inglês	Espanhol	Mista	Sim	Sim
MeLLANGE	Dinamarquês Espanhol Francês Inglês	Catalão Dinamarquês Espanhol Francês Inglês Italiano	Materna	Sim	Sim
RusLTC	Inglês	Russo	Materna	Sim	Somente consulta
KOPTE	Francês	Alemão	Materna	Sim	Não
UPF LTC	Inglês	Catalão	Materna	Sim	Não
MISTiC	Francês Inglês	Italiano	Materna	Não	Não
CELTraC	Tcheco	Inglês	Estrangeira	Sim	Não

CAT de Yasmin S. Brito¹⁰	Português	Inglês	Estrangeira	Não	Não
CAT de José T. Oliveira¹¹	Inglês	Português	Materna	Não	Não

Fonte: A autora, 2016.

A seguir, os dados expostos na Tabela 01 são detalhados e explicitados. Cada CAT listado na Tabela 01 é descrito, bem como seus respectivos subcorpora. A resenha abrange características relacionadas a estrutura; tamanho; língua ou par de línguas envolvidas; e pesquisas que já foram ou ainda estão sendo desenvolvidas a partir desses CAT.

1.2.1 CAT da Universidade de Leipzig

A primeira menção ao uso de um CAT na área de ensino e pesquisa dos Estudos da Tradução é a do corpus de Robert Spence da Universidade de Leipzig. Iniciado em 1988, conta com 49 textos originais em alemão e suas respectivas traduções para o inglês, feitas por aprendizes. Com mais de 1.200 traduções, é possível concluir que foram feitas diversas traduções para os 49 textos originais.

Spence (1998, p. 11) tinha objetivos de curto, médio e longo prazos a alcançar com esse corpus. Em curto prazo, seu objetivo era projetar um curso baseado em suas descobertas, com a contribuição do corpus. Seu objetivo em médio prazo era avaliar a qualidade de traduções e de sistemas de memórias de tradução. Já em longo prazo, Spence almejava ensinar tradução ou língua estrangeira, auxiliado pelo computador. Em sua apresentação em 1998, na Universidade de Hong Kong, Spence cita tais objetivos. No entanto, não houve mais publicações a respeito da utilização desse CAT.

1.2.2 PELCRA

¹⁰ O CAT de Yasmin de Sousa Brito, da UNESP, foi inserido na Tabela 01 para fins de identificação.

¹¹ O CAT de José T. Oliveira, mestre pela USP, foi inserido na Tabela 01 para fins de identificação.

Depois do corpus compilado por Spence (1998), o segundo Corpus de Aprendizes de Tradução do qual temos registro é o PELCRA¹². O projeto PELCRA foi desenvolvido na Universidade de Lodz entre os anos de 1997 e 2005.

O objetivo inicial era construir um Corpus de Aprendizes de Tradução. No entanto, devido a demandas da própria formação de seus universitários, os pesquisadores terminaram por construir dois corpora dentro do projeto: um corpus monolíngue polonês e um corpus de aprendizes de inglês. Desta forma, os alunos teriam um suporte para tomar decisões relacionadas à tradução ao consultar o CAT, evitando erros já cometidos por outros aprendizes, bem como teriam um ponto de referência linguística ao usar o corpus monolíngue polonês.

O PELCRA tem um total de três milhões de palavras que estão divididas em três seções (UZAR, 2008, p. 239-240): a primeira seção conta com três textos originais em inglês, sendo um artigo de jornal, um artigo de revista e um documento sobre contabilidade. Na segunda seção constam três textos originais em polonês com conteúdo equivalente aos textos da primeira seção. A terceira seção é composta por sessenta traduções dos originais da segunda seção, no par linguístico polonês-inglês, feitas por aprendizes – totalizando 180 traduções.

Essa organização torna o PELCRA um corpus comparável, o que permite, por exemplo, que o pesquisador faça análises de traduções de vários alunos comparadas a um texto original em inglês. Outra possível comparação seria entre traduções de diferentes textos feitas pelo mesmo tradutor aprendiz, por exemplo (UZAR, 2008, p. 240).

Os dados do corpus PELCRA começaram a render material de pesquisa já em 1997, com aplicações nas áreas de ensino de línguas, tradução, lexicografia e processamento de linguagem. No site do projeto¹³ estão disponíveis listas de corpora, dicionários eletrônicos e ferramentas de busca e de processamento de linguagem – todos desenvolvidos por pesquisadores do PELCRA. No ano de 2008, o corpus PELCRA passou a constituir o Corpus Nacional do Polonês.

¹² *Polish and English Language Corpora for Research and Applications* – Corpora Inglês-Polonês para Pesquisas e Aplicações (tradução minha).

¹³ Disponível em: http://pelcra.pl/new/pelcra_corpus (último acesso em 16/12/2015).

Com todo esse produto transformado em material de pesquisa, o PELCRA poderia ter-se tornado um dos mais confiáveis meios para pesquisa com base em corpus no mundo, tanto pelo conteúdo do corpus em si, quanto pelo potencial para exploração em termos de análise linguística, elaboração de material ou mesmo para servir como exemplo para criação de CAT envolvendo outras línguas. O PELCRA também possui uma tipologia de erros própria, à qual não temos acesso. Uma característica diferenciada dessa tipologia é a existência de uma etiqueta para *feedback* positivo, com a qual são marcadas escolhas tradutórias feitas pelo tradutor aprendiz que são avaliadas como bem-sucedidas pelo pesquisador (UZAR, 2008). No entanto, seu conteúdo não está disponível on-line, impossibilitando o acesso de outras pessoas que não sejam os pesquisadores envolvidos no projeto.

1.2.3 STA

O STA¹⁴ é um CAT paralelo cuja compilação teve início em 1997 pelos pesquisadores Lynne Bowker e Peter Bennison. O motivo para a criação desse CAT veio dos corpora de aprendizes de língua estrangeira existentes à época. Eles despertaram nos pesquisadores a ideia de compilarem “uma coleção eletrônica de traduções de aprendizes” (BENNISON; BOWKER, 2003, p. 104).

O STA abriga textos das áreas jurídica e médica escritos originalmente em espanhol ou francês e traduzidos para o inglês, língua materna dos tradutores aprendizes. Para a compilação do corpus, Bennison e Bowker (2003) desenvolveram um formulário no qual constavam metadados referentes às traduções e aos tradutores aprendizes. Assim, proporcionaram a documentação do STA, o que possibilita o recorte do corpus com base em características específicas dos textos através de um programa chamado STTS¹⁵. Com esse *software*, o pesquisador pode selecionar textos para, por exemplo, acompanhar o desempenho de um único aluno em suas traduções ao longo de um período de tempo; ou analisar todas as traduções provenientes de um mesmo texto-fonte.

É fundamental destacar que o STA não possui anotação nem uma tipologia de erro própria. No entanto, fica evidente no artigo “Student Translation Archive and Student

¹⁴ *Student Translation Archive*.

¹⁵ *Student Translation Tracking System*.

Translation Tracking System: Design, Development and Application” (BENNISON; BOWKER, 2003) que esse não era o objetivo dos pesquisadores. O interesse recaía sobre o progresso feito pelos alunos envolvidos no projeto ao longo de um período. Por isso deram forma ao STTS, priorizando uma extensa documentação dos textos que compoem o STA.

1.2.4 ENTRAD

O ENTRAD¹⁶ é um corpus paralelo que abrange três subcorpora. O primeiro é um subcorpus de textos originais em inglês derivados de jornais e revistas, em sua maioria. O segundo subcorpus é composto por traduções em espanhol produzidas por alunos. O terceiro contém uma tradução para cada texto-fonte (do primeiro subcorpus) que é feita em sala de aula e “negociada” pelo grupo (FLORÉN; LORÉS, 2008, p. 438).

Apesar de o corpus estar disponível on-line para consulta, não se sabe ao certo seu tamanho. A única informação disponível nas publicações relacionadas ao ENTRAD é que há 45 textos-fonte em inglês e as 45 traduções negociadas. No entanto, não é conhecido o número de traduções realizadas pelos aprendizes individualmente, ou seja, as que não foram negociadas em sala de aula.

Os textos do subcorpus de inglês estão documentados em relação a autor, data e lugar de publicação, e quaisquer alterações julgadas necessárias pelos pesquisadores na versão original – relacionadas a tamanho, omissão, mudança ou adição – e o motivo pelo qual a alteração aconteceu. Já no subcorpus de espanhol, a documentação contém informações sobre o tradutor aprendiz, como nome, idade, gênero e sua língua nativa, bem como a data em que a tradução foi enviada ao professor da disciplina.

Segundo Celia Florén e Rosa Lorés (2008, p. 439), a documentação favorece a “busca nas traduções por motivos pedagógicos [...] e torna o material acessível para propósitos de pesquisa [...]”¹⁷. Em seu artigo, Florén e Lorés ilustram a possibilidade de avaliação do desempenho individual do aluno e a investigação do desempenho de tradutores aprendizes estrangeiros comparando o desempenho conforme o gênero.

¹⁶ *Enseñanza de la Traducción*. Disponível em: <http://ice.unizar.es/entrad/> (último acesso em 23/12/2015).

¹⁷ No original: “the search for translations for pedagogical reasons [...] and also makes material accessible for research purposes [...]” (tradução minha).

Atualmente, o ENTRAD também dispõe de sua própria tipologia de erros dividida em cinco categorias (erros pragmáticos, semânticos, gramaticais, ortográficos e estratégicos), cada uma encerrando subcategorias com as devidas explicações e exemplos¹⁸. A marcação dos erros foi feita baseada na própria tipologia criada pelas pesquisadoras do ENTRAD. No entanto, ela se baseia em um sistema de cores e não pode ser interpretada por um programa de análise de corpus devido à incompatibilidade de formato dos textos e códigos de cores, que não podem ser lidos nos formatos geralmente aceitos por esse tipo de programa¹⁹. Isso desaceleraria o trabalho de um pesquisador interessado em analisar erros no corpus, por exemplo.

1.2.5 MeLLANGE

O projeto *MeLLANGE*²⁰ deu forma a um CAT multilíngue alinhado e anotado que é considerado o Corpus de Aprendizes de Tradução mais desenvolvido do mundo (CASTAGNOLI et al., 2011, p. 2). Ele comporta 429 traduções de tradutores aprendizes e profissionais e seus respectivos textos originais cujo tamanho varia em torno de 350 palavras. Os pares de línguas envolvidos nos textos-fonte e alvo são ensinados pelos colaboradores do projeto (inglês, francês, alemão, e espanhol) e provêm de quatro áreas diferentes (Legal, Técnica, Administrativa e Jornalística). O corpus está disponível on-line no site do projeto²¹ e esteve em constante expansão durante os anos nos quais o projeto permaneceu ativo, entre 2004 e 2007.

O objetivo inicial do projeto foi criar um corpus de traduções produzidas por tradutores aprendizes e profissionais que fosse alinhado e anotado. A partir desse tratamento inicial, o corpus foi aprimorado com anotações linguísticas de classe gramatical e de erros. O objetivo na etapa de anotações foi identificar e classificar erros sem foco na avaliação ou na

¹⁸ Disponível em: http://ice.unizar.es/entrad/minisite_tipologia/semanticos.php?pagina=semanticos (último acesso em 26/12/2015).

¹⁹ O formato .txt.

²⁰ *Multilingual e-Learning in LANGUAGE Engineering*.

²¹ Disponível em: <http://mellange.eila.jussieu.fr/index.en.shtml> (último acesso em 14/12/2015).

qualidade do produto²². Desse modo, o corpus poderia contribuir para a formação dos tradutores aprendizes na medida em que promoveria o aumento da responsabilidade e da autonomia ao traduzir. A partir da construção e design do corpus, o objetivo seguinte foi criar uma metodologia que auxiliaria na elaboração de materiais de ensino em plataforma eletrônica nos campos da tradução e da tecnologia, combinados, para adequar a maneira de treinar tradutores e profissionais às novas necessidades do mercado de trabalho (CASTAGNOLI et al., 2011, p. 1).

Dentre os Corpora de Aprendizes de Tradução até aqui ilustrados, o MeLLANGE é o que apresenta a tipologia de erros mais ampla, com duas categorias (conteúdo e linguagem) e 45 tópicos e subtópicos no total (v. Anexo I desta dissertação). O objetivo da criação dessa tipologia de erros foi o de servir para todas as línguas envolvidas no projeto (CASTAGNOLI et al., 2011, p. 5). Em suas subcategorias, a tipologia de erros do MeLLANGE abrange classificações a respeito de forma, como intromissão das línguas-fonte ou alvo na transferência entre as duas, sintaxe, terminologia, registro e variações de gênero; e de conteúdo, como escolhas lexicais impróprias, falsos cognatos e expressões não equivalentes a expressões da língua-alvo. A característica mais notável da tipologia de erros do MeLLANGE, no entanto, não é sua amplitude, mas sim a existência, em cada subcategoria, de um tópico chamado “definido pelo usuário”²³. Esse tópico pode ser escolhido a critério do pesquisador, sempre que o erro encontrado não se encaixar em nenhuma das subcategorias existentes. Considero essa possibilidade de adequar uma tipologia fixa a diferentes evidências nos corpora uma forma muito democrática que garante autonomia ao pesquisador. Assim, o pesquisador pode classificar todo o tipo de erro que não couber nas subcategorias preexistentes sem precisar criar novos tópicos a cada erro encontrado.

1.2.6 RusLTC

O RusLTC²⁴ é um corpus paralelo que abrange textos técnicos originais em inglês e suas traduções para o russo, língua nativa dos aprendizes, e vice-versa. Estima-se haver um

²² Disponível em: http://mellange.eila.jussieu.fr/public_doc.en.shtml (último acesso em 11/01/2016).

²³ *User-defined*, em inglês.

²⁴ *Russian Learner Translator Corpus*.

milhão de *tokens*²⁵, o que o torna o terceiro CAT no mundo a estar disponível on-line (ainda que apenas para consulta), vindo após o ENTRAD e o MeLLANGE (KUNILOVSKAYA; KUTUZOV, 2014, p. 317). O RusLTC tem aproximadamente 260 textos-fonte e 1.795 traduções, com textos de dez gêneros textuais diferentes, dentre acadêmico, informativo, técnico, ficção, educacional, enciclopédia, carta, resenha, discurso e propaganda.

O corpus não tem uma tipologia de erros, o que não é considerado positivo por alguns pesquisadores (ŠTĚPÁNKOVÁ, 2014, p. 19; PRAVEC, 2002, p. 81). No entanto, ele foi documentado em relação a dez itens (KUTUZOV et al., 2012, p. 367). São eles: 1) gênero do tradutor; 2) nível universitário; 3) nota recebida pelo aluno para aquela tradução; 4) rascunho ou versão final; 5) gênero textual; 6) exercício ou avaliação; 7) onde a tradução foi feita (em casa ou na universidade); 8) data da tradução; 9) a qual subcorpus pertence; e 10) universidade que enviou os textos.

Logo, o RusLTC pode ser considerado um CAT de grande representatividade²⁶ na área dos Estudos da Tradução com base em corpus. O RusLTC possui anotação de classe gramatical e, apesar de não possuir tipologia nem anotação de erros, que podem ser criadas posteriormente, ele está disponível on-line e documentado, o que o torna um CAT valioso para pesquisadores dessa área. Pela possibilidade de o RusLTC ser acessado livremente, ele permite que outros pesquisadores explorem sua estrutura e conteúdo, podendo promover pesquisas na área de estudos com base em corpus ou mesmo servir como inspiração para a montagem de um CAT com a mesma estrutura, porém com outra língua-alvo. Outra vantagem do total acesso on-line ao RusLTC é que coincide com o objetivo inicial do projeto, que era “criar um recurso representativo e confiável que possa ser usado em pesquisas na área de Estudos da Tradução e colaborar para a pedagogia da tradução” (KUNILOVSKAYA; KUTUZOV, 2014, p. 315). O objetivo futuro para o RusLTC é que ele possa auxiliar na criação de exercícios individualizados para cada tipo de problema de aprendizagem que influencie no processo de tradução, e que esses exercícios sirvam de materiais de autoconsulta

²⁵ *Tokens* são as palavras dentro de um corpus que podem aparecer repetidas vezes ao longo do texto. Essa nomenclatura é usada na Linguística de Corpus para se referir à contagem de palavras ou ao tamanho do corpus. Disponível em: <http://wmtang.org/corpus-linguistic-glossary/> (último acesso em 03/01/2016).

²⁶ Segundo Almut Koester (2010, p. 70), um corpus representativo é aquele cuja variedade de achados linguísticos converse diretamente com os objetivos do pesquisador que o está analisando. Tony Berber Sardinha (2000, p. 342-345) também define o conceito de representatividade de um corpus, como sendo um critério não objetivo e relativo na etapa da construção. Segundo ele, não é possível definir o que seria uma amostra representativa, por ela depender da frequência de palavras, dos objetivos da pesquisa com o corpus em questão e do sentido das palavras, por exemplo.

e estudo. Acredito que esses exercícios poderão funcionar como uma estratégia para promover a autonomia do tradutor aprendiz.

1.2.7 KOPTE

O KOPTE²⁷ é um corpus paralelo que foi criado em 2009 por Andrea Wurm, da universidade alemã de Saarland, como parte de um projeto do Departamento de Linguística Aplicada, Tradução e Interpretação. O projeto envolveu alunos que puderam escrever suas monografias e dissertações utilizando o KOPTE como corpus de pesquisa, ao passo que contribuíam para a geração da anotação do corpus (WURM, 2013, p. 2).

No KOPTE, o par linguístico envolvido no corpus é o francês-alemão, sendo mais de setenta textos-fonte em francês para mais de 970 traduções para o alemão feitas por aprendizes. Os textos-fonte são, em sua maioria, artigos de jornais. Cada texto conta com até 2.400 caracteres, tornando o KOPTE um corpus médio em relação a seu tamanho (BERBER SARDINHA, 2000, p. 346).

Uma característica diferenciada deste CAT é que tanto a documentação como as anotações do corpus em termos de *POStagging* e outros aspectos linguísticos foram realizadas pelos próprios aprendizes de tradução como parte de suas pesquisas acadêmicas, enriquecendo a experiência dos tradutores em treinamento tanto em aspectos relacionados à prática tradutória como a enfoques teóricos da formação e da área de Estudos da Tradução. Esse envolvimento dos alunos com o corpus já havia resultado em oito publicações oficiais até o ano de 2013 (WURM, 2013, p. 4). Logo, as características mais notáveis do KOPTE são o envolvimento dos aprendizes de tradução com a pesquisa, contribuindo para a pesquisa na área de corpus dentro da própria universidade onde estudam e o número de publicações que se originaram deste projeto.

²⁷ *Korpus studentischer Übersetzungen.*

1.2.8 UPF LTC

O *UPF Learner Translation Corpus* é outro exemplo de CAT paralelo. Começou a ser desenvolvido na Espanha, na Universidade Pompeu Fabra, em 2006, e engloba dez textos-fonte em inglês e 194 traduções. O subcorpus de textos-fonte tem 8.448 palavras e o de traduções tem 200.187. As traduções são feitas para o catalão, língua nativa dos aprendizes (ESPUNYA, 2014, p. 35-36).

Para cada texto-fonte há mais de uma tradução equivalente, podendo variar entre duas a 42 traduções feitas por alunos diferentes. Os aprendizes de tradução cujas traduções fazem parte do UPF LTC cursaram Tradução e Interpretação entre os anos de 2006 e 2011 e estavam no primeiro ou terceiro ano de curso (ESPUNYA, 2014, p. 35).

Esse CAT, assim como a maior parte dos corpora descritos até agora, também possui sua própria tipologia de erros, que não se encontrava disponível on-line nem em publicações sobre o UPF LTC até o término desta dissertação. Ela é composta por 25 categorias simples, sem subdivisões. O UPF também passou por *POStagging*, anotação de função sintática e informações morfossintáticas (ESPUNYA, 2014, p. 36-37). Por possuírem sua própria tipologia de erros e uma anotação linguística detalhada e abrangente, as publicações referentes aos CAT KOPE e UPF LTC em termos de aplicações do CAT em sala de aula analisam erros e problemas de linguagem dos aprendizes (ESPUNYA, 2013; KUNILOVSKAYA; KUTUZOV, 2014, p. 317). Dentre as limitações para a consulta ao UPF LTC está o fato de ele não se encontrar disponível on-line, assim como sua tipologia de erros, como mencionei anteriormente. Isso impede que pesquisadores possam utilizar essa tipologia como base para elaborarem sua própria tipologia, de acordo com o CAT que estão construindo.

Em minha opinião, o grande ponto positivo e diferencial do UPF LTC é a investigação feita a partir da tipologia e anotação de erros do corpus. A pesquisadora Anna Espunya concentra-se nas aplicações do corpus que partem dos erros de tradução envolvendo dificuldades lexicais, como falsos cognatos e escolhas lexicais imprecisas (ESPUNYA, 2013, p. 129). Apesar de ser um trabalho investigando diferenças entre o inglês e o catalão, línguas envolvidas no UPF LTC, ela demonstra que, “ao analisar diferentes dados, o professor pode obter informações valiosas que posteriormente podem ser aplicadas na criação de metodologias específicas para tratar de dificuldades enfrentadas por alunos” (2013, p. 136). Espunya alega, também, que “listas de frequência de palavras conectadas a erros específicos podem revelar não só onde se encontra a dificuldade no processo de aprendizagem

do aluno, mas também podem auxiliar na descoberta de lacunas lexicais ou conceituais entre as duas línguas, além de diferenças nas composições lexicais entre as línguas” (2013, p. 136). Ou seja, a pesquisadora tenta usar a dificuldade do aluno em prol de seu desenvolvimento como tradutor, não apenas apontando o erro, mas o identificando para que, dessa forma, o aluno – e tradutor aprendiz – possa repensar suas escolhas lexicais e aprimorar seu produto final, através do conhecimento do processo de tradução.

1.2.9 MISTiC

O MISTiC²⁸ é um corpus multilíngue alinhado da Universidade de Pisa desenvolvido por Sara Castagnoli (2009) para ser usado como o corpus de análise de sua pesquisa de doutorado. O CAT conta com pouco mais de 480 traduções para o italiano de 59 textos-fonte originais, sendo trinta em inglês e 29 em francês.

As traduções foram feitas por 107 pós-graduandos italianos, estudantes de tradução especializada. O corpus totaliza 176.000 *tokens* na língua-alvo. Ressalto que não há informações em publicações sobre o MISTiC que divulguem o número de *tokens* total do corpus. Entretanto, segundo Tony Berber Sardinha (2000, p. 346), o MISTiC já pode ser considerado um CAT de tamanho médio, apenas pelo número de *tokens* na língua-alvo, o italiano.

Sobre a construção desse corpus, é possível supor que cada texto-fonte possui mais de uma tradução equivalente e que todos os alunos contribuíram com mais de uma tradução, já que as traduções coletadas para esse corpus são provenientes de avaliações feitas em sala de aula. Ou seja, as traduções que fazem parte do MISTiC não foram feitas com esse objetivo. Os textos-fonte utilizados como tarefa de aula que se tornaram o subcorpus de italiano também não foram selecionados especificamente com a finalidade de se tornarem parte desta pesquisa (CASTAGNOLI, 2009, p. 5).

Cada arquivo foi anotado em relação à classe gramatical²⁹. Em seguida, documentado com metadados relacionados a língua, gênero textual, assunto do texto, tarefa tradutória e o

²⁸ *Multiple Italian Student Translation Corpus*.

²⁹ *POS-tagging*.

tradutor. Além disso, cada tradução foi alinhada ao texto-fonte equivalente, bem como a outras traduções do mesmo texto.

O MISTiC não possui uma tipologia de erros, pois foi construído com um objetivo específico: a pesquisa de Castagnoli (2009) foi sobre explicitação³⁰ e requeria apenas um corpus tradicional. Para a análise almejada pela pesquisadora, então, não era necessária a anotação de erros, tampouco a criação de uma tipologia própria.

1.2.10 CELTraC

O CELTraC³¹ é um CAT bilíngue unidirecional³² que foi desenvolvido com o objetivo de auxiliar professores e tradutores aprendizes do departamento de Estudos Ingleses e Americanos da Universidade de Masaryk. A anotação de erros do CELTraC foi feita através da plataforma Hypal que, segundo Kristýna Štěpánková (2014, p. 65), comporta uma interface que possibilita a anotação de erros. Ele é um corpus paralelo composto por dois subcorpora. O primeiro conta com textos-fonte em tcheco, do gênero textual artigo, que falam sobre a herança cultural tcheca. No segundo subcorpus, aparecem as versões em inglês feitas pelos aprendizes de tradução que cursam o mestrado em língua inglesa.

Ao todo, há três textos-fonte em tcheco e 28 versões para cada texto-fonte, totalizando 84 versões para o inglês. Os textos-fonte foram passados como tarefa de casa aos alunos com o objetivo de “desenvolver e praticar habilidades tradutórias relacionadas às humanidades e à mídia” (ŠTĚPÁNKOVÁ, 2014, p. 35).

O corpus foi documentado com informações como gênero do tradutor, área de pesquisa e nível de experiência. A documentação é uma etapa importante da construção do corpus, pois permite que o pesquisador se informe sobre os principais dados dos textos que compõem cada corpus³³. É interessante citar que esse corpus segue uma tipologia de erro que

³⁰ A explicitação ocorre quando o tradutor acrescenta elementos no texto traduzido, afim de evitar ambiguidades. O exemplo mais claro desse conceito é o acréscimo de pronomes quando o tradutor está versando um texto para a língua inglesa, por exemplo.

³¹ *Czech-English Learner Translation Corpus*.

³² Um corpus unidirecional é um corpus constituído por um subcorpus com textos originais somente na L1 e um subcorpus em L2, que são as traduções.

³³ A documentação será explicada de forma mais detalhada na seção 2.3 desta dissertação.

não é própria, e sim uma versão adaptada da tipologia proposta pelo MeLLANGE LTC (v. Anexo II desta dissertação). Entretanto, foram adicionadas algumas categorias e subcategorias para que a tipologia contemplasse os erros achados no CELTraC (ŠTĚPÁNKOVÁ, 2014, p. 65). Na categoria “Língua”, foi adicionada a subcategoria “Determinante errado” e a subcategoria “Coesão e Coerência” foi acrescida à categoria “Conteúdo”. Além disso, Kristýna Štěpánková também acrescentou uma categoria de etiquetas positivas, chamada *Feedback Positivo*³⁴. A criação dessa categoria foi inspirada no PELCRA LTC (v. subseção 1.2.2), com o intuito de anotar soluções tradutórias interessantes.

1.3 CAT no Brasil

Durante minha pesquisa, pude encontrar dois CAT em desenvolvimento no Brasil: o primeiro, criado na UNESP e o segundo, desenvolvido por Joacyr Tupinambás de Oliveira, da USP, como tema de sua dissertação de mestrado, defendida em 2015. Nenhum deles está disponível para consulta on-line.

A respeito do CAT desenvolvido em 2014 pela aluna de graduação e bolsista de iniciação científica Yasmin de Sousa Brito, na Universidade Estadual Paulista Julio de Mesquita Filho, pude encontrar apenas o resumo de seu trabalho no XXV Congresso de Iniciação Científica na própria universidade³⁵. No resumo há informações superficiais sobre o corpus. Baseando-me nessas informações, descobri que o CAT não parece ter um nome; que ele é constituído por versões de textos jornalísticos no par linguístico português-inglês; e que o trabalho realizado com esse CAT tem o objetivo de verificar se colocações que aparecem com maior frequência são utilizadas por falantes nativos de língua inglesa, com o auxílio do COCA, o *Corpus of Contemporary American English*. Não há informações sobre o tamanho do corpus nem sobre o tratamento que a ele foi dado. Tampouco consegui achar outras publicações ou pesquisas envolvendo esse corpus. De qualquer forma, ele serve como uma

³⁴ *Positive Feedback*, em inglês.

³⁵ Disponível em:

http://prope.unesp.br/cic/admin/ver_resumo.php?area=100076&subarea=22618&congresso=35&CPF=40960999884 (último acesso: 11/01/2016).

pequena amostra de que o trabalho com CAT está sendo desenvolvido no Brasil, mesmo que esse trabalho pareça ainda um pouco tímido.

O CAT de J. T. Oliveira (2015), que também não tem nome específico, foi compilado com dois objetivos. O primeiro era a organização e alinhamento do corpus, logo após a compilação. O segundo objetivo seria um levantamento de opções feitas pelos tradutores aprendizes e o que os motivou a fazê-las, tanto lexical como morfossintaticamente. O corpus conta com oito textos em inglês que contêm de duzentas a trezentas palavras e suas traduções para o português, englobando aproximadamente cem traduções para cada texto-fonte. Essas traduções foram feitas por tradutores aprendizes nos primeiro e segundo períodos de uma faculdade particular do estado de São Paulo, entre 2010 e 2013. Para documentar o corpus, o pesquisador traçou um perfil dos participantes com informações relacionadas a faixa etária, sexo, conhecimento de inglês, vivência no exterior, idioma nativo, experiência como tradutor e experiência em outras áreas de conhecimento (contato com outro curso superior por pelo menos dois anos). A característica mais relevante desse CAT para minha pesquisa é que o alinhamento foi feito manualmente e utilizando o programa Microsoft Excel, mesmo programa utilizado para verificar o alinhamento do CATUERJ, inicialmente feito pelo *LFAligner*, que será discutido no capítulo 2 desta dissertação.

1.4 Comentários finais

A partir de uma tabela³⁶ disponibilizada pelos pesquisadores responsáveis pelo RusLTC, tomei conhecimento de outros dois corpora que são denominados Corpus de Aprendizes de Tradução, o RuTLC (*Russian Translation Learner Corpus*) e o NEST (*Norwegian-English Student Translation corpus*). Como explica a própria tabela, esses CAT não foram compilados a partir de materiais produzidos exclusivamente por tradutores em treinamento, mas também por alunos de língua inglesa de universidades e faculdades. Portanto, esses CAT e suas características não foram levados em consideração para esta pesquisa.

³⁶ Disponível em: http://rus-ltc.org/download/LTC_RelatedResearch_March2014.pdf (último acesso em 26/12/2015).

É possível notar que a pesquisa com corpora de aprendizes em diversas áreas de conhecimento vem crescendo e se aprimorando conforme novas tecnologias e *softwares* de análise linguística são criados. A título de exemplo, cito um artigo de Norma Pravec (2002) intitulado *Survey of learner corpora* no qual ela apresenta uma lista dos corpora de aprendizes existentes até a publicação. Na lista, aparecem os seguintes corpora de aprendizes: CLC, HKUST, ICLE, JEFLL, JPU, LLC, MELD, PELCRA, TSLC e USE. Ou seja, dez corpora são enumerados, dentre corpora de aprendizes e de aprendizes de tradução. É interessante, porém, notar que apenas o PELCRA é um Corpus de Aprendizes de Tradução nessa lista.

Pouco mais de dez anos depois do artigo de Pravec, em uma lista disponível na página virtual da Universidade de Louvain (UCL)³⁷, situada na Bélgica, é possível comparar a quantidade de Corpora de Aprendizes à de CAT que foram encontrados pelas pesquisadoras Amandine Dumont e Sylviane Granger. A página havia sido atualizada pela última vez apenas uma semana antes de meu último acesso, indicando estar em constante atualização e sendo coerente com sua descrição. Dentre os mais de 140 Corpora de Aprendizes, apenas os citados na Tabela 01 aparecem como corpora de aprendizes de tradução, e ainda com exceção do CELTraC.

Logo, é possível perceber que as pesquisas com CAT vêm ganhando espaço na área de Estudos da Tradução com base em corpus, bem como na área de Linguística de Corpus. No entanto, na tabela de corpora de aprendizes da UCL, apenas um total menor que cinco por cento dos corpora descritos são de aprendizes de tradução. Isso indica que ainda há um longo caminho a ser explorado na área, tanto em relação à variedade de pares de línguas envolvidos e analisados como CAT, quanto à vasta gama de características passíveis de análise.

Como foi possível perceber, os corpora descritos a partir da Tabela 01 diferenciam-se em termos de abordagens, técnicas e *softwares* utilizados para design, consulta e análise de aspectos específicos. A maioria dos estudiosos que criam os CAT parece optar por corpora com tradução para a língua materna dos tradutores aprendizes com a finalidade de pesquisar fenômenos relacionados à tradução em vez de versão, que seria a opção menos escolhida.

A maior parte das pesquisas com CAT parece ser dedicada à análise de erros e dificuldades linguísticas dos aprendizes nas traduções. Isso se deve, possivelmente, à importância dada ao erro no processo de aprendizagem vista não só no processo tradutório (ESPUNYA, 2013; ZANETTIN, 2002), mas também no de aprendizagem e aquisição de línguas. Na análise desenvolvida nesta dissertação, o foco recai nas escolhas tradutórias

³⁷ Disponível em: <http://www.uclouvain.be/en-cecl-lcworld.html> (último acesso em 07/11/2015).

positivas e para ressaltar outras possibilidades de utilização e aplicação desses corpora. Eles também poderiam, por exemplo, motivar pesquisas sobre estratégias de tradução e escolhas tradutórias a partir de conteúdos trabalhados durante a formação.

Os CAT podem fornecer dados que possibilitem ou facilitem o “desenvolvimento de materiais feitos sob medida para alunos de tradução” (ŠTĚPÁNKOVÁ, 2014, p. 03) ou que sejam um “ponto de partida para trabalhos empíricos” (ESPUNYA, 2013, p. 136). Os CAT são compilados, anotados quanto a erros e classes gramaticais e fornecem conclusões de ordem estatística. No entanto, parece haver uma carência de trabalhos práticos que não sejam sobre erros e dificuldades dos aprendizes. Os pesquisadores que detêm os dados dos CAT parecem preferir a análise de erros à elaboração de exercícios empíricos relacionados ao conteúdo desses corpora, por exemplo. Essa afirmação é fundamentada pelo comentário feito por Sylviane Granger, em uma lista de e-mail para pesquisadores sobre corpora, de que gostaria de recolher mais informações sobre uso real dos corpora, já que muitas publicações apenas descrevem o uso potencial (SANTOS, 2014, p. 70).

Ressalto, portanto, que o foco da minha pesquisa não está nos erros e dificuldades de aprendizes, como fez a maioria dos pesquisadores dos CAT aqui elencados; sim nos acertos. Visando corroborar com a área de ensino de tradução, esta pesquisa debruça-se sobre as contribuições que o conteúdo do CATUERJ³⁸ pode trazer para o ensino de tradutores aprendizes e em treinamento dentro da própria universidade na área de língua inglesa, do mesmo modo que os corpora de aprendizes vêm trazendo para o ensino de língua estrangeira em todo o mundo.

³⁸ O CATUERJ será explicado de forma mais detalhada na seção 2.2.

2 METODOLOGIA DE PESQUISA

Neste capítulo, abordo a construção do CATUERJ, o Corpus de Aprendizizes de Tradução de uma universidade pública do estado do Rio de Janeiro. Descrevo o CATUERJ, suas principais características e as etapas de sua construção. Em seguida, explico como foi feita a documentação do CATUERJ, quais critérios utilizados e quais metadados foram adicionados. Posteriormente, relato o uso do *software* de alinhamento *LFAligner* e ilustro os estágios do alinhamento do CATUERJ. Por último, justifico por que a abordagem guiada pelo corpus foi adotada para essa pesquisa, comparando as duas abordagens existentes – a guiada pelo corpus e a baseada no corpus.

2.1 O Escritório Modelo de Tradução Ana Cristina César

O projeto para criação do EscrTrad foi concebido no ano de 1996 e a inauguração do Escritório Modelo aconteceu em dezembro de 1999. Desde então, o EscrTrad realiza traduções e versões nas línguas alemã, espanhola, francesa, inglesa e italiana para as comunidades interna e externa.

A sala, que faz parte das instalações de uma unidade acadêmica de uma universidade pública do Rio de Janeiro, conta com computadores ligados em rede, impressoras, scanner e outros equipamentos multimídia. Além disso, suas prateleiras contêm um acervo bibliográfico voltado para os Estudos da Tradução, não só com títulos de temas gerais sobre a disciplina, mas também com livros específicos para cada um dos idiomas envolvidos no projeto e obras de referência da língua portuguesa.

A primeira conquista do Grupo de Trabalho que projetou o EscrTrad foi começar a oferecer disciplinas eletivas de Introdução aos Estudos da Tradução de um curso de graduação em Letras³⁹. A oferta de tais disciplinas levou aos estudantes de Letras uma nova perspectiva

³⁹ A partir de 1999, as seguintes disciplinas eletivas foram oficializadas e começaram a ser oferecidas em um curso de Letras da Universidade do Estado do Rio de Janeiro: Introdução aos Estudos de Tradução I e II – Língua Alemã; Introdução aos Estudos de Tradução I e II – Língua Espanhola; Introdução aos Estudos de Tradução I e II – Língua Francesa; Introdução aos Estudos de Tradução I e II – Língua Inglesa; Introdução aos Estudos de Tradução I e II – Língua Italiana.

universitária, com novos objetivos a serem almeçados a partir da possibilidade de uma formação mais diversificada e completa.

Concretizando os três pilares do projeto – Ensino, Pesquisa e Extensão –, além das disciplinas oferecidas na graduação e o trabalho com a prática tradutória, o EscrTrad firmou convênios com outras universidades e conseguiu fomentos de instituições públicas de pesquisa. Sua equipe planeja, organiza e participa de eventos. São oferecidas, por vezes, palestras com tradutores convidados para os bolsistas que atuam no EscrTrad e os professores coordenadores dos idiomas envolvidos. São desenvolvidas discussões a respeito da prática do tradutor como mediador de culturas na sociedade, da importância do resgate cultural na contemporaneidade, e do estágio atual dos Estudos Culturais, Literários, Comparatistas, Pós-Coloniais e de Gênero.

Durante o tempo que atuei como estagiária do par linguístico inglês-português no EscrTrad, de outubro de 2012 a fevereiro de 2014, tive contato com diferentes áreas do mercado de tradução e da pesquisa acadêmica. Participei de dois congressos fora da UERJ ao longo desse período para os quais pude levar a rotina de pesquisas e trabalhos de dentro do escritório para os olhos do público externo da UERJ. Em troca, recebi críticas e sugestões construtivas, ideias inovadoras de pesquisadores que viam de fora, com outra experiência, minha trajetória como estagiária. Esse período de dezesseis meses contribuiu para minha formação acadêmica e pessoal de maneira imensurável, pois posso dizer que meu estágio no EscrTrad foi o despertar do meu interesse por tradução, que resultou no desenvolvimento de diversos trabalhos na área, além do ingresso no mestrado, em março de 2013.

2.2 O CATUERJ e sua construção

O Corpus de Aprendizes de Tradução da UERJ (CATUERJ) é composto por textos originais em inglês e suas respectivas traduções para o português, feitas por estagiários do EscrTrad, entre os anos de 2006 e 2014. Os estagiários cujas traduções foram reunidas trabalham com o par linguístico inglês-português. No EscrTrad, eles executam trabalhos propostos por clientes de fora do Escritório Modelo e também tarefas designadas pela coordenadora de inglês do projeto.

Os trabalhos encomendados ao EscrTrad vêm das comunidades interna e externa. Esses trabalhos podem ser traduções, versões e transcrições (de texto oral para escrito, por

exemplo). O trabalho chega para o estagiário, que discute com o coordenador a melhor forma de executá-lo. O processo de tradução é iniciado, passando pelas mãos do coordenador quantas vezes quanto forem necessárias, até que estagiário e coordenador cheguem a um produto final. A tradução pronta é, então, entregue ao cliente. Esse tipo de trabalho é encomendado por clientes reais – não se tratando de uma simulação.

Por sua vez, as tarefas pedidas pela coordenadora de inglês podem englobar pesquisas, produções de texto e traduções. As pesquisas são na área de tradução e podem ser de caráter teórico ou prático, como sobre cursos de tradução on-line e ementas dos cursos de tradução oferecidos em universidades, bem como sobre publicações com temas específicos na área de tradução e Estudos da Tradução. As produções textuais também têm a tradução como tema na maior parte das vezes, podendo ser uma resenha de um livro ou artigo sobre tradução, ou mesmo um relatório sobre as pesquisas realizadas. As traduções, por sua vez, são atividades tradutórias propriamente ditas, como traduções de artigos científicos; artigos de páginas virtuais; capítulos de livros sobre tradução; e receitas. Os gêneros textuais propostos para que o estagiário realize a tradução são diversificados, pois o objetivo principal das tarefas de tradução é melhorar a prática tradutória do tradutor em treinamento.

As traduções e os textos originais compilados para o CATUERJ foram todos provenientes das tarefas tradutórias definidas pela coordenadora de inglês do EscrTrad, entre os anos de 2006 e 2014, por tratarem de um mesmo tema – tradução. O CATUERJ é um corpus paralelo constituído por dois subcorpora. O primeiro, denominado CATUERJ-Ing, é um subcorpus com os textos em inglês (L2), que foram tirados de capítulos de livros sobre tradução. Na Tabela 02, abaixo, há referências bibliográficas dos capítulos e os respectivos livros que fazem parte deste subcorpus.

Tabela 02 – Referências dos textos originais que compõem o CATUERJ

Nome do arquivo	Referência bibliográfica
TO01	SCHIFFRIN, A. Preface. In: _____. <i>The business of books: How International Conglomerates Took Over Publishing and Changed the Way We Read</i> . Londres: Verso, p. 01-14, 2001.
TO02	VENUTI, L. How to read a translation. In: VENUTI, L. <i>Translation changes everything: theory and practice</i> . Londres: Routledge, 2013.
TO03	BASSNETT, S. The meek or the mighty. In: ÁLVAREZ, R. e VIDAL, M. (Ed.) <i>Translation, Power, Subversion</i> . Clevedon: Multilingual Matters Ltd. p. 10-24, 1996.

TO04 e TO05	BASSNETT, S. Original Sin. In: _____. <i>Reflections on Translation</i> . Bristol: Multilingual Matters, p. 12-15, 2011.
TO06	RAZMJOU, L. To be a good translator. <i>Second International Conference on Critical Discourse Analysis: the Message of the Medium</i> . Iémen, Universidade de Hodeidah, out., 2003. Disponível em: < http://www.translationdirectory.com/article106.htm >. Acesso em: 18/01/2016.
TO07	BASSNETT, S. Translation or Adaptation? In: _____. <i>Reflections on Translation</i> . Bristol: Multilingual Matters, p. 40-43, 2011.
TO08	SHIELDS, K. <i>Gained in Translation: Language, Poetry, and Identity in Twentieth-Century Ireland</i> . Berna: Peter Lang, 2000.

Fonte: A autora, 2016.

O outro subcorpus, o CATUERJ-Tra, é onde estão armazenadas as traduções dos textos que fazem parte do CATUERJ-Ing para o português, língua materna dos estagiários. O CATUERJ, portanto, é um corpus paralelo unidirecional inglês-português.

Como mostra a Tabela 03, os textos em inglês contam com um total de 16.484 *tokens* em sete textos originais diferentes, enquanto os textos traduzidos englobam 18.572 *tokens*, divididos entre oito traduções. Apesar de os textos 4 e 5, no CATUERJ-Ing, aparecerem com números diferentes, trata-se do mesmo texto que foi duplicado como forma de ser relacionado mais facilmente no corpus alinhado com as respectivas traduções. Ao todo, o corpus é constituído por 35.056 *tokens*.

Tabela 03- *Tokens* de textos originais e traduzidos do CATUERJ-Tra

	CATUERJ-Ing	CATUERJ-Tra
Texto 1	1,754	1,823
Texto 2	2,840	2,911
Texto 3	5,727	5,833
Texto 4	1,457	1,325
Texto 5		1,597
Texto 6	1,625	1,909
Texto 7	1,553	1,593
Texto 8	1,528	1,581
Total	16,484	18,472
35,056		

Fonte: A autora, 2016.

Um aspecto que chama a atenção, a partir de um olhar mais detalhado para a Tabela 03, é a pouca diferença do número de *tokens* entre os textos originais e os traduzidos, o que pode indicar a preocupação do tradutor aprendiz em relação à manutenção da quantidade de palavras do original. Além disso, nota-se ainda a quantidade de *tokens* dos TT 04 e 05, que foram traduções de um mesmo original, embora o TT04 tenha quase trezentos *tokens* a mais que o TT05; e o número de *tokens* do TT04 também é menor que o original, fato considerado incomum em textos traduzidos, que tendem a ser maiores que o original em número de *tokens*.

2.3 A documentação

A documentação é uma forma de registrar informações estruturais do corpus, ou seja, dados externos aos textos, como informações bibliográficas, de publicação e catalogação. Segundo a pesquisadora Cláudia Freitas,

a documentação deve incluir informações quantitativas como o número de palavras; informações relativas à proveniência do material compilado (datas de acesso e ou de criação; sites ou locais de onde os textos foram retirados); critérios de seleção da amostra; tipo de texto; registro; e pode conter ainda informação relacionada à autoria dos textos, ao gênero dos autores, e todas as informações que os criadores do corpus considerarem relevantes [...] A explicitação de tais informações possibilita um melhor entendimento dos resultados, tornando mais fácil a comparação com outros corpora, por exemplo. (FREITAS, 2015, p. 34).

A partir da descrição de todos os CAT na seção 1.2, é possível afirmar que a tipologia de erros é um tratamento valioso dado ao corpus em termos de pesquisa, porém não primordial. Por outro lado, como já ressaltado em publicações da área de linguística de corpus (BAKER, 2004; PRAVEC, 2002; REPPEN 2010), a documentação é uma etapa fundamental (FREITAS, 2015, p. 34) no tratamento inicial do corpus. As informações, anexadas ao texto em forma de cabeçalho, facilitam a recuperação do texto, a avaliação da validade daquele corpus para a pesquisa a ser realizada e a geração de outros subcorpora, selecionando textos de um mesmo autor, de certo ano ou de certo gênero. Na documentação de um texto, os metadados da documentação podem anexar ao texto informações externas do próprio texto, como explicita a seção 2.2. Na documentação de um CAT pode haver, também, metadados referentes às traduções e aos tradutores aprendizes. As diferentes documentações dos CAT da seção 1.2 demonstram que são comuns informações sobre o tradutor aprendiz, como nome, gênero, idade, grau de experiência e a data na qual a tradução foi realizada.

Segundo Celia Florén e Rosa Lorés (2008, p. 439), a documentação favorece a “busca nas traduções por motivos pedagógicos [...] e torna o material acessível para propósitos de pesquisa [...]”⁴⁰. A documentação de um CAT também permite que o pesquisador esteja ciente de detalhes inerentes ao processo de criação daquele corpus, melhorando a comunicação entre estudante e corpus. Traçando-se um paralelo com a área médica, a documentação pode funcionar como uma espécie de prontuário, no qual se encontram informações referentes ao passado daquele paciente, sendo o paciente o corpus.

A documentação do corpus CATUERJ-Tra foi feita visando à reutilização do corpus para pesquisas futuras, facilitando a interpretação de resultados por outros pesquisadores de forma mais consistente, não somente pelos criadores do corpus. Para definir os itens a serem inseridos nos metadados da documentação do CATUERJ, adaptei os critérios utilizados na documentação dos CAT RusLTC, MISTiC e CELTraC. Esse processo de adaptação e seleção resultou em oito itens, sendo eles: (a) Título do TO; (b) Autor do TO; (c) Gênero textual; (d) Ano de publicação; (e) Editora; (f) Idioma do TO; (g) Número de *tokens* do arquivo; (h) Número de segmentos alinhados; e (i) Nome do subcorpus do qual o texto faz parte. Nos metadados da documentação do CATUERJ constam informações sobre o texto.

A escolha de não acrescentar informações sobre os tradutores aprendizes foi feita por questões de ordem prática. Seria necessária uma pesquisa mais extensa para recrutar dados sobre os estagiários que realizaram as traduções do subcorpus CATUERJ-Tra. Além disso, essa é uma etapa que pode ser modificada a qualquer momento, podendo o pesquisador adicionar novas informações que lhe sejam úteis para a pesquisa que vai desenvolver.

2.4 O alinhamento

O alinhamento de textos na linguística de corpus pode ser definido como a busca pela correspondência entre segmentos dos textos-alvo e fonte de um corpus paralelo. Ou seja, ao realizar o alinhamento, o pesquisador busca no texto-alvo – a tradução – o trecho correspondente a determinada passagem do texto-fonte – o original.

O alinhamento pode ocorrer considerando diferentes recortes. É de fundamental importância não considerar que a frase seja a unidade de tradução, pois, em diversas

⁴⁰ No original: “the search for translations for pedagogical reasons [...] and also makes material accessible for research purposes [...]” (tradução minha).

situações, a frase pode perder seu sentido original se tirada do contexto (GAUSSIÉ et al., 2000, p. 272). Então, os corpora paralelos podem ser alinhados em nível de palavra, de frase, de sentença, de parágrafo, ou por segmentos não específicos, que podem ser trechos de textos de tamanhos diferentes. Os corpora serão alinhados, portanto, da forma que melhor couber ao objetivo planejado para cada pesquisa.

Os corpora podem ser alinhados manualmente pelo próprio grupo de pesquisa ou automaticamente, com auxílio de um *software*. Tanto o alinhamento manual como o automático têm vantagens e desvantagens. No alinhamento manual, o pesquisador tem a oportunidade de estar próximo de seu corpus em todo o processo, lendo cada segmento de texto a ser alinhado e pode, assim, conhecer melhor seu corpus de pesquisa. A desvantagem está na demora desse processo. O pesquisador precisa ler trecho a trecho tanto no subcorpus de L1 como no subcorpus de L2, consumindo muitas horas hábeis de sua pesquisa.

Este gasto de tempo não acontece quando o corpus é alinhado automaticamente. O alinhamento automático de corpora paralelos, por ser feito pelo *software*, é um processo quase instantâneo. O pesquisador seleciona os corpora ou subcorpora que deseja alinhar, inserindo-os no *software*. Na realidade, essa é a etapa mais demorada. A etapa seguinte, o alinhamento em si, feito pelo *software*, acontece quase imediatamente, com um gasto de tempo menor que um minuto, independente do tamanho e quantidade de corpora ou subcorpora a serem alinhados. Por ser mais rápido, o alinhamento automático é mais usado quando há grande quantidade de material.

Andras Farkas⁴¹, idealizador do *LFAligner* – que será descrito na seção 2.4.1 –, *software* utilizado para alinhar o CATUERJ, afirma que o alinhamento automático tem em média noventa e cinco por cento de precisão. No entanto, Farkas reconhece que textos que foram autoalinhados devem ser revisados pelo pesquisador para certificar-se de que o resultado foi satisfatório⁴². Há, ainda, a possibilidade de os textos não serem adequados para o autoalinhamento, pois pode haver um alto grau de incompatibilidade de segmentos devido à troca de ordem de frases, por exemplo.

Independente do tipo de alinhamento que seja feito no corpus, o alinhamento de textos paralelos é uma forma de obter informações proveitosas que podem ser aplicadas de diferentes maneiras. No campo da atividade tradutória, ele pode ser transformado em conteúdo para bancos de dados de memórias de tradução, por exemplo. Já na área de

⁴¹ Disponível em: <http://www.farkastranslations.com/> (último acesso em 12/01/2016).

⁴² Disponível em: <http://www.farkastranslations.com/alignment.php> (último acesso em 11/01/2016).

terminologia, pode auxiliar a construção de dicionários bilíngues. Por fim, os corpora ou subcorpora alinhados podem, também, servir para a elaboração de material didático para ensino de línguas.

2.4.1 O software LFAaligner

O *LFAaligner* é um *software* desenhado para fazer alinhamento automático de corpora paralelos que fica hospedado no endereço virtual⁴³ da comunidade colaborativa *sourceforge*, cujo objetivo é promover projetos gratuitos para que sejam bem-sucedidos⁴⁴. Sendo assim, o *LFAaligner* é um programa gratuito que funciona em computadores com sistema Windows, Linux e Mac.

Desenvolvido por Andras Farkas, o programa lê arquivos em txt, rtf, doc, docx, pdf, HTML, páginas virtuais e inclusive documentos referentes à legislação da União Europeia e ao Parlamento Europeu. Esses arquivos podem ser extraídos do programa nos formatos .txt, .tmx e .xls. Ao serem inseridos no programa, os arquivos são avaliados automaticamente e o programa informa ao usuário se a realização do alinhamento será possível.

O *LFAaligner* executa o autoalinhamento em nível de sentenças, as quais são denominadas *segments*, ou segmentos, em tradução livre. O usuário tem, também, a opção de converter o alinhamento para a segmentação por parágrafos. É importante ressaltar, porém, que o programa permite ainda que o usuário faça modificações no resultado final. Além disso, sua base de dados compreende uma combinação de mais de oitocentas línguas que podem ser alinhadas.

2.4.2 O alinhamento do CATUERJ

O CATUERJ está armazenado eletronicamente em dois formatos: .doc e .txt. O alinhamento foi feito automaticamente e em nível de sentenças, utilizando o *software*

⁴³ Disponível em: <http://sourceforge.net/projects/aligner/?source=navbar> (último acesso em 17/01/2016).

⁴⁴ Disponível em: <http://sourceforge.net/about> (último acesso em: 17/01/2016).

LFAligner (v. seção 2.4). Os textos e suas respectivas traduções foram inseridos no *software* e, em seguida, relacionados em segmentos de frases e alinhados pelo próprio programa – o segmento do texto original e seu correspondente traduzido. A Tabela 04 mostra o número de segmentos divididos e alinhados automaticamente em cada texto dos dois subcorpora, o CATUERJ-Ing e o CATUERJ-Tra.

Tabela 04 – Números de segmentos dos subcorpora CATUERJ-Ing e CATUERJ-Tra

	Segmentos no CATUERJ-Ing (L2)	Segmentos no CATUERJ-Tra (L1)
TO01-TT01	79	81
TO02-TT02	107	109
TO03-TT03	219	222
TO04-TT04	56	59
TO05-TT05	56	62
TO06-TT06	79	89
TO07-TT07	61	62
TO08-TT08	54	57

Fonte: A autora, 2016.

Após obter os dados da Tabela 04, referentes ao alinhamento automático realizado pelo *LFAligner*, modifiquei os pares correspondentes, igualando o número de segmentos de cada subcorpus. Essa etapa foi realizada manualmente com o auxílio do programa Microsoft Excel, tendo como inspiração a metodologia desenvolvida e adotada por Oliveira (2015), brevemente apresentado na seção 1.3, que construiu seu CAT nessa mesma ferramenta. No entanto, a revisão manual do alinhamento só foi possível devido ao tamanho reduzido do corpus, o que possibilitou a leitura de todas as passagens do CAT tornou viável essa pesquisa. Dessa forma, no subcorpus CATUERJ-Tra há segmentos constituídos por mais de uma frase, para que cada segmento esteja equivalente ao conteúdo e sentido do segmento correspondente do CATUERJ-Ing.

É possível notar que, em alguns dos pares, há uma diferença de até dez segmentos a mais do TO para o TT. Isso ocorre devido à quebra que o tradutor em treinamento faz de uma frase longa em inglês ao traduzi-la para o português. A Tabela 05, a seguir, mostra três exemplos dessa ocorrência em diferentes pares dos subcorpora.

Tabela 05 – Exemplos de quebras de frases do inglês para o português no CATUERJ

Exemplo 1: TO03-TT03
[58] Central to the polysystems approach were certain key assumptions about translation, most crucial of which was the recognition of the role played by translation in shaping the literary polysystem.
[58] Fundamental para a abordagem dos polissistemas eram certas suposições chave sobre a tradução. A mais crucial delas era o reconhecimento do papel desempenhado pela tradução na modelagem do polissistema literário.
Exemplo 2: TO04-TT04
[45] Apart from restricted supplies of manuscripts, scribes to copy them and teachers to disseminate knowledge, not all rules promoted learning, and disease, wars and Viking raiders made consistent study difficult.
[45] À parte de uma oferta restrita de manuscritos, escribas para copiá-los e professores para disseminar o conhecimento, nem todos os governantes promoviam o ensino. Além disso, doenças, guerras e piratas viquingues tornavam difícil o estudo consistente.
Exemplo 3: TO06-TT06
[63] Another important point is that successful translators usually choose one specific kind of texts for translating and continue to work only in that area; for example a translator might translate only literary works, scientific books, or journalistic texts.
[63] Um outro ponto importante a ressaltar é que tradutores consagrados normalmente escolhem um modelo específico de texto para traduzir e tendem a permanecer trabalhando somente nesta área. Por exemplo, um tradutor pode especializar-se em traduzir apenas trabalhos literários, livros científicos ou textos jornalísticos.

Fonte: Segmentos extraídos do CATUERJ, 2016.

A exceção à explicação dada antes da Tabela 05 é o TT05. Ele tem seis segmentos a mais que o TO05, entretanto, ao analisar o par de textos alinhados pelo *LFAligner*, percebi que o motivo do aumento do número de segmentos foi um deslize do tradutor aprendiz, que deslocou um trecho que seria do final do texto mais para o início, fazendo com que os segmentos 06 e 13 – oito, no total – ficassem sem correspondência e, em seguida, tornou a traduzir os segmentos correspondentes ao TO. A Tabela 06 ilustra essa ocorrência.

Tabela 06 – Ilustração do deslocamento de oito segmentos no par TO05 – TT05

Número do segmento	TO05	TT05
[04]	Most translators would immediately opt for the second option, being all too aware of the pitfalls of the word-for-word approach.	A maioria dos tradutores, familiarizados com o perigo do uso da palavra-por-palavra, optaria de imediato pela segunda opção.
[05]	After all, a translation that is too literal can be simply unreadable.	Afinal, uma tradução literal demais pode acabar por ilegível.
[06]		A história da tradução literal através da glosa interlinear dos antigos manuscritos não é apenas uma narrativa especializada para acadêmicos, mas a história do nascimento do inglês escrito.
[07]		De modo parecido, a glosa interlinear em outras línguas europeias deu origem a outras formas escritas de vernáculo.
[08]		Isso significa que a literatura oral que tinha circulado por séculos, como os grandes épicos germânicos, as canções, as charadas e as histórias, podia ser registrada em línguas que, no mínimo, equiparavam-se ao latim, ainda que não o suplantassem em estilo, àquela época.
[09]		É interessante refletir sobre o papel da tradução literal no desenvolvimento das competências linguísticas.
[10]		Comecei menosprezando as versões palavra por palavra e ainda defendo que uma boa tradução vai além do literal.
[11]		Entretanto, a transposição exata de um texto serve a um propósito bem definido.
[12]		Os escribas que faziam as anotações anglo-saxãs eram, acredito, muito melhores linguistas do que meu filho.

[13]		Mas o princípio de alinhamento de palavras, usado para entender como línguas diferentes funcionam, ainda é válido através dos séculos.
[14]	Inexperienced translators seem to go for word-for-word renderings, and it seems to be a universal truth that translation in the tourist industry worldwide is pretty dire.	Já tradutores inexperientes parecem optar pela adoção da palavra-por-palavra, e é quase uma verdade universal que a tradução na indústria de turismo é péssima no mundo inteiro.
[15]	Here are a couple of word-for-word items, one from an Indonesian hotel brochure and the other from a pamphlet produced by the city of Salamanca tourist office:	Encontramos a seguir alguns exemplos do uso da palavra-por-palavra, um retirado da brochura de um hotel na Indonésia, e outro de um panfleto produzido pela secretaria de turismo de Salamanca:

Fonte: Segmentos extraídos do CATUERJ, 2016.

Nesse mesmo par de textos (TO05-TT05), também acontece o oposto do que aconteceu em vários outros pares – a condensação de duas frases do TO em só uma no TT, como exemplifica a Tabela 07.

Tabela 07 – Condensação de frases do TO05 para o TT05

Número do Segmento	TO05	TT05
[25]	If I translate a phrase such as ‘the onset of darkness’, the context will tell me whether ‘darkness’ is being used literally or figuratively, that is, to indicate a state of mind. If it is being used figuratively, then depending on the language into which I am translating, I will have to use a different word from that which renders the physical condition of nightfall.	Ao traduzir uma frase como “a chegada da escuridão”, o contexto me dirá se a palavra “escuridão” está sendo usada de forma literal ou no sentido figurado, e então, dependendo da língua para a qual a estou traduzindo, terei de usar uma palavra diferente daquela que descreve a condição física do anoitecer.

Fonte: Segmentos extraídos do CATUERJ, 2016.

Em seguida, os pares de textos alinhados e revisados foram salvos no formato .xls, ficando disponíveis para consultas futuras através do programa Microsoft Excel. A partir dos arquivos armazenados em formato .xls, o CATUERJ passou a ser um corpus paralelo

alinhado sentencialmente e de fácil compreensão para o pesquisador que se debruçar sobre ele, independente de seu objetivo.

2.5 As abordagens com uso de corpus

Na Linguística de Corpus, há duas abordagens que são utilizadas para pesquisas com corpora. Trata-se, na verdade, de maneiras de relacionar o corpus às teorias. Tais abordagens são denominadas abordagem baseada em corpus (*corpus-based*) ou guiada por corpus (*corpus-driven*).

A abordagem baseada em corpus é apresentada em Tânia Shepherd (2009, p. 103) como uma metodologia que se aproveita do corpus, principalmente para expor ou testar hipóteses e exemplificar teorias e descrições linguísticas pré-existentes. Ela é escolhida quando o pesquisador já tem suas hipóteses e perguntas de pesquisa, e utilizará o corpus para tentar exemplificá-las ou descrevê-las a partir de exemplos daquele corpus.

A abordagem guiada ou direcionada pelo corpus, por sua vez, parte da observação desinteressada de um corpus de pesquisa. A partir dessa observação, o pesquisador dá forma a suas hipóteses e, conseqüentemente, a generalizações a respeito de padrões e frequências lexicais. Na abordagem guiada pelo corpus, há a “referência a dados ou fenômenos que emergem do corpus” (FREITAS, 2015, p. 28).

Minha dissertação teve como norte a pesquisa dirigida pelo corpus. A análise, nessa abordagem, é inteiramente baseada no conhecimento e experiências prévias do pesquisador. Comecei a ler o corpus alinhado para fazer as adaptações de segmentos mencionadas na seção 2.4.2, para que os pares de textos tivessem o mesmo número de segmentos. Sabendo que realizaria meu estágio docente⁴⁵ em uma turma da disciplina Introdução aos Estudos da Tradução no semestre seguinte, já havia lido o artigo chamado “Uma experiência de planejamento de ensino de tradução: conjugando competência e capacidade de renovação na formação de tradutores autocríticos” da professora Bianca Walsh, publicado no periódico *Cadernos de Tradução* em 2015. Ao analisar mais detalhadamente do CATUERJ, percebi que

⁴⁵ Requisito obrigatório para o término do mestrado. Consiste em quinze horas de envolvimento em uma turma de graduação, dentre observações, coparticipações e participações.

havia diversos trechos que poderiam incitar discussões sobre uma das aulas da disciplina, a Aula 03⁴⁶, sobre a importância da interpretação na tradução.

Logo, após o alinhamento do corpus, pude conhecê-lo melhor e trazer à tona a pergunta desta pesquisa, que consistia no envolvimento que o CATUERJ poderia ter com algumas aulas da disciplina Introdução aos Estudos da Tradução, e de que forma. Então, considerando o processo de descoberta do CATUERJ a partir de minhas próprias experiências, posso afirmar que esta dissertação debruça-se na abordagem direcionada pelo corpus, por “fazer referência a dados ou fenômenos que emergem do corpus” (FREITAS, 2015, p. 28) levando-me a construir reflexões e perguntas de pesquisa.

2.6 Comentários finais

O CATUERJ pode ser considerado um CAT pequeno, por conter menos de vinte mil palavras (BERBER SARDINHA, 2000; KOESTER, 2010), especializado e representativo. Considero o CATUERJ representativo de acordo com a noção de representatividade discutida por Almut Koester (2010, p. 70), como sendo a garantia de que o corpus tenha uma variedade de construções linguísticas nele presentes que possam dialogar com o objetivo do pesquisador. Por outro lado, Berber Sardinha (2000, p. 342-345) alega que, quanto maior o corpus, possivelmente maior será sua representatividade. Isso não acontece no caso do CATUERJ pois, por ser um corpus especializado, sua amostra linguística corresponde à compilação de toda a produção dos estagiários do EscrTrad. Logo, o CATUERJ é o mais representativo possível dentro de seu recorte e seu caráter especializado.

Quanto ao CATUERJ ser considerado um corpus pequeno, Randi Reppen (2010) alega que não há um número de palavras fixo para que um corpus seja compilado. Para ela, o tamanho do corpus deve variar de acordo com o propósito para o qual ele servirá. O pesquisador deve apenas se preocupar com as variáveis de praticidade e representatividade. Para considerar a primeira variável – a praticidade –, a preocupação deve ser em relação ao prazo para que a pesquisa seja realizada e finalizada, entre outros detalhes de ordem prática, como datas limites para publicações e emissão de documentos, por exemplo. Já para a

⁴⁶ As aulas selecionadas para serem relacionadas com o corpus nesta dissertação serão explicadas no capítulo 4.

segunda variável – a representatividade –, o pesquisador deve determinar se já coletou textos suficientes para representar o item a ser investigado.

A extensão do CATUERJ, ao longo de sua compilação, pareceu ser suficiente e elucidativa para o desenvolvimento desta pesquisa. Talvez, porém, para uma pesquisa com outro objetivo, seu tamanho não seja satisfatório. Uma característica positiva do CATUERJ, no entanto, é que sua existência está diretamente ligada ao EscrTrad. Portanto, o pesquisador pode compilar outros textos produzidos por estagiários em anos mais recentes ou com outros assuntos, desde que dentro de seu objetivo de pesquisa, e acrescentá-los ao CATUERJ para que produções futuras de estagiários também possam dialogar entre si, bem como com as já existentes no CATUERJ, aprofundando o aprendizado e o conhecimento sobre o processo de tradução e sobre os alunos interessados no estágio em tradução.

Dentre os motivos pelos quais um corpus pode ser especializado, cito os que Lynne Flowerdew (2004, p. 21) apresenta como sendo os principais: a compilação do corpus com o objetivo de investigar um item lexical ou gramatical, em especial; um único gênero textual; um tipo de texto; um assunto; um registro de linguagem; ou um contexto específico, como os participantes, o propósito comunicativo do texto, ou condições particulares para a criação do texto. O CATUERJ, então, é especializado pois foi compilado visando apenas o contexto de traduções realizadas por tradutores em treinamento que atuam como estagiários em um escritório universitário em particular.

Por último, é fundamental ressaltar que o CATUERJ também é representativo no contexto de estudos com CAT no Brasil. Apesar de não poder ser disponibilizado on-line por motivos legais, já que todos os textos originais estão regidos por direitos autorais, o CATUERJ é um dos primeiros CAT a ser compilado e utilizado para pesquisas de forma concreta no Brasil, ilustrando algumas das possíveis aplicações do conteúdo de um CAT na sala de aula de disciplinas voltadas para formação de tradutores.

3 A DISCIPLINA: INTRODUÇÃO AOS ESTUDOS DA TRADUÇÃO I

Neste capítulo, descrevo a disciplina Introdução aos Estudos da Tradução I, disciplina na qual cumpri meu estágio docente. Em seguida, relato os objetivos gerais a serem alcançados durante as aulas e explico o foco do planejamento proposto por Bianca Walsh (2014). Após isso, menciono os pressupostos teóricos adotados como arcabouço para a elaboração dos planos de aula para o programa da disciplina e, nas seções 3.1 a 3.5, refiro-me de forma mais detalhada às cinco aulas escolhidas para participarem da análise dessa pesquisa, tendo seus conteúdos relacionados às traduções que compõem o CATUERJ. É importante ressaltar que, nesta dissertação, os fundamentos teóricos selecionados para a elaboração do planejamento da disciplina Introdução aos Estudos da Tradução I serão explicados neste capítulo de forma superficial. O objetivo ao apresentá-los é apenas contextualizar o leitor, e não estabelecer pressupostos teóricos ou reflexões aprofundadas, por não tratarem diretamente do tema proposto nesta pesquisa.

A disciplina Introdução aos Estudos da Tradução foi inserida no currículo do curso de graduação em Letras de uma universidade pública do Rio de Janeiro em 1999. Foi projetada como uma disciplina eletiva com dois módulos denominados Introdução aos Estudos da Tradução I e II. A disciplina Introdução aos Estudos da Tradução II dá continuidade a temas abordados no primeiro módulo, bem como aborda novos temas para discussão. No entanto, não há obrigatoriedade de cursar as duas disciplinas, apesar de o aluno precisar cursar e ser aprovado na primeira para que esteja apto a cursar a segunda. Ela não será abordada em detalhes nesta dissertação por não ter interesse direto com o recorte selecionado para a pesquisa.

O planejamento da disciplina Introdução aos Estudos da Tradução I feito por Walsh (2014, p. 209) visa a formar tradutores críticos e competentes. Segundo ela,

algumas ações gerais contribuem para alcançar essa meta: (1) que o instrutor trabalhe desvelando crenças sobre a tradução; (2) que os aprendizes sejam levados a pesquisar e a compreender a pesquisa como parte integrante de sua prática; (3) que as aulas sejam orientadas para a resolução de problemas; (4) que o curso deixe um legado teórico sobre os Estudos da Tradução. As ações 1 e 4 atendem ao propósito de construir a visada crítica, a ação 3 atende diretamente à construção de múltiplas competências, já a ação 2 serve simultaneamente aos dois propósitos (WALSH, 2014, p. 209).

É interessante ressaltar que, para cursar essa disciplina, o aluno de graduação precisa ter cursado previamente as disciplinas Língua Inglesa I, II, III e IV como pré-requisito. Isso indicaria que todos os alunos que se inscrevem na disciplina têm um mínimo de experiência acadêmica em comum, não havendo obstáculos de caráter linguístico, como a compreensão de textos em inglês que apareçam ao longo do semestre.

A carga horária da disciplina Introdução aos Estudos da Tradução I é explicitada por Walsh (2014) como sendo quinze aulas no semestre, divididas semanalmente, cada uma com duração de uma hora e trinta minutos. A disciplina foi dividida em quatro momentos. Primeiro, houve um “trabalho de levantamento e avaliação de crenças sobre componentes que envolvem a tradução: o ato tradutório, o tradutor, os atores envolvidos e a sociedade em que se traduz” (WALSH, 2014, p. 210). O segundo momento foi o mais longo; abrangeu nove aulas, nas quais os tradutores aprendizes traduziriam textos e cujo foco era a “construção de um aparato de estratégias” (WALSH, 2014, p. 210). As aulas seguintes, em um terceiro momento, serviriam para apresentar correntes e/ou controvérsias na área de Estudos da Tradução, que seriam “discutidas em confronto com a experiência prática do segundo momento” (WALSH, 2014, p. 210). As duas aulas finais da disciplina, o quarto momento, foram dedicadas a seminários.

Os fundamentos teóricos que permeiam a disciplina iniciam pela reflexão sobre o ensino de tradução, entendendo a língua como atividade integrada a um contexto e comprometida com o uso (WALSH, 2014, p. 210-211). A autora Heloísa Barbosa (1990, p. 11) concebe tradução como “uma atividade humana realizada através de estratégias mentais empregadas na tarefa de transferir significados de um código linguístico para outro”. Bianca Walsh, considerando o contexto da disciplina, decide por afastar-se da ideia de “transferir significados”, optando pelo termo “transposição”, e referir-se a

uma complexa tarefa de movimentação entre dois diferentes sistemas linguísticos, cientes das transformações no processo, que dependem não só da escolha do tradutor, mas das pressões de agentes externos e restrições dos sistemas linguísticos/culturais em contato (WALSH, 2014, p. 211).

A pesquisadora também adota dois imperativos para serem considerados no que tange ao ensino de tradução. O primeiro é a capacitação para a ação de traduzir, juntamente com um repertório de competências, e o segundo é “o fomento a uma perspectiva crítica dessa ação” (WALSH, 2014, p. 212).

Para conduzir o primeiro imperativo em seu planejamento, Walsh opta pela abordagem do Ensino de Línguas Baseado em Tarefas, na qual “a tarefa (de solução de problemas) é a unidade principal de planejamento de ensino” (WALSH, 2014, p. 212). A abordagem do Ensino de Línguas Baseado em Tarefas, bem como a Abordagem Comunicativa foram selecionadas como base para a proposta de ensino desse planejamento, visando à resolução de problemas e ao desenvolvimento de competências para a tradução⁴⁷. Outros autores cujas ideias foram adotadas por Walsh na concepção do planejamento da disciplina foram Heloísa Barbosa (1990) e Fábio Alves, Célia Magalhães e Adriana Pagano (2000) na literatura brasileira e Dorothy Kelly (2005) na literatura estrangeira sobre Estudos da Tradução.

O segundo imperativo foi fundamentado pela reflexão com as crenças relacionadas à esfera da tradução, tanto em aspectos teóricos, como com os conceitos de estrangeirização *versus* domesticação de Lawrence Venuti e patronagem de Lefevere (MARTINS, 2010), como em aspectos da prática tradutória (ALVES; MAGALHÃES; PAGANO, 2000, p. 11). O principal objetivo das atividades que envolviam a reflexão e discussão das crenças era “que diferentes posicionamentos fossem postos em discussão em sala para que os alunos amadurecessem seus próprios posicionamentos” (WALSH, 2014, p. 216). Walsh alega que a maioria das aulas tem como objetivos desenvolver competências e criticidade e que, para essa finalidade, as aulas que têm esses objetivos como foco estão no começo e fim do curso. Segundo ela, “o desenvolvimento da prática é balizado pela visada crítica” (WALSH, 2014, p. 216).

Dentre os objetivos dos autores do livro *Traduzir com autonomia* (ALVES; MAGALHÃES; PAGANO, 2000), dois me tomaram a atenção, por dialogarem diretamente com a proposta de trabalhar com o CATUERJ aplicado em sala de aula. O primeiro diz que a abordagem de tradução do livro é “centrada em estratégias ou ações que conduzem à resolução, de forma eficaz e adequada, de problemas tradutórios” (p. 7). Ou seja, esse objetivo está diretamente ligado aos objetivos da disciplina Introdução aos Estudos da Tradução I, que envolvem o estímulo da pesquisa por parte do tradutor aprendiz para que ele possa construir sua confiança e, assim, poder trabalhar na solução de problemas. O segundo objetivo é o de que o tradutor aprendiz possa desenvolver estratégias de tradução e amplificar sua consciência enquanto tradutor. Esse segundo objetivo pode conectar-se com a intenção da construção de

⁴⁷ As competências principais deste trabalho foram a estratégica; a comunicativa/textual; a cultural/intercultural; e a profissional/instrumental.

múltiplas competências e do ensinamento teórico conferido aos alunos, formando tradutores mais críticos e corroborando com as ideias dos autores do livro, já que, segundo eles,

a ideia de levar o tradutor em formação a desenvolver estratégias de tradução está imbuída do espírito de conscientizá-lo da complexidade do processo tradutório e da necessidade de monitorar suas ações e examinar com cuidado as decisões tomadas ao longo do processo tradutório. A conscientização desse tradutor envolve um redimensionamento do conceito de aprender, o qual passa a demandar que o aprendiz se torne diretamente responsável pelo seu próprio processo de aprendizagem. Em outras palavras, espera-se que o aprendiz se torne autônomo para escolher o caminho mais adequado, para selecionar e gerenciar as ações que melhor respondam a seus interesses e necessidades... (ALVES; MAGALHÃES; PAGANO, 2000, p. 7).

É fundamental destacar os dois critérios que estabeleci para selecionar as cinco aulas da disciplina Introdução aos Estudos da Tradução I, que são explicadas em detalhe nas seções 3.1, 3.2, 3.3, 3.4 e 3.5. O primeiro motivo para selecionar as aulas foi a minha presença. Isto é, em primeira instância, foram selecionadas aulas nas quais estive presente durante meu estágio docente, tendo a oportunidade de observar a discussão dos temas e o decorrer da aula. O segundo critério para escolha referiu-se a aulas com foco na competência comunicativa/textual. Tais aulas foram escolhidas por terem um objeto de discussão que pudesse ser textualmente destacado a partir da análise do texto do CATUERJ sem a dependência de analisar o processo da tradução, como a busca por recursos de apoio externos (WALSH, 2014, p. 218), por exemplo. Para conduzir a análise do CATUERJ, portanto, selecionei segmentos do corpus em que aparecessem um ou mais temas trabalhados nessas aulas, para servirem tanto para exemplo como para questionamentos.

3.1 Aula 03

A Aula 03 tem como foco a interpretação na tradução. Ela trabalha os valores de visada crítica da competência comunicativa/textual. Os recursos para essa aula foram retirados do primeiro capítulo do livro *Traduzir com autonomia* (ALVES; MAGALHÃES; PAGANO, 2000, p. 23-25). Tais recursos consistem em duas propagandas com dificuldades de interpretação que se tornam explícitas na hora de traduzir, por diferentes motivos. Esses textos contêm “ambiguidades, forte ancoragem textual e polissemia” (WALSH, 2014, p. 217). Destaco que o conceito de ancoragem textual não será levado em consideração para a

condução da análise no capítulo 5 desta dissertação, já que ancoragem textual diz respeito à relação texto-imagem (NÖTH; SANTAELLA, 2008, p. 55).

A título de ilustração, cito o bordão de uma das propagandas (v. Anexo III), que pertence a uma associação de cardiologia, na qual se lê “hit the bars after work” logo acima da imagem de um ciclista segurando o guidom de uma bicicleta. O jogo de palavras fundamenta-se no fato de que a palavra “bar” é homônima na língua inglesa. Ela pode tanto significar a barra do guidom, quanto referir-se ao bar como lugar de lazer. No entanto, não pode acontecer o mesmo na língua portuguesa, já que as palavras “barra” e “bar” são diferentes. Sendo assim, os alunos devem discutir possíveis maneiras de interpretação e apresentar às outras duplas e, junto com toda a turma, optarem pela melhor saída.

Durante minha observação dessa aula, os alunos ressaltaram diversas características presentes na propaganda. As principais menções sintetizadas pela professora ao final da discussão foram as relações metafóricas com os trechos “turn out the lights” e “go home”, que foram comparados ao fechamento de um dia de trabalho; a construção de campos semânticos e seus cruzamentos metafóricos – as palavras/expressões “bar”, “happy hour” e “hangover” versus o mundo do esporte –; e a relação da imagem do guidom da bicicleta com o bordão do anúncio – “hit the bars”.

3.2 Aula 04

A Aula 04, assim como a Aula 03, trabalha a criticidade e a competência comunicativa/textual. Seu objetivo geral é “identificar o impacto da variação do contexto e tipo de situação comunicativa no texto” (WALSH, 2014, p. 217). Já os objetivos específicos da aula consistem em considerar “escolhas linguísticas/discursivas; características de gênero; estratégias de persuasão; efeito polissêmico; jogo com as palavras etc.” (WALSH, 2014, p. 217).

A aula ocorre da seguinte maneira: a turma é dividida em dois grandes grupos que recebem textos com diferentes contextualizações. Na aula a que assisti, particularmente, foram distribuídos um e-mail e uma carta formal. Os alunos deveriam identificar o gênero e, em seguida, no grande grupo, apresentar suas interpretações para tentar reconhecer o contexto e o tipo de situação comunicativa de cada correspondência, tanto a escrita manualmente como a eletrônica.

No segundo momento da aula, a professora distribuiu textos de diferentes gêneros. Os gêneros selecionados para essa ocasião foram receita, cardápio, carta informal e e-mail informal. Os alunos formaram grupos e analisaram superficialmente cada texto, tentando identificar os gêneros e suas respectivas situações comunicativas. Ao final da discussão entre os grupos, a turma tornou a se reunir para concatenar ideias e sintetizar as discussões.

3.3 Aula 06

A Aula 06 gira em torno de dois assuntos principais: as características do gênero e de padrões retóricos e o impacto de ambos na tradução. O valor se concentra na competência comunicativa/textual. Para tanto, o professor utiliza como base dois textos retirados do quinto capítulo do livro *Traduzir com autonomia* (ALVES; MAGALHÃES; PAGANO, 2000, p. 78) que compartilham as mesmas informações, porém cada um em um gênero diferente, sendo o primeiro (texto 5.7) pertencente ao gênero informativo e o segundo (texto 5.8) ao gênero “bathroom readers” ou “leitores de banheiro”, em português. A tarefa dos alunos, divididos em pequenos grupos, era comparar os textos, levando em consideração a alternância dos gêneros e a compatibilidade entre gêneros e padrões retóricos. Em seguida, eles levavam a discussão para o grande grupo, compartilhando suas conclusões.

A tarefa seguinte consiste na mesma análise envolvendo “gênero e padrões retóricos utilizados na tradução de um rascunho de carta em português para um ofício em inglês” (WALSH, 2014, p. 219). Na aula da qual participei, essa atividade se fez possível porque a professora retirou o texto para a segunda tarefa de seu arquivo pessoal de tradução. No entanto, percebi que não seria a única possibilidade de realizar a atividade com sucesso, já que outros textos poderiam ser pesquisados na internet ou no arquivo pessoal de tradução de outro professor; podendo inclusive ser uma alternância entre outros gêneros.

3.4 Aula 07

O principal objetivo da Aula 07 era que os alunos percebessem o impacto da estrutura gramatical na tradução. Esse, talvez, tenha sido o tema mais concreto visualmente para os

alunos da turma. O valor de competência comunicativa/textual foi trabalhado enfaticamente durante toda a aula.

No planejamento da disciplina, a proposta é analisar uma versão bilíngue publicada, que é selecionada pelo professor da turma. Essa versão deve conter problemas relacionados à estrutura gramatical, principal foco da aula. Ao distribuir os textos, os problemas são destacados pelos alunos e comentados por eles e pelo professor. Ao final, a tarefa dos alunos é produzir uma versão alternativa.

Como nessa aula realizei uma participação, fui encarregada de selecionar uma versão para o inglês na qual houvesse problemas relacionados a escolhas gramaticais. Para me manter próxima da área acadêmica, optei pelo gênero resumo acadêmico e levei para o encontro o texto original e sua versão para o inglês⁴⁸. O texto, tanto em português como em inglês, continha sérios problemas de tradução, como também palavras traduzidas de maneira errada e até mesmo sem tradução, com a manutenção da palavra em português na versão. Os alunos usaram, em sua maioria, seus telefones celulares pessoais para consultar e fundamentar os problemas constatados nos textos e opinaram fortemente sobre as escolhas feitas na versão. Ao final, os alunos produziram versões alternativas e grande parte deles fizeram melhorias no próprio texto original.

3.5 Aula 08

A Aula 08, junto com a Aula 07, trabalha temas repletos de conteúdo teórico. Na oitava aula do curso, a atenção recai sobre o impacto do léxico na tradução. Sendo essa a última aula a desenvolver a competência comunicativa/textual, os alunos já passam a compreender a dinâmica de tarefas e discussão em grupos. São trabalhadas as noções de colocação lexical, expressão idiomática, polissemia, sinonímia, efeito expressivo, efeito metafórico e prosódia semântica.

A partir de uma propaganda da Shell retirada do capítulo seis do livro *Traduzir com autonomia* (ALVES; MAGALHÃES; PAGANO, 2000, p. 99-100), os alunos “examinam fenômenos lexicais envolvidos na tradução” (WALSH, 2014, p. 220) em pequenos grupos.

⁴⁸ Apesar de o artigo estar disponível on-line, a fonte não será divulgada nesta dissertação por razões éticas.

Em seguida, o professor analisa um texto junto com os alunos para trabalhar as nuances de sentido.

Na aula que observei, a professora perpassou de forma sintética os conceitos a serem trabalhados, dando atenção especial somente para a prosódia semântica, noção sobre a qual os alunos se mostraram mais inseguros. Em seguida, trabalhou um trecho do texto original e da tradução “Rasas da Yoga”, retirado de seu arquivo pessoal. Os alunos comentaram as escolhas lexicais tendo em vista as nuances de sentido e reconheceram a dificuldade da tradução de um texto como aquele, carregado de escolhas lexicais específicas da Yoga.

3.6 Comentários finais

É importante, então, lembrar que o foco da aula é principalmente voltado ao desenvolvimento de competências, à resolução de problemas e à integração da pesquisa como parte de sua prática. Por esse motivo, os alunos recebem previamente toda a ementa e programação do curso, para que possam antecipar as leituras teóricas referentes aos temas a serem trabalhados em cada aula. Sendo assim, durante a aula, a atenção pode recair sobre os aspectos da prática tradutória fundamentada pela teoria.

No capítulo 3, portanto, discorri sobre a disciplina Introdução aos Estudos da Tradução I. Nele, foram expostos os objetivos gerais da disciplina, o foco de seu planejamento e os pressupostos teóricos selecionados para a elaboração dos planos de aula para o programa da disciplina. Por último, foram detalhadas as cinco aulas escolhidas como o recorte para a análise desta dissertação e foram ressaltados os aspectos teóricos trabalhados em cada uma das aulas consideradas durante a análise do CATUERJ. No capítulo 4, a seguir, desenvolvo a análise do CATUERJ relacionada aos conteúdos apresentados nas seções 3.1 a 3.5.

4 ANÁLISE

A análise do CATUERJ à luz das pesquisas já realizadas com Corpora de Aprendizes de Tradução pelo mundo e do planejamento da disciplina Introdução aos Estudos da Tradução I proposto por Walsh (2014) foi realizada em cinco etapas. A primeira, já explicada no capítulo 3, foi a seleção das aulas da disciplina que julguei estarem relacionadas ao CATUERJ. A segunda etapa consistiu na listagem dos conceitos que seriam considerados durante a busca dos segmentos do CATUERJ de acordo com cada aula escolhida da disciplina Introdução aos Estudos da Tradução I. A partir dessa listagem, o terceiro passo foi o recorte de segmentos a serem selecionados em cada par de textos dos subcorpora CATUERJ-Ing e CATUERJ-Tra. No quarto estágio da análise ocorreu a relação entre os segmentos escolhidos e as possíveis aulas das quais eles poderiam ser aplicados em forma de atividade para discussão em sala. No quinto e último passo, os achados do CATUERJ foram contrapostos com as pesquisas e possíveis aplicações dos CAT já existentes que foram expostos na seção 2 desta dissertação.

Agora, portanto, destacarei quais conceitos de cada aula da disciplina Introdução aos Estudos da Tradução I foram levados em consideração para a busca de segmentos do CATUERJ que servissem como exemplo prático de tais conceitos. Em seguida, explicarei qual recorte foi estabelecido para cada par de textos dos subcorpora CATUERJ-Ing e CATUERJ-Tra, elencando os pares de segmentos selecionados. Durante o procedimento de análise do CATUERJ, explicitarei os pressupostos teóricos relacionados aos exemplos do corpus somente de forma superficial. Isso porque o foco desta dissertação recai sobre as possibilidades de aplicação do CATUERJ. Portanto, trabalhar outros pressupostos teóricos de forma aprofundada não dialoga diretamente com o tema desta pesquisa.

Para relacionar os pares de segmentos do CATUERJ às Aulas 03, 04, 06, 07 e 08, procurei exemplos dos conceitos teóricos trabalhados durante os procedimentos de cada aula. A seguir, na Tabela 08, aponto os conceitos que selecionei para serem destacados a partir do conteúdo do CATUERJ.

Tabela 08 – Aulas selecionadas para a pesquisa e seus respectivos conceitos

Aula	Pressupostos Teóricos
Aula 03	Ambiguidade; polissemia e trechos com diferentes possibilidades de interpretação.
Aula 04	Escolhas linguísticas e discursivas; características de gênero; efeito polissêmico; e presença de jogos de palavras.
Aula 06	Compatibilidade de gênero e padrões retóricos.
Aula 07	Estrutura e escolhas gramaticais.
Aula 08	Colocação lexical; expressão idiomática; polissemia; sinonímia; efeito expressivo; efeito metafórico; prosódia semântica.

Fonte: Adaptada de WALSH, 2014, p. 217-220.

A Aula 03 ressalta a interpretação como condição para a tradução, então optei por concentrar minha seleção de segmentos do CATUERJ que ilustrassem ambiguidade, polissemia e/ou apresentassem mais de uma possibilidade de interpretação. A Aula 04, que trabalha a variação do contexto e o tipo de situação comunicativa no texto, pede que os alunos analisem escolhas linguísticas e discursivas, características de gênero, efeito polissêmico e jogo de palavras. Portanto, na análise do CATUERJ relacionada à Aula 04, decidi destacar segmentos que pudessem identificar esses conceitos. Já para relacionar a Aula 06 com o CATUERJ, optei por manter a seleção dos segmentos próxima de exemplos que ressaltassem a presença de padrões retóricos. Na Aula 07, transpondo os conceitos de estrutura e escolha gramatical para o contexto do CATUERJ, procurei por escolhas que refletissem exemplos de soluções tradutórias ou da falta dessas soluções. E para conectar a Aula 08, que trata de léxico, ao CATUERJ, precisei selecionar segmentos que exemplificassem alguns dos fenômenos lexicais trabalhados com os aprendizes em sala de aula.

Para conduzir a pesquisa dos conceitos no CATUERJ, foi necessário estabelecer um recorte condizente com esta dissertação. Portanto, optei por selecionar pouco mais de dez por cento de pares de segmentos do CATUERJ para esta análise, sendo o segmento do subcorpus CATUERJ-Ing, juntamente com seu equivalente do subcorpus CATUERJ-Tra, identificado pelo *software LFAaligner*. Logo, a partir dos dados de alinhamento do CATUERJ expostos na seção 2.4, montei a Tabela 09, que apresenta o número de segmentos escolhidos de cada par de textos alinhados, totalizando oitenta pares de segmentos selecionados de todo o CATUERJ.

Tabela 09 – Relação do recorte de segmentos de cada par de textos do CATUERJ

Par de textos alinhados	Número de segmentos alinhados	Número de segmentos escolhidos para análise
TO01-TT01	79	10
TO02-TT02	107	10
TO03-TT03	219	10
TO04-TT04	56	10
TO05-TT05	58	10
TO06-TT06	79	10
TO07-TT07	61	10
TO08-TT08	54	10

Fonte: A autora, 2016.

A fim de enumerar de forma organizada os pares de segmentos selecionados, criei dentro das seções 4.1 a 4.7 as Tabelas 10 a 16, sendo uma para cada par de textos do subcorpora. Na primeira coluna de cada tabela, enumero os segmentos selecionados respeitando a numeração proposta pelo *LFAaligner* e revisada por mim, cujo número final de segmentos de cada par de textos está na Tabela 09. Na segunda coluna, exponho o conteúdo do segmento referente ao número que aparece na primeira coluna. E na terceira coluna, disponho o segmento traduzido correspondente ao original que aparece na segunda coluna. De antemão, explico que a Tabela 13 é a única que contém quatro colunas, sendo a quarta coluna o segmento do texto traduzido denominado TT05 do subcorpus CATUERJ-Tra correspondente aos textos 04 e 05, do subcorpus CATUERJ-Ing, já que os textos TT04 e TT05 são traduções de um mesmo texto. Além disso, a quarta coluna da Tabela 13 também se diferencia por apresentar o número do segmento antecedendo o texto traduzido, que difere do segundo segmento observado na Tabela 05 no capítulo 2. Como já explicado no capítulo 2, os textos TO04 e TO05 receberam nomenclaturas diferentes, mas representam um só texto⁴⁹. Nas seções 4.1 a 4.7, a seguir, após a disposição de cada tabela, relaciono os pares de segmentos apresentados com os conteúdos das Aulas 03, 04, 06, 07 e 08 da ementa da disciplina Introdução aos Estudos da Tradução I. Outro aspecto a ser destacado em relação à

⁴⁹ Referências na Tabela 02 desta dissertação.

formatação é que alguns dos trechos comentados dentro dos segmentos foram sublinhados por mim, a fim de facilitar a visualização e identificação desses trechos durante a leitura.

4.1 TO01 – TT01

Tabela 10 – Análise dos segmentos do par TO01 – TT01 do CATUERJ

Nº do segmento	TO01	TT01
[01]	When Random House bought the venerable publisher Alfred A. Knopf in 1960, the story was reported on the <u>front page of The New York Times.</u>	Quando a Random House comprou a venerável editora Alfred A. Knopf em 1960, a notícia foi publicada na <u>primeira página do New York Times.</u>
[02]	<u>Its appearance</u> caused the general attorney general’s office to call Bennet Cerf, <u>the head of Random House.</u>	<u>Isso</u> levou a procuradoria-geral a telefonar para Bennet Cerf, <u>o diretor da Random House.</u>
[04]	Just the other week a similar story made the front page <u>of the Times</u> and other papers around the world.	Outro dia, uma história semelhante chegou às primeiras páginas <u>da Times</u> e de <u>outros periódicos</u> em todo o mundo.
[18]	<u>In 1999</u> , the top twenty publishers accounted for 93 percent of sales, and the ten largest had 75 percent of revenues.	<u>Em 1999</u> , as vinte maiores editoras eram responsáveis por 93% das vendas, e as dez maiores ficavam com 75% dos rendimentos.
[19]	Time Warner, which owns Little, Brown and Company as well as the Book-of-the-Month Club, <u>is the largest of the media conglomerates</u> , with close to <u>\$31 billion</u> in sales.	A Time Warner, proprietária da Little, Brown and Company, assim como do Book-of-the-Month Club, <u>é o maior dos conglomerados midiáticos do mundo</u> , com faturamento de quase <u>31 bilhões de dólares</u> em vendas.
[30]	<u>After all, some</u> seventy thousand books were published in the United States last year.	<u>Afinal, cerca de</u> 70 mil livros foram publicados ano passado nos Estados Unidos.
[45]	<u>I will begin</u> in the early 1940s, when <u>my father</u> , Jacques Schiffrin, helped to found a small exile publishing house in New York called Pantheon Books, which, in the twenty years of its independent <u>existence</u> , brought much European writing to the United States.	Eu começo no início da década de 1940, <u>quando meu pai</u> , Jacques Schiffrin, no exílio, ajudou a fundar em Nova York uma pequena editora chamada Pantheon Books, que, nos vinte anos de sua <u>vida independente</u> , trouxe muita literatura europeia para os Estados Unidos.
[61]	Would Garcia Marquez have become an	Garcia Márquez e Umberto Eco teriam

	international best-seller had he not had good translators? Would Umberto Eco?	se tornado bestsellers internacionais se não tivessem contado com bons tradutores?
[64]	One of the earliest pioneers who attempted to translate the Bible into English was John Wycliffe.	Um dos primeiros a tentar traduzir a Bíblia para o inglês foi John Wycliffe, que teve sorte o bastante para escapar da fogueira.
[65]	He was fortunate enough to avoid the stake, <u>but</u> 44 years after his death in 1384, his writings were condemned as heresy and his bones dug up and burned.	<u>Mas</u> , 44 anos após sua morte em 1834, seus escritos foram condenados por heresia e seus ossos desenterrados e queimados.

Fonte: Segmentos extraídos do CATUERJ, 2016.

Em relação aos segmentos elencados na Tabela 10, os segmentos 01 e 04 se complementam em termos de sentido, podendo ser ligados aos temas das Aulas 04 e 08, que tratam sobre escolhas linguísticas e de léxico. No segmento 01, o aprendiz traduziu o nome do jornal por “New York Times”, deixando de inserir o artigo “The” em sua tradução, que deveria fazer parte do nome por aparecer em letra maiúscula no texto original, mantendo-se, assim, “The New York Times”. À Aula 08, que trata de colocação e escolhas lexicais, também relacionei as escolhas lexicais do tradutor ao traduzir “head” como “diretor” no segmento 02; no segmento 30, no qual a palavra “some” é traduzida por “cerca de”; e no segmento 45, onde “existence” torna-se “vida” no texto em português.

Adiante, no segmento 04, o tradutor troca o artigo ao se referir ao jornal como “da Times”, podendo fazer com que o leitor interprete como a revista americana *Times magazine*. Além disso, o tradutor caracteriza o *The New York Times* como um periódico, termo que geralmente remete o leitor à área acadêmica. Essa passagem, portanto, também poderia ser relacionada à importância da interpretação na tradução, tema da Aula 03. Além disso, ela ilustra uma dificuldade lexical do aprendiz, que Espunya denomina uma “imprecisão lexical” (2013, p. 134, v. seção 1.2.8), pois o aluno não identificou como essa troca afetaria o contexto.

Os trechos sublinhados dos segmentos 02, 18, 30, 45 e 65 podem ser comentados em relação à Aula 06, que trata de compatibilidade de gênero e padrões retóricos, na qual os alunos discutem sobre contextualização, coesão, coerência (ALVES; MAGALHÃES; PAGANO, 2000, p. 71). Comparando os exemplos originais e traduzidos, é possível perceber que o tradutor aprendiz manteve as redes de coesão e coerência criadas pelo autor do texto original, além de ter mantido, também, as passagens que contextualizam o capítulo. O

segmento 19 também pode ser conectado com a Aula 06, que discute os conceitos de compatibilidade de gêneros e padrões retóricos, pois o tradutor acrescenta em sua tradução o tipo de moeda do faturamento de vendas do conglomerado midiático *Time Warner*, o dólar.

Por fim, os segmentos 61, 64 e 65 podem ser relacionados à Aula 07, na qual os alunos discutem estruturas e escolhas gramaticais. O tradutor aprendiz transformou as duas perguntas presentes no segmento 61 em uma só, na tradução, além de ter invertido a ordem de palavras a fim de aproximar a pergunta da forma mais natural em português, isto é, sem a inversão de verbo e sujeito que existe na língua inglesa. Ele também mesclou os segmentos 64 e 65, transferindo o início da passagem 65 para o fim da frase do segmento 64. Escolhas acertadas como essa poderiam gerar um *feedback* positivo da parte do professor em sala de aula, bem como proporcionaria a anotação do segmento com a etiqueta “*Feedback* positivo”, caso o professor pesquisador adotasse a nomenclatura da categoria criada para o PELCRA (v. seção 1.2.2).

4.2 TO02 – TT02

Tabela 11 – Análise dos segmentos do par TO02 – TT02 do CATUERJ

Nº do segmento	TO02	TT02
[08]	The answer has less to do with a <u>fancied reincarnation</u> than with the fact that literary taste changes.	A resposta não tem tanto a ver com uma <u>reencarnação idealizada</u> , mas com o fato de que o gosto literário muda.
[11]	The translator is no stand-in or ventriloquist for the foreign author, but a resourceful imitator who rewrites the original to appeal to another audience in a different language and culture, often in a different <u>period</u> .	O tradutor não é nenhum suplente ou ventríloquo do autor estrangeiro, mas um engenhoso imitador que reescreve o original a fim de atrair um novo público em outra língua e outra cultura, muitas vezes em um <u>período histórico</u> diferente.
[12]	This audience ultimately takes priority, insuring that the verbal <u>clothing</u> the translator cuts for the foreign work never fits exactly.	Este público enfim tem a prioridade, fazendo com que a <u>roupagem</u> verbal escolhida pelo tradutor para o trabalho importado nunca fique devidamente ajustada.

[15]	I would describe it as an attempt to compensate for an irreparable loss by controlling an exorbitant gain.	Eu descreveria isso como uma tentativa de compensar uma perda irreparável mantendo o controle sobre um ganho exorbitante.
[16]	The foreign language is <u>the first thing to go</u> , the very sound and order of the words, and along with them all the resonance and allusiveness that they carry for the native reader.	A língua estrangeira é <u>o ponto de partida</u> , precisamente os sons e ordens de palavras, e junto com eles toda a ressonância e referência que eles carregam para o leitor nativo.
[31]	<u>Hence</u> the first rule of reading translations: Don't just read for meaning, <u>but</u> for language too; <u>appreciate</u> the formal features of the translation.	<u>Portanto</u> a primeira regra para a leitura de traduções é: <u>não</u> leia <u>somente</u> para compreender o significado, <u>mas</u> preste atenção à língua <u>também</u> ; <u>tenha em mente</u> as características formais da tradução.
[38]	Consider a passage from <u>Margaret Jull Costa's version of The Man of Feeling</u> by the Spanish novelist Javier Marías.	Considere uma passagem <u>da versão feita por Margaret Jull Costa The Man of Feeling</u> , do romancista espanhol Javier Marías.
[80]	<u>Nonetheless</u> , the translator's interpretation remains partial, <u>both</u> incomplete in omitting irrecoverable aspects of the foreign text and slanted towards what is intelligible and interesting in the receiving culture.	<u>No entanto</u> , a interpretação do tradutor se mantém parcial, incompleta em omitir aspectos do texto estrangeiro que não podem ser cobertos e inclinada ao que é inteligível e interessante na cultura receptora.
[97]	The <u>figures</u> are staggering, even if we set aside the immediate worldwide translation of bestsellers like Stephen King, Danielle Steel, and Tom Clancy and focus on authors with literary reputations.	Os <u>números</u> são impressionantes, mesmo que desconsideremos a tradução mundial de bestsellers como Stephen King, Danielle Steel, e Tom Clancy e nos foquemos em autores com reputação literária.
[102]	<u>My aim has been to describe</u> ways of reading translations which increase rather than diminish the pleasures that only reading can offer.	<u>Meu objetivo é descrever</u> maneiras de ler traduções que aumentem ao invés de diminuir os prazeres que apenas a leitura pode oferecer.

Fonte: Segmentos extraídos do CATUERJ, 2016.

Os segmentos de número 31, 38, 80 e 102 da Tabela 11 podem servir de apoio à Aula 06. Eles servem como exemplos de padrões retóricos, contextualização e dos conceitos coesão e coerência. No segmento 80 há também a palavra “both” com a finalidade de atribuir coesão

ao texto. No entanto, na Tabela 11 é possível identificar que o tradutor optou por retirar o elemento coesivo na tradução. Nesse caso, os alunos poderiam discutir o processo pelo qual o tradutor aprendiz passou para traduzir o segmento, questionando se essa foi uma dificuldade gramatical ou lexical, por exemplo, ou mesmo se o tradutor aprendiz reconhece a palavra “both” como elemento coesivo, podendo o aprendiz não ter atingido as expectativas relacionadas ao conhecimento linguístico (ESPUNYA, 2013, p. 132, v. seção 1.2.8) de um tradutor do EscrTrad, que precisa ter cursado um mínimo de disciplinas obrigatórias antes de estar habilitado ao estágio. No segmento 11, o tradutor transforma “period” em “período histórico”, seguindo a mesma estratégia do tradutor do TT01 no segmento 19 (Tabela 10). Portanto, o segmento 11 também pode ser relacionado à Aula 06.

No segmento 12, há uma metáfora com o uso da palavra “clothing”, no TO, que foi traduzida como “roupagem”. O tradutor aprendiz, portanto, manteve a construção metafórica no mesmo campo semântico do original, apenas traduzindo literalmente a escolha vocabular do autor. Logo, essa passagem tem relação com a Aula 08, na qual são discutidos efeitos metafóricos.

Os segmentos 08, 16 e 97 podem remeter às Aulas 03 e 08 por ilustrarem escolhas lexicais diferenciadas do tradutor. No primeiro, é possível perceber o cuidado da escolha da palavra “idealizada” como tradução para “fancy” ao consultar um dicionário bilíngüe⁵⁰, no qual aparecem suas possíveis traduções, sendo “fantasia: a) imaginação. b) obra de imaginação. c) ideia, concepção, pensamento, parecer, opinião. d) noção, suposição, ilusão, imagem mental, ideia visionária. e) capricho, extravagância, veneta, desejo singular, gosto passageiro. f) gosto pessoal, vontade, preferência, arbítrio. g) ideia fixa, obsessão”. Ou seja, o tradutor se apropriou do sentido da frase ao traduzi-la para o português. O mesmo aconteceu com a palavra “figures”, que aparece no segmento 97. Na língua inglesa, “figure” é definida no dicionário como “um símbolo que representa um número; um valor expressado em números; ou uma pessoa ou animal que pode ser visto apenas em sua forma ou contorno”. O tradutor aprendiz, portanto, atentou para o contexto ao interpretar “figures” como “números”. Já na passagem “the first thing to go” do segmento 16, o sentido foi mantido, sendo a expressão traduzida como “o ponto de partida”, normalmente utilizada na língua portuguesa. Aqui, como na seção 4.1, há a motivação para que o professor pesquisador dê seu “*Feedback* positivo”, categoria proveniente da tipologia de erros do PELCRA (v. seção 1.2.2).

⁵⁰ Disponível em: http://michaelis.uol.com.br/escolar/ingles/definicao/ingles-portugues/fancy_12226.html (último acesso em 17/01/2016).

Da mesma forma que o tradutor aprendiz se ateuve ao conteúdo e ao sentido, também se preocupou com escolhas e estruturas gramaticais. O segmento 15 ilustra essa afirmação, pois o tradutor fez uma tradução literal mas inverteu a ordem dos adjetivos, preservando a ordem das palavras no português. Assim, as locuções “irreparable loss” e “exorbitant gain” foram traduzidas como “perda irreparável” e “ganho exorbitante”. Esse exemplo demonstra que o tradutor aprendiz do TT02 não tem o comportamento comum entre tradutores aprendizes, que traduzem de forma bem literal e acabam muitas vezes por preservar a ordem das palavras na língua do texto original. No entanto, o segmento 38 parece estar incompleto à interpretação, já que o equivalente da preposição “of” foi omitido na tradução, o que pode ser tópic de discussão tanto para a Aula 07 como para a Aula 03, que discutem os conceitos de estruturas e escolhas gramaticais e diferentes possibilidades de interpretação, respectivamente. Os exemplos acima, tanto do segmento 15, quanto do de número 38, permitem a discussão entre professor e alunos e criam a oportunidade de desconstruírem crenças (WALSH, 2014, p. 215), como a de que “a tradução é uma atividade prática que requer apenas um conhecimento da língua ou um bom dicionário” (ALVES; MAGALHÃES; PAGANO, 2000, p. 11)⁵¹. Dessa forma, ao passo que os alunos analisam os segmentos, o professor promove a oportunidade de o processo tradutório se tornar tema para discussão e desenvolve, assim, a consciência crítica nos aprendizes (v. capítulo 3).

4.3 TO03 – TT03

Tabela 12 – Análise dos segmentos do par TO03 – TT03 do CATUERJ

Nº do segmento	TO03	TT03
[37]	Recent translation theory <u>has stressed</u> the vital role of the translator in the interpretive process, and <u>has moved away</u> from the old idea of the translator as a betrayer of the pure source text.	A teoria mais atual da tradução <u>tem enfatizado</u> o papel vital do tradutor no processo interpretativo, e <u>se afastou</u> da antiga idéia do tradutor como um traidor do imaculado texto original.

⁵¹ Mesmo consciente de que a Aula 01, que trata de crenças sobre tradução, não faz parte do recorte desta pesquisa, faço menção aqui apenas para ilustrar a flexibilidade do trabalho e da aplicação do CATUERJ.

[58]	<u>Central</u> to the polysystems approach <u>were</u> certain key assumptions about translation, most crucial of which was the recognition of the role played by translation in shaping the literary polysystem.	<u>Fundamental</u> para a abordagem dos <u>polissistemas</u> eram certas suposições chave sobre a tradução. A mais crucial delas era o reconhecimento do papel desempenhado pela tradução na modelagem do polissistema literário.
[69]	<u>Firstly</u> , the translator must ‘understand perfectly the meaning and the subject matter of the author he translates’.	<u>Em primeiro lugar</u> , o tradutor deve ‘entender perfeitamente o significado e o assunto abordado pelo autor que ele traduz’.
[71]	<u>Secondly</u> , the translator must have perfect knowledge of the source language and have ‘achieved the same excellence in the language he wants to translate into’.	<u>Em segundo lugar</u> , o tradutor deve possuir total conhecimento da língua fonte e ter ‘alcançado a mesma excelência na língua para qual deseja traduzir’.
[73]	<u>Thirdly</u> , and significantly, the translator must ‘ <u>not enter into slavery</u> ’ by translating word for word.	<u>Em terceiro lugar</u> , e de maneira significante, o tradutor ‘ <u>não deve se escravizar</u> ’ traduzindo palavra por palavra.
[74]	<u>This metaphor</u> reinforces Dolet’s assertion that the role of the translator is an active one, and that the relationship between writer and translator is one of equality and not of subservience.	<u>Esta metáfora</u> reforça a asserção de Dolet de que o papel do tradutor é um papel ativo, e que a relação entre escritor e tradutor é de igualdade e não de subserviência.
[75]	<u>This leads him on to his fourth point</u> , which concerns the development of vernacular languages, a crucially important issue in Renaissance Humanist Europe.	<u>Isto o conduz a seu quarto ponto</u> , que diz respeito ao desenvolvimento das línguas vernáculas, um tema crucialmente importante na Europa Renascentista Humanista.
[78]	<u>Finally</u> , the translator should ‘observe the figures of speech’ and arrange words ‘with such sweetness that the soul is satisfied and the ears are <u>pleased</u> ’.	<u>Finalmente</u> , o tradutor deve ‘respeitar as figuras de linguagem’ e organizar as palavras ‘com tamanha doçura que a alma seja satisfeita e os ouvidos <u>agradados</u> ’.
[177]	The idea of translation as a lowlier kind of writing persisted for centuries, and <u>to some extent</u> is still with us today.	A idéia da tradução como um gênero inferior de escrita permaneceu por séculos, e <u>em alguma amplitude</u> se mantém conosco hoje.
[178]	Translators are <u>frequently</u> poorly	Os tradutores são <u>freqüentemente</u> mal

	paid, their work <u>often</u> regarded as inferior, despite the insistence by a huge number of eminent writers throughout the world on the complexity of translation and on its importance.	pagos, seu trabalho é <u>frequentemente</u> considerado inferior, a despeito da insistência de um grande número de escritores célebres em todo o mundo na complexidade da tradução e na sua importância.
--	---	--

Fonte: Segmentos extraídos do CATUERJ, 2016.

Em minha análise da Tabela 12, seis dentre os dez segmentos remetem às Aulas 04 e 06, que falam respectivamente sobre escolhas linguísticas e discursivas e compatibilidade de gêneros e padrões retóricos, sendo eles os de número 69, 71, 73, 74, 75 e 78. A primeira característica que chama atenção é a manutenção de todos os elementos coesivos que se estendem ao longo do capítulo. Outra característica é a clara tendência do tradutor aprendiz a evitar o advérbio terminado com o sufixo “-mente”. Segundo Vera e Danilo Nogueira⁵², tradutores profissionais renomados e experientes, os advérbios terminados em “-mente” soam mais pesados que seus equivalentes em inglês, geralmente terminados em “-ly”. Esse peso ocorre devido à terminação ter duas sílabas em português, enquanto em inglês só possui uma. Outra razão para se evitar advérbios terminados em “-mente” seria a potencial ocorrência de ecos ou rimas intratextuais desagradáveis.

Outra observação feita durante a análise é a escolha lexical e estrutura gramatical dos segmentos 37, 58, 78, 177 e 178. No segmento 37 há uma inconsistência na manutenção dos tempos verbais, aparecendo, na tradução, as locuções “tem enfatizado”, em uma tradução literal, e “se afastou”, remetendo ao sentido do texto original. Já no segmento 58 há um erro de concordância, pois a frase traduzida deveria começar com o adjetivo no plural: “fundamentais”; para concordar com o substantivo “suposições”, que aparece logo adiante, na mesma frase. Traçando um paralelo com a tipologia de erros do MeLLANGE, por exemplo, esse equívoco estaria inserido na categoria Língua – Flexão – Número (v. Anexo I).

A última palavra do segmento 78 aparece no texto original como “pleased”, com sua tradução equivalente dentro da locução “ouvidos agradados”. Ao consultar um dicionário bilíngue⁵³, constatei que, segundo ele, as possíveis definições para a palavra *pleased* são

⁵² Apresentação de Vera e Danilo Nogueira no I Congresso Internacional da ABRATES (Rio de Janeiro, 2005). Disponível em: http://www.iscap.ipp.pt/elgg/anaf/files/-1/416/Tecnicas_de_traducao%5B1%5D.pdf (último acesso em 17/01/2016).

⁵³ Disponível em: <http://dictionary.cambridge.org/pt/dicionario/ingles-portugues/pleased> (último acesso em 17/01/2016).

“satisfeito” e “contente”. Como é possível ver na Tabela 12, a palavra “satisfeita” foi usada logo antes na frase traduzida. De qualquer forma, o tradutor poderia ter optado pela palavra “contente”, para evitar o estranhamento da locução “ouvidos agradados”.

Ao final do segmento 177, o aprendiz traduz *to some extent* por “em alguma amplitude”, fazendo uma opção por um equivalente não tão frequente como a expressão “até certo ponto”, por exemplo. Por último, no segmento 178 o advérbio “frequentemente” aparece por duas vezes na mesma frase. Ou seja, apesar da tendência de evitar os advérbios terminados em “-mente” ao longo de todo o texto, o aprendiz repete este último advérbio de modo, o que poderia ter sido evitado caso ele optasse pela expressão “muitas vezes”.

Há, também, a presença de uma metáfora no segmento 73, que aparece entre aspas tanto no texto original como na tradução, por se tratar de uma citação. Ela é explicitada e explicada no segmento seguinte, de número 74. Portanto, ambos os segmentos estão relacionados à Aula 08, que fala sobre efeito metafórico. Outro detalhe que não prejudica a comunicação do texto, mas deve ser tratado com atenção pelo tradutor, é a utilização das regras do novo acordo ortográfico, visto que o tradutor aprendiz manteve o acento agudo em “ideia”, no segmento 37, e utilizou por duas vezes a palavra “frequentemente” com o trema no segmento 178. Ressalto que, nos exemplos acima, foi necessário apontar os erros do aprendiz por se tratarem de erros relacionados à forma e à estrutura da língua, neste caso, a acentuação.

4.4 TO04/TO05 – TT04 e TT05

Tabela 13 – Análise dos segmentos do trio TO04/TO05 – TT04 – TT05 do CATUERJ

Nº do segmento	TO04/TO05	TT04	TT05
[07]	Here are a couple of word-for-word items, one from an Indonesian hotel brochure and the other from a pamphlet produced by the city of Salamanca tourist office:	Abaixo estão dois exemplos traduzidos palavra por palavra. Um é de uma brochura de um hotel indonésio e o outro de um panfleto produzido pela secretaria de turismo da cidade de Salamanca.	Encontramos a seguir alguns exemplos do uso da palavra-por-palavra, um retirado da brochura de um hotel na Indonésia, e outro de um panfleto produzido pela secretaria de turismo de Salamanca.
[10]	From both of <u>these</u> paragraphs we get an	Em ambos os parágrafos <u>acima</u> , temos a ideia do	Em ambos os parágrafos, é possível termos uma

	idea of what is being described, although the obscurity of expression is due to the translator having adhered too closely to the original even, in the case of Spanish, down to translating inverted commas.	que é descrito, mesmo que a obscuridade de expressão se deva ao fato de o tradutor se prender demais ao original, chegando até, no caso do espanhol, a traduzir o texto entre aspas.	ideia do que foi descrito, porém através de uma tradução confusa, feita por um tradutor que se manteve próximo demais do original, deixando até as aspas do texto que estava em espanhol.
[13]	With so many examples of the <u>inadequacies</u> of literal translation everywhere we go, it is <u>hardly surprising that many translators are wary of it.</u>	Com tantos exemplos da <u>inadequação</u> da tradução literal em toda parte, <u>não é nenhuma surpresa que muitos tradutores sejam cautelosos quanto a ela.</u>	Com tantos exemplos dos <u>problemas</u> causados pela tradução literal, <u>é perfeitamente compreensível a cautela de muitos tradutores quanto ao seu uso.</u>
[21]	Nevertheless, the <u>inadequacy</u> of machine translation in dealing with certain kinds of text opened up debates about forms of translation that still continue.	Contudo, a <u>inadequação</u> da tradução mecânica em lidar com certos tipos de texto abriu debates, que ainda continuam, sobre as formas de tradução.	Entretanto, a <u>ineficiência</u> de uma máquina ao lidar com a tradução de certos tipos de textos gerou debates que ainda existem a respeito dos métodos de tradução.
[22]	So with all this in mind, why would anyone want to defend literal translation?	Logo, com tudo isso em mente, porque iria alguém querer defender a tradução literal?	Tendo tudo isso em mente, por que alguém iria defender a tradução literal?
[23]	<u>Can it ever be useful?</u>	<u>Pode ela ser sequer útil?</u>	<u>Ela pode vir a ser útil?</u>
[24]	Well, yes, it can.	Bem, pode sim.	Sim, ela pode.
[31]	All I will say is that after some <u>discussion</u> , the difference between the English and German sentence to emerge, and through the errors of literal translation he was able to see an alternative.	Vou dizer apenas que, depois de alguma <u>discussão</u> , a diferença entre as sentenças em inglês e alemão começou a emergir e, através dos erros da tradução literal, ele conseguiu ver uma alternativa.	Tudo o que direi é que, após alguma <u>discussão</u> , a diferença entre as sentenças em inglês e alemão começou a aparecer, e através dos erros da tradução literal, meu filho conseguiu ver uma alternativa.
[42]	<u>Apart from</u> restricted supplies of manuscripts, scribes to copy them and teachers to disseminate knowledge, not all rules promoted learning, and disease, wars and Viking raiders made consistent study difficult.	<u>À parte de</u> uma oferta restrita de manuscritos, escribas para copiá-los e professores para disseminar o conhecimento, nem todos os governantes promoviam o ensino. Além disso, doenças, guerras e piratas	<u>Com exceção dos poucos</u> manuscritos, dos escrivães para copiá-los e dos professores que disseminavam o conhecimento, nem todos os regentes promoviam o aprendizado, e as doenças, guerras e invasões vikings

		viquingues tornavam difícil o estudo consistente.	tornavam o estudo constante difícil.
[46]	Through literal translation, understanding spread and gradually Anglo- Saxon began to acquire status in its own right as a written language.	Através da tradução literal, o conhecimento se disseminou e, gradualmente, o anglo- saxão começou a adquirir o próprio direito ao status de língua escrita.	Através da tradução literal, o conhecimento se espalhou e aos poucos o anglo-saxão foi adquirindo por direito próprio a condição de língua escrita.

Fonte: Segmentos extraídos do CATUERJ, 2016.

A Tabela 13, como mencionado na introdução deste capítulo, diferencia-se das outras tabelas por conter quatro colunas e duas traduções do mesmo texto. Em minha análise, percebi que todos os segmentos da Tabela 13 podem ser enumerados em termos de escolhas linguísticas e discursivas, tema da Aula 04, e escolhas e estruturas gramaticais, tema da Aula 07. Isso porque há duas traduções feitas por diferentes tradutores aprendizes, tornando-se inevitável a comparação entre as duas traduções.

É possível perceber que o tradutor aprendiz do TT04 se atém mais às escolhas lexicais e estruturas gramaticais preferidas pelo autor do texto original que o tradutor aprendiz do TT05, principalmente nos segmentos 13, 21, 23 e 42. Destaco que a opção por traduzir de forma mais literal feita pelo tradutor aprendiz do TT04 funciona, de certa forma, como uma ironia talvez inconsciente. Isso porque o autor do TO04/TO05 critica justamente as técnicas de tradução literal e tradução palavra por palavra e suas irregularidades, como ilustram os segmentos 10, 13, 22 e 23. No segmento 46, no entanto, o autor do TO04/TO05 cita benefícios trazidos para a área da tradução como consequência da tradução literal. Esse momento pode servir como uma reflexão posta aos alunos pelo professor, para que eles possam analisar o processo em conjunto e desenvolver a criticidade ao discutir as técnicas de tradução utilizadas nesse e em outros casos. Algumas questões poderiam ser levantadas, dentre elas: “O estágio no EscrTrad não tem como uma de suas funções desenvolver um tradutor consciente de quais estratégias está utilizando?”; ou “As escolhas feitas por esses tradutores aprendizes são conscientes?”; ou mesmo “Podemos detectar dificuldades no processo de aprendizagem desse tradutor através dos exemplos do CATUERJ? Quais são elas?” (ESPUNYA, 2013, v. seção 1.2.8). Nas traduções da Tabela 13, apesar de haver diferentes escolhas, os dois tradutores aprendizes traduziram o texto sem perdas de significado que pudessem impedir a leitura dos textos em português.

Finalizando, comento o segmento 31, no qual aparece um exemplo malsucedido de manutenção da prosódia semântica da palavra “discussion” em inglês, que foi traduzida como “discussão” por ambos os tradutores aprendizes. Esse erro não necessariamente impede a comunicação do texto, mas uma opção mais condizente seria a palavra “conversa”, por exemplo. Esse segmento pode ser relacionado com a Aula 08, que tem a prosódia semântica como um dos temas.

4.5 TO06 – TT06

Tabela 14 – Análise dos segmentos do par TO06 – TT06 do CATUERJ

Nº do segmento	TO06	TT06
[02]	<u>In addition to</u> being a member of our country, we are members of the world community, and this gives us a global identity.	<u>Além de</u> cidadãos do nosso país, acrescente-se o fato de que somos membros da comunidade internacional, o que nos proporciona uma identidade global.
[12]	<u>As mentioned earlier</u> , the whole world is undergoing complex changes in different areas such as technology and education.	<u>Como foi mencionado anteriormente</u> , a humanidade como um todo está passando por mudanças complexas em diferentes áreas, como por exemplo, a tecnológica e a [de educação ou educacional?].
[25]	<u>Therefore</u> , receptive skills should be developed before the productive ones; i.e. <u>by reinforcing their passive knowledge</u> , students will eventually improve their active knowledge.	<u>Sendo assim</u> , as habilidades receptivas devem ser desenvolvidas antes das habilidades produtivas; i.e. <u>ao fortalecerem [fortalecemos conhecimentos?]</u> seus <u>conhecimentos passivos</u> , os estudantes irão finalmente melhorar seus conhecimentos ativos.
[31]	The articles will <u>not only</u> improve the students' reading skill in general, <u>but also</u> give them insights which will subconsciously be applied when actually translating.	Os artigos <u>não apenas</u> irão desenvolver a habilidade dos estudantes para a leitura de uma forma geral, <u>mas também</u> darão a eles a percepção adequada que será aplicada por eles de forma inconsciente na hora de traduzir.
[36]	<u>Moreover</u> , translation trainees should have a <u>good ear</u> for both source and target languages; i.e.	<u>Além disso</u> , os estagiários de tradução devem ter um <u>“bom ouvido”</u> tanto na língua fonte quanto na língua alvo; i.e.

	they should be alert to pick up various expressions, idioms, and specific vocabulary and their uses, and <u>store them in their minds</u> to be used later.	devem estar atentos para capturar várias frases, expressões idiomáticas, o vocabulário específico e seu uso, e <u>armazena-los em suas mentes</u> para que sejam utilizados posteriormente.
[41]	<u>One of the most important points</u> to consider in the act of translating is understanding the value of the source text within the framework of the source-language discourse.	<u>Um dos pontos mais importantes</u> a serem considerados no ato da tradução é a compreensão do valor do texto fonte dentro da estrutura do discurso da língua fonte.
[43]	<u>Therefore</u> , the hidden structure of the source text should be discovered through the use of various discursal strategies by the translator.	<u>Por isso</u> , a estrutura oculta do texto fonte deve ser descoberta pelo tradutor através do uso das várias estratégias discursivas.
[63]	<u>Another important point</u> is that successful translators usually choose one specific kind of texts for translating and continue to work only in that area; <u>for example</u> a translator might translate only literary works, scientific books, or journalistic texts.	<u>Um outro ponto importante</u> a ressaltar é que tradutores consagrados normalmente escolhem um modelo específico de texto para traduzir e tendem a permanecer trabalhando somente nesta área. <u>Por exemplo</u> , um tradutor pode especializar-se em traduzir apenas trabalhos literários, livros científicos ou textos jornalísticos.
[68]	<u>In an academic environment</u> , recently published articles, journals and books on translation are available to the trainees, <u>who</u> thus become familiar with good translators and their work by reading them and then comparing them with the original texts.	<u>Considerando o ambiente acadêmico</u> , nele estarão disponíveis aos alunos artigos de publicação recente, periódicos e livros sobre tradução. Isto torna possível a <u>estes</u> conhecerem bons tradutores e seus trabalhos, possibilitando a leitura destes trabalhos e a comparação com os textos originais.
[75]	<u>Finally</u> , it is important to know that it takes much more than a dictionary to be a good translator, and translators are not made overnight.	<u>Finalmente</u> , é importante saber que é necessário muito mais do que um dicionário para que alguém venha a tornar-se um bom tradutor. E mais: tradutores não são formados da noite para o dia.

Fonte: Segmentos extraídos do CATUERJ, 2016.

A Tabela 14, que lista os segmentos selecionados dos textos TO06 e TT06 do CATUERJ, mostra uma característica particular desse tradutor aprendiz em relação às outras oito traduções. Nos segmentos 12 e 25, ele adiciona ao longo do texto observações entre colchetes nas quais questiona suas próprias escolhas linguísticas. É importante que esses questionamentos sejam levados aos alunos da disciplina, principalmente remetendo às Aulas 03, 04 e 07, que discutem possibilidades de interpretação, escolhas linguísticas e discursivas e estruturas e escolhas gramaticais.

De modo geral, os alunos podem discutir sobre como as opções trazidas pelo tradutor aprendiz no TT06 poderiam afetar a interpretação ou a estrutura gramatical e lexical do texto, dependendo da escolha final. Além, é claro, de os próprios alunos terem a possibilidade de levantar outras opções e discutirem uma solução consensual, o que poderia ser feito em relação a qualquer um dos textos do CATUERJ. Além disso, todos os segmentos selecionados para comporem a Tabela 14 podem ser relacionados à Aula 06, cujo tema é a compatibilidade de gêneros e padrões retóricos. Dessa forma, todos os segmentos da Tabela 14 podem ser relacionados às Aulas 03, 04, 06 e 07, demonstrando que os conteúdos e temas das aulas são interdependentes e funcionam em um conjunto, cujo produto final é a tradução. Essa relação também poderia levar os alunos a concluírem que o CAT ou o CATUERJ, nesse caso mais específico, favorece a reflexão sobre o processo da tradução e as etapas pelas quais o tradutor aprendiz passa desde o início do processo tradutório até o produto final. Mais uma vez, destaco a oportunidade de aplicação do CATUERJ na sala de aula em prol do desenvolvimento da visada crítica, objetivo do curso de Introdução aos Estudos da Tradução I (v. capítulo 3), e da possibilidade de olhar de perto para o processo tradutório, um dos objetivos do livro *Traduzir com autonomia* (ALVES, MAGALHÃES & PAGANO, 2000, v. capítulo 3).

Na análise da Tabela 14, é indispensável comentar o único erro que poderia comprometer a interpretação do texto, dentre os segmentos selecionados. No segmento 68, o tradutor transforma uma frase do TO06 em duas, e acaba perdendo sentido ao usar o pronome “estes” para se referir a “alunos”, palavra que aparece anaforicamente. Esse erro pode ser inserido na tipologia de erros do MeLLANGE na categoria “Língua”. No entanto, não há uma subcategoria específica preestabelecida na qual ele pode ser encaixado, levando-o a ser um erro “Definido pelo usuário”.

Neste ponto da pesquisa, atento para o fato de que, dentre as tipologias de erros vistas e comentadas nesta dissertação (v. capítulo 1), pouco poderia ser aproveitado para esta análise. Por outro lado, o destaque dado a boas escolhas tradutórias, incluindo a alta

frequência de elementos coesivos que realçam padrões retóricos, juntamente com a menção a dificuldades lexicais dos aprendizes, demonstra que o CATUERJ pode ser aplicado na sala de aula para discutir o processo tradutório tanto no caso de dificuldades lexicais e gramaticais (v. seção 1.2.8) e, principalmente, da visão crítica do tradutor que o leva a fazer boas escolhas (v. capítulo 3).

4.6 TO07 – TT07

Tabela 15 – Análise dos segmentos do par TO07 – TT07 do CATUERJ

Nº do segmento	TO07	TT07
[09]	Debates <u>about when</u> a translation stops being a translation and becomes an adaptation have rumbled on for decades, <u>but I have yet to meet anyone</u> who can give me an adequate definition of the difference between the two.	Debates <u>acerca de quando</u> o tradutor deixa de traduzir e passa a adaptar têm se alastrado por décadas a fio, <u>mas eu ainda nunca encontrei alguém</u> que pudesse me dar uma definição adequada sobre a diferença entre as duas coisas.
[11]	The problem is, <u>though</u> , how close do you have to be, and how far away do you have to move before the labels change?	Temos <u>então</u> um problema: como saber qual o nível de proximidade e distância em que devemos estar para delimitar o momento em que mudamos o rótulo da embalagem?
[12]	I have always had problems with <u>this</u> distinction.	Eu sempre tive problemas com <u>essa</u> distinção.
[19]	<u>For surely</u> what Balmer is doing is exercising her own creativity while engaging with the creative work of another writer.	<u>Sem dúvidas</u> o que Balmer fez foi exercer sua própria criatividade enquanto se envolvia com a obra criativa de outro escritor.
[33]	<u>When</u> you are translating a legal document, <u>for example</u> , nobody is going to complain that you have produced an adaptation if the two texts are clearly different; the client will acknowledge that you have rewritten the document in accordance with the style and the conventions of the culture	<u>Quando</u> se traduz um documento legal, <u>por exemplo</u> , ninguém vai reclamar que você produziu uma adaptação de dois textos que são evidentemente diferentes; o cliente vai reconhecer que você reescreveu o documento de acordo com o estilo e convenções da cultura alvo na qual os leitores estão

	of the readers for whom it is intended.	inseridos.
[35]	<u>Yet</u> , nobody complains about a letter being adapted, rather it is accepted that this is a necessary adjustment.	<u>Ainda assim</u> , ninguém reclama da adaptação de uma carta, que é amplamente aceita como um ajuste necessário.
[36]	<u>So</u> why are we so obsessive when it comes to literary works, what is it that fuels the debate and leads intelligent people to hunker down in trenches, refusing even to contemplate that there can be degrees of freedom for all translators?	<u>Então</u> por que essa obsessão toda quando se trata de obras literárias? Quais fatores alimentam esse debate e levam pessoas inteligentes a se agacharem atrás de trincheiras na teimosia de não acreditarem que possa haver graus diferentes de liberdade para todos os tradutores?
[52]	<u>In the 1960s</u> , Ted Hughes and Daniel Weissbort started a <u>journal</u> , Modern Poetry in Translation.	<u>Nos anos 60</u> , Ted Hughes e Daniel Weissbort iniciaram um <u>caderno</u> , Modern Poetry in Translation.
[53]	<u>This</u> journal enabled a great host of poets from around the world to reach an English readership.	<u>Essa</u> publicação permitiu que um grande número de poetas ao redor do mundo alcançasse os leitores anglófonos e, <u>ainda hoje</u> , ela garante recursos valiosos e leituras prazerosas.
[54]	It is <u>still</u> a valuable resource and a pleasure to read, <u>and</u> what I particularly enjoy are the brief comments by translators and their work.	Do que eu pessoalmente mais gosto são os comentários dos tradutores sobre seu trabalho.

Fonte: Segmentos extraídos do CATUERJ, 2016.

Na Tabela 15, selecionei dez pares de segmentos dos textos TO07 e TT07. Todos esses segmentos puderam ser relacionados com a Aula 06, na qual são discutidos características de gênero e padrões retóricos. Além disso, relacionei o segmento 52 às Aulas 03 e 04, que tratam de interpretação e escolhas linguísticas e efeito polissêmico, respectivamente. Nesse segmento aparece a palavra “journal”, podendo gerar diferentes interpretações de acordo com o contexto, dentre eles “diário”; “jornal”; ou uma revista que trata de assuntos que interessam a um grupo específico de pessoas. Logo, o tradutor aprendiz optou pela palavra “caderno” em português, demonstrando que adotou a terceira definição da palavra “journal”.

Os segmentos 53 e 54 também poderiam ser discutidos na Aula 04, que trata de escolhas linguísticas e discursivas e características de gênero, pois há uma mescla de informações na tradução em relação ao texto original, porém sem perda de sentido ou

informação. O tradutor aprendiz transfere a primeira parte do segmento 54 para o final do segmento 53, adicionando o advérbio “hoje” à tradução, podendo agregar elementos à coesão e contextualização do texto.

As Aulas 07 e 08, que discutem estrutura e escolhas gramaticais e colocação lexical e efeito metafórico, respectivamente, também podem ser debatidas a partir da Tabela 15, por meio dos segmentos 09 e 36, respectivamente. No segmento 09, o trecho “I have yet to meet anyone” é traduzido por “eu ainda nunca encontrei ninguém”. O tradutor aprendiz usa as palavras “ainda”, “nunca” e “ninguém” na mesma frase, o que não só acarreta problemas na estrutura gramatical da oração, como também no entendimento do leitor, pois tais palavras comprometem a clareza da frase. No segmento 36, o autor do texto cria uma relação metafórica entre a aceitação da adaptação em traduções e o contexto de guerra e liberdade. Tal relação é respeitada e mantida pelo tradutor aprendiz.

4.7 TO08 – TT08

Tabela 16 – Análise dos segmentos do par TO08 – TT08 do CATUERJ

Nº do segmento	TO08	TT08
[02]	At a dinner table <u>the other week</u> the <u>conversation</u> moved on to a <u>discussion</u> about translation.	<u>Há algumas semanas</u> , enquanto jantávamos, a <u>conversa</u> à mesa rumou para uma <u>discussão</u> sobre tradução.
[03]	One of the people round the table was very <u>dismissive</u> – <u>never yet read</u> a good translation, you have to recognize that something is always lost in translation, better by far to read the original.	Uma das pessoas ao redor da mesa tinha uma opinião muito <u>incrédula</u> - <u>ainda nunca li na vida</u> uma tradução decente, é inegável que sempre se perde alguma coisa quando se traduz, então é muito melhor ler o original.
[08]	A good <u>translation</u> takes the reader into the world of the (translated) book in the same way that a writer took his or her readers into the original book in the first instance.	Um bom <u>tradutor</u> faz o leitor embarcar no mundo do livro (traduzido) do mesmo jeito que o autor fez com que seus leitores ficassem imersos no livro original.
[09]	There is loss in translation, <u>but</u> there is also gain, and this elementary fact <u>seems</u> so often <u>to be forgotten</u> .	Existe perda na tradução sim, <u>mas</u> também existe um ganho, e <u>muitos parecem se esquecer</u> desse fato

		fundamental.
[21]	<u>However</u> , <u>this</u> was an editorial decision taken in clear knowledge of <u>where</u> the poem was destined to be published, which was in an anthology of women's poetry on the environment throughout the ages.	<u>Essa</u> , <u>porém</u> , foi uma decisão editorial tomada com pleno conhecimento <u>acerca de</u> onde o poema estava destinado a ser publicado, no caso, uma coletânea de poesias femininas sobre o meio ambiente ao longo dos séculos.
[25]	There is <u>also</u> another kind of translation reading, however, and that involves translating in order to get to know a writer more completely.	Há <u>também</u> outro tipo de leitura para tradução, uma que requer a tradução como forma de conhecer um escritor mais a fundo.
[46]	<u>This</u> is fine, <u>but</u> it separates the soul from its verb, <u>so that</u> the sense of the all-powerful soul is diminished <u>because</u> the phrase <u>el alma vence</u> has been split.	<u>Isso</u> é válido, <u>mas</u> separa a alma de seu verbo, <u>então</u> a noção de alma todo-poderosa acaba reduzida <u>devido à</u> divisão do sintagma <u>El alma vence</u> .
[51]	Machado is a poet who is <u>both</u> profoundly spiritual <u>and</u> in possession of an astonishing pictorial ability that enables him to create images of the Spanish landscape that have made him one of the best-loved 20th century poets in his homeland.	Machado é um poeta de <u>tamanha</u> profundidade espiritual e surpreendente habilidade pitoresca, capaz de criar imagens da paisagem hispânica que fez dele um dos mais bem-amados poetas do século XX em sua terra natal.
[53]	I am gradually <u>getting to know Machado better</u> , and perhaps eventually I might venture to try and <u>polish</u> some of the <u>scribbles</u> into readable versions.	Eu estou, aos poucos, <u>conhecendo melhor Machado</u> , e talvez, algum dia eu possa embarcar na empreitada de <u>envernizar</u> esses <u>rabiscos</u> até transformá-los em versões legíveis.
[54]	<u>For now, though</u> , I am appreciative of the work of other translators who have been bold enough to tackle this marvelous, difficult, more readable of poets.	<u>Mas por ora</u> , eu me contento em apreciar o trabalho de outros tradutores que foram corajosos a ponto de lidar com esse poeta, que é um dos mais admiráveis, difíceis e prazerosos de que se tem notícia.

Fonte: Segmentos extraídos do CATUERJ, 2016.

Os segmentos 02, 09, 21, 25, 46 e 54, listados na Tabela 16, puderam ser relacionados à Aula 06. Eles contêm exemplos de ferramentas discursivas de coesão e coerência, como pronomes e marcadores, que facilitam a compreensão do texto (ALVES; MAGALHÃES; PAGANO, 2000, p. 78).

À Aula 04, que fala sobre escolhas linguísticas e discursivas, pude relacionar o segmento 03, no qual o tradutor aprendiz traduziu “dismissive” como “opinião incrédula”. Nos segmentos 08 e 09, as escolhas discursivas e gramaticais do tradutor aprendiz diferem das do autor. No primeiro, o autor usa o substantivo “translation”, enquanto o tradutor opta por usar o agente da ação, o “tradutor”. Logo em seguida, no segmento 09, o tradutor opta por trocar a voz passiva usada no original pela voz ativa, transformando “to be forgotten” em “muitos parecem se esquecer”. Esses dois exemplos podem ser conectados não só à Aula 04, mas também às Aulas 03 e 07, por envolverem mudanças de estruturas gramaticais, que podem impactar a interpretação do texto. Nesse caso, não há perda de sentido, apesar da mudança estrutural.

Uma observação interessante no segmento 03 é a tradução do trecho “never yet read” por “ainda nunca li na vida”, pois remete à mesma escolha de palavras feita pelo tradutor aprendiz no segmento 09 da Tabela 15. Há, também, outro exemplo de prosódia semântica no segmento 02, com a palavra “discussão”. Ao contrário do segmento 31 da Tabela 13, dessa vez a palavra “discussão” foi utilizada na tradução para estabelecer a continuidade da cena, sendo uma sequência possível da “conversa”.

Por último, destaco o segmento 53, onde há um exemplo de metáfora que remete às Aulas 04 e 08, por também envolver escolhas linguísticas. O autor do texto utiliza as palavras “polish” e “scribbles”, as quais são traduzidas como “envernizar” e “rabiscos”. No entanto, a palavra “scribbles”, no contexto da escrita, também poderia ser traduzida como “garranchos”, já que a continuação da frase diz “transformá-los em versões legíveis”.

4.8 Comentários finais

Optei por trabalhar com o CATUERJ identificando boas escolhas tradutórias feitas pelos aprendizes, assim como dificuldades, para aplicá-las na sala de aula. Nesse trabalho, portanto, conjugo a utilização do CATUERJ para apontar pontos fracos e fortes desse grupo de alunos e trabalhá-los com outro grupo de alunos, a fim de fomentar a visada crítica e manter os objetivos da disciplina na qual o CATUERJ seria utilizado. Outra utilização proporcionada pelo CATUERJ é a aplicação direta das informações retiradas do próprio CAT na sala de aula da disciplina Introdução aos Estudos da Tradução I. Essas duas possibilidades de utilização do CATUERJ corroboram com as utilizações possíveis de corpora de aprendizes

e de aprendizes de tradução definidas por Bennison e Bowker; Nesselhauf; Zanettin (v. seções 1.1 e 1.2), bem como com os objetivos definidos por Spence, responsável pelo CAT desenvolvido na Universidade de Leipzig, que pretendia usar seu CAT para ensinar tradução ou língua estrangeira com o auxílio do corpus eletrônico de aprendizes (v. seção 1.2.1).

Após a relação dos segmentos escolhidos para o recorte desta pesquisa com as aulas planejadas por Walsh (2014), é possível perceber que há uma relação explicitamente forte entre o conteúdo dos pares de textos com os temas trabalhados na disciplina Introdução aos Estudos da Tradução I, bem como com algumas publicações sobre outros CAT já produzidos. Além disso, percebi que pode ser estabelecida uma ponte entre os segmentos aqui comentados e a pesquisa de Espunya (2013) a respeito de dificuldades lexicais apresentadas na seção 1.2.8 desta dissertação. A discussão de tais dificuldades lexicais em sala de aula poderia servir como um pontapé inicial para introduzir os temas trabalhados em sala de aula, além de o termo “dificuldades lexicais” não ser considerado o mesmo que o termo “erro”, pois aquele carrega uma conotação mais positiva, em minha opinião. Isso estimula o aprendiz a questionar o processo tradutório, para que a dificuldade existente possa ser superada em traduções futuras, estando o tradutor aprendiz consciente desse processo.

A partir da aplicação deste recorte do CATUERJ em sala de aula, os aprendizes poderiam se beneficiar não só do contato com produções reais de outros aprendizes, mas também do diálogo fundamentado nos temas referentes a cada aula para analisarem os textos do CATUERJ. Isso mostra que, como afirmou Espunya (2013, p. 136), o CAT pode revelar onde mora a dificuldade no processo de aprendizagem do tradutor aprendiz que, na discussão em sala de aula mediada pelo professor, pode levar os alunos a refletirem sobre as escolhas dos tradutores aprendizes e desenvolverem sua visão crítica.

O objetivo desta análise foi ilustrar as possíveis aplicações do CATUERJ ao relacioná-lo com as aulas elaboradas por Walsh (2014) para a disciplina Introdução aos Estudos da Tradução I e com as pesquisas com corpora de aprendizes e de aprendizes de tradução já realizadas e expostas no capítulo 1 desta dissertação. O CATUERJ, portanto, possibilita o acesso ao processo de tradução e realça etapas desse processo como uma possibilidade de aperfeiçoamento, e não dificuldades como aspectos negativos do tradutor aprendiz, visando a contribuir para uma construção positiva da imagem desse tradutor aprendiz e, ao mesmo tempo, formando tradutores mais críticos e conscientes de suas escolhas.

CONSIDERAÇÕES FINAIS

Esta dissertação investigou a criação de possibilidades de aplicação do CATUERJ em uma sala de aula da disciplina eletiva Introdução aos Estudos da Tradução I (WALSH, 2014), a partir da análise de um recorte com oitenta segmentos desse CAT contrapostos com teorias e pesquisas já desenvolvidas ou ainda em desenvolvimento na área de Estudos da Tradução com base em corpora.

Selecionei sete textos em inglês, cujo tema era tradução, e oito traduções dos respectivos textos originais para o português feitas por tradutores aprendizes durante o período em que estagiaram no EscrTrad. Esses textos foram alinhados com o auxílio do *LFAAligner* e os segmentos foram revisados manualmente, no programa Microsoft Excel. Um total de oitenta pares de segmentos foram selecionados como recorte e contrapostos às teorias e pesquisas já publicadas sobre CAT, assim como relacionados aos temas de cinco aulas da disciplina eletiva Introdução aos Estudos da Tradução I. O software *LFAAligner* se mostrou eficiente e cumpriu com a proposta do alinhamento em nível de frase. Após sua utilização, foi necessária a revisão manual devido às condensações e divisões de frases feitas pelos tradutores aprendizes. A revisão manual foi realizada apenas para que os textos originais e traduzidos pudessem ter o mesmo número final de segmentos, com seu alinhamento de acordo com o conteúdo de cada segmento. Dessa forma, a disposição visual da análise e referência aos segmentos foi facilitada pelo uso do *software*. O alinhamento manual do corpus teria prolongado demais o processo. Além do alinhamento, o CATUERJ também foi documentado com informações a respeito dos textos que o integram. No entanto, a documentação desta pesquisa não focou em informações sobre os estagiários do EscrTrad que realizaram as traduções selecionadas para compor o CATUERJ. Então, não há informações sobre gênero, idade e experiência com tradução dos tradutores aprendizes na documentação do CATUERJ, por exemplo.

O embasamento teórico, metodológico e prático oferecido pela linguística de corpus orientou todo o processo de criação, documentação e alinhamento do CATUERJ. As metodologias já desenvolvidas por pesquisas com corpora de aprendizes abriram espaço para outro tipo de corpora, os corpora de aprendizes de tradução. A partir da criação dos CAT, as pesquisas em Estudos da Tradução com base em corpora ganharam novas perspectivas e possibilidades de caminho. O suporte teórico essencial à análise veio da linguística de corpus, dos Estudos da Tradução com base em corpora e, também, do planejamento da disciplina

Introdução aos Estudos da Tradução I, elaborado por Bianca Walsh (2014) e permitiu que ela decorresse de forma mais completa, pois o recorte do CATUERJ pôde, ao mesmo tempo, ser analisado à luz de teorias fundamentadas por outros CAT já existentes e por teorias já em vias práticas, para que a possibilidade da aplicação do CATUERJ em sala de aula pudesse ser ilustrada de forma mais concreta.

O resultado mais evidente é que as teorias que fundamentaram esta pesquisa puderam ser contrapostas com todos os segmentos selecionados para a análise de forma extensiva, tanto em relação a pesquisas realizadas com outros CAT como aos conteúdos da ementa da disciplina de tradução. Outra consideração interessante é o destaque dado aos pontos positivos das traduções dos aprendizes, principalmente nos achados sobre elementos de coesão e coerência, presentes extensivamente ao longo dos textos traduzidos. Além disso, a tentativa de salientar aspectos positivos das traduções teve a finalidade de desconstruir a imagem negativa e o enfoque nos erros cometidos pelos tradutores aprendizes, reforçando o estigma da quantidade de erros em produções de aprendizes. O presente estudo também permitiu observar que os CAT considerados pequenos podem servir como material para aplicação na própria universidade onde foi desenvolvido, por comportarem conteúdo de origem específica, tornando-se CAT especializados. Outra contribuição evidente dos CAT seria para pesquisas das áreas de ensino e formação de tradutores e Estudos da Tradução com base em corpus. Esta dissertação mostrou que os estudos com base em corpora de aprendizes de tradução encontram-se em fase inicial, havendo apenas dois CAT sendo desenvolvidos e estudados no Brasil. Ambos haviam sido veiculados ao público até a finalização desta pesquisa, além do CATUERJ. Esta pesquisa é, também, o primeiro trabalho sobre o CATUERJ.

Eu não poderia concluir esta dissertação sem ressaltar a característica inovadora do trabalho por enfatizar as aplicações do CATUERJ em sala de aula, destacando o potencial de utilização do CAT para o ensino de tradução. O caráter inovador deste trabalho no Brasil perpassa dois motivos, como mencionado na introdução, que se concretizaram ao final da pesquisa. Primeiramente, a oportunidade de fomentar o diálogo entre pesquisas feitas na universidade e as aplicações de teorias e abordagens criadas nessas pesquisas em salas de aula trazem à universidade uma ponte profícua entre os pilares teóricos e práticos nela desenvolvidos. O segundo motivo é o foco dado aos pontos fortes nas produções dos tradutores aprendizes, colaborando para a construção de uma imagem mais positiva do tradutor em treinamento e ressaltando o processo de aprendizagem cujo acesso é proporcionado pela utilização do CAT como ferramenta didática em sala de aula. Dessa forma, a partir das relações entre teoria e prática, a formação do tradutor aprendiz é

beneficiada pela possibilidade de os alunos analisarem o processo da tradução, para produzirem textos, trabalharem em conjunto e aprimorarem suas traduções, tudo sob o monitoramento do professor na sala de aula, como defende Granger (v. seção 1.1).

Para o desenvolvimento deste estudo, enfrentei limitações de ordem prática, como a grande quantidade de etapas em cada momento da pesquisa. O início da pesquisa, com a listagem dos CAT existentes no mundo, foi a etapa inicial e mais longa do processo. Ela acabou entremeando-se pelas outras etapas da pesquisa, pois, na medida em que uma nova publicação sobre CAT era encontrada, com a publicação vinha a descoberta de mais um CAT. Nessa primeira etapa, além de descobrir novos CATs, também precisei conhecer cada um da forma mais aprofundada possível, para que chegasse à próxima etapa – a compilação, organização e o alinhamento do CATUERJ. Na segunda etapa, a constante comparação entre as informações listadas na primeira etapa e a construção do CATUERJ foi fundamental para que eu compilasse um CAT com estrutura coerente e sólida. No terceiro momento da pesquisa, acompanhei aulas de uma disciplina eletiva sobre tradução (WALSH, 2014), tratando-se de um processo demorado, pois o decorrer da pesquisa dependia do decorrer do semestre no qual fiz minha observação e participação. Já a quarta e última etapa foi laboriosa, pois consistiu na fusão de todas as etapas anteriores para a concepção desta pesquisa, para atingir o objetivo principal da pesquisa – criar possibilidades de aplicação do CATUERJ nas aulas dessa disciplina eletiva sobre tradução (WALSH, 2014).

Outra dificuldade foi a impossibilidade de acesso – parcial ou integral – à maioria dos CAT existentes, pois não estão disponíveis on-line. Além disso, o difícil alcance a publicações relacionadas aos CAT, por confidencialidade ou não gratuidade de acesso a artigos, bem como a descentralização de informações acabaram por prejudicar o recrutamento de fontes para pesquisa de maneira mais eficaz, mesmo sendo de forma eletrônica.

Por fim, restam algumas possibilidades futuras de trabalho com o CATUERJ. A título de ilustração, cito a anotação de erros e de classes gramaticais, que já foi mais explorada em outros CAT já existentes; a análise da lista de frequências de palavras para que os dados do CATUERJ direcionem uma pesquisa sobre aspectos gramaticais, sintáticos ou semânticos, como já vem sendo feito com corpora em geral (de aprendizes ou não); e até mesmo a construção de um ou mais subcorpora do CATUERJ visando à disponibilização desse CAT on-line. Outra possibilidade futura para pesquisa seria a construção de materiais que levem os resultados dessa pesquisa à sala de aula, destacando pontos positivos da produção dos tradutores aprendizes, como uma espécie de seguimento do presente estudo. Antecipo, também, uma indagação que permanece ao final desta dissertação: como desconstruir a

imagem negativa do tradutor em treinamento? Este trabalho, por sua vez, se desenrolou mantendo em sua perspectiva salientar aspectos positivos das traduções com o intuito de desconstruir a imagem negativa e o enfoque nos erros cometidos pelos tradutores aprendizes, o que reforça o estigma da quantidade de erros em produções de aprendizes.

REFERÊNCIAS

- ALVES, F.; MAGALHÃES, C.; PAGANO, A. *Traduzir com autonomia: estratégias para o tradutor em formação*. São Paulo: Contexto, 2000.
- BAKER, M. A corpus-based view of similarity and difference in translation. *International Journal Of Corpus Linguistics*, v. 9, n. 2, p. 167-193, 2004.
- BARBOSA, H. G. *Procedimentos técnicos da tradução*. Campinas: Pontes, 1990.
- BASSNETT, S. Original Sin. In: _____. *Reflections on Translation*. Bristol: Multilingual Matters, 2011, p. 12-15.
- BASSNETT, S. The meek or the mighty. In: ÁLVAREZ, R.; VIDAL, M. (Ed.). *Translation, Power, Subversion*. Clevedon: Multilingual Matters, 1996. p. 10-24.
- BASSNETT, S. Translation or Adaptation? In: _____. *Reflections on Translation*. Bristol: Multilingual Matters, 2011. p. 40-43.
- BENNISON, P.; BOWKER, L. Student Translation Archive and Student Translation Tracking System. Design, Development and Application. In: BERNARDINI, S.; STEWART, D.; ZANETTIN, F. (Ed.). *Corpora in translator education*. Manchester: St. Jerome Publishing, 2003. p. 103-118.
- BERBER SARDINHA, A. P. Linguística de Corpus: Histórico e Problemática. *D.E.L.T.A.*, v. 16, n. 2, p. 323-367, 2000.
- CASTAGNOLI, S. *Regularities and variations in learner translations: A corpus-based study of conjunctive explicitation*. Tese (Doutorado em Linguística Geral) – Universidade de Pisa, 2009.
- CASTAGNOLI, S. et al. Designing a learner translator corpus for training purposes. In: KUBLER, N. (Org.). *Corpora, Language, Teaching and Resources: From Theory to Practice*. Berna: Peter Lang, 2011.
- ESPUNYA, A. The UPF learner translation corpus as a resource for translator training. *Language Resources and Evaluation*, v. 48, n. 1, p. 33-43, 2014.
- ESPUNYA, A. Investigating lexical difficulties of learners in the error-annotated UPF learner translation corpus. In: GRANGER, S.; GILQUIN, G.; MEUNIER, F. (Ed.). *Twenty Years of Learner Corpus Research: Looking back, Moving ahead*. Corpora and Language in Use – Proceedings. Presses Universitaires de Louvain, 2013.
- FLORÉN, C.; LORÉS, R. The application of a parallel corpus English-Spanish to the teaching of translation (ENTRAD project). In: MUÑOZ CALVO, M. et al. (Ed.). *New Trends in Translation and Cultural Identity*. Cambridge: Cambridge Scholars Publishing, 2008. p. 433-444.

- FLOWERDEW, L. The Argument for Using English Specialized Corpora to Understand Academic and Professional Settings. In: CONNOR, U.; UPTON, T. (Ed.). *Discourse in the Professions: Perspectives from Corpus Linguistics*. Amsterdam: John Benjamins, 2004. p. 11-33.
- FREITAS, C. Corpus, Linguística Computacional e as Humanidades Digitais. In: LEITE, M.; TERESA GABRIEL, C. (Org.) *Linguagem, discurso, pesquisa e educação*. 1. ed. Petrópolis: FAPERJ, 2015. p. 23-56.
- FROTA, M. P. Erros e lapsos de tradução: um tema para o ensino. *Cadernos de Tradução*, Santa Catarina, v. 1, n. 17, p. 141-156, jan/jun., 2006. Disponível em: <<https://periodicos.ufsc.br/index.php/traducao/article/view/6859/6411>>. Acesso em: 26 dez. 2015.
- GAUSSIER, E., HULL, D., AÏT-MOKTHAR, S. Term alignment in use: Machine aided human translation. In: VÉRONIS, J. (Org.). *Parallel text processing*. [S.l.]: Kluwer Academic Publishers, 2000. p. 253-74.
- GRANGER, S. A bird's eye view of learner corpus research. In: _____; S.; HUNG, J; PETCH-TYSON, S. (Ed.). *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching*. Amsterdam: Benjamins, 2002. p. 3-33.
- KELLY, D. *A Handbook for Translator Trainers: A Guide to Reflective Practice*. New York: Routledge, 2005.
- KOESTER, A. Building small specialized corpora. In: O'KEEFFE, A. McCARTHY, M. (Ed.). *The Routledge Handbook of Corpus Linguistics*. London: Routledge, 2010.
- KUNILOVSKAYA, M.; KUTUZOV, A. Russian Learner Translator Corpus: Design, Research Potential and Applications. In: *Text, Speech and Dialogue*. Brno: Springer, 2014. p. 315-323.
- KUTUZOV, A. et al. Russian learner parallel corpus as a tool for translation studies. *Computational linguistics and intellectual technologies*. Disponível em: <<http://publications.hse.ru/en/chapters/74015006>>. Acesso em: 11 jan. 2016.
- MARTINS, M. A. P. As Contribuições de André Lefevere e Lawrence Venuti para a Teoria da Tradução. *Cadernos de Letras*, n. 27, p. 59-72, 2010.
- MAY, P. *First Certificate Trainer*. Ed. 6. Cambridge: Cambridge English, 2012.
- NESSSELHAUF, N. Learner corpora and their potential for language teaching. In: SINCLAIR, J. *How to use corpora in language teaching*. Amsterdam: John Benjamins, 2004. p 125-153.
- NOGUEIRA, D.; NOGUEIRA, V. Algumas técnicas para resolver problemas de tradução do inglês para o português. In: CONGRESSO INTERNACIONAL DA ABRATES, 1., 2005, Rio de Janeiro. *Anais...* Rio de Janeiro: [s.n.], 2005. Disponível em: <http://www.iscap.ip.pt/elgg/anaf/files/-1/416/Tecnicas_de_traducao%5B1%5D.pdf>. Acesso em: 17 jan. 2016.

NOTH, W.; SANTAELLA, L. *Imagem: cognição, semiótica, mídia*. São Paulo: Iluminuras, 2008.

OLIVEIRA, J. T. *A linguística de corpus na formação do tradutor: compilação e proposta de análise de um corpus paralelo de aprendizes de tradução*. Dissertação (Mestrado em Letras) – Universidade de São Paulo, 2015.

PRAVEC, N. Survey of learner corpora. *ICAME Journal*. v. 26, p. 81-114, 2002.

RAZMJOU, L. To be a good translator. *Second International Conference on Critical Discourse Analysis: the Message of the Medium*. Iémen, Universidade de Hodeidah, out, 2003. Disponível em: <<http://www.translationdirectory.com/article106.htm>>. Acesso em: 18 jan. 2016.

REPPEN, R. Building a Corpus: What are the key considerations? In: O'KEEFFE, A. McCARTHY, M. (Ed.). *The Routledge Handbook of Corpus Linguistics*. Londres: Routledge, 2010.

SANTOS, D. PoNTE: apontando para corpos de aprendizes de tradução avançados. *LínguaMÁTICA*, v. 6, n. 1, p. 69-86, jul. 2014.

SCHIFFRIN, A. Preface. In: _____. *The business of books: How International Conglomerates Took Over Publishing and Changed the Way We Read*. Londres: Verso, 2001. p. 01-14.

SHEPHERD, T. O Estatuto da Linguística de Corpus: Metodologia ou Área da Linguística? *Matraga*, Rio de Janeiro, v. 16, n. 24, p. 150-172, jan./jun. 2009.

SHIELDS, K. *Gained in Translation: Language, Poetry, and Identity in Twentieth-Century Ireland*. Berna: Peter Lang, 2000.

SPENCE, R. A Corpus of Student L1-L2 Translations. In: GRANGER, S.; HUNG, J. (Ed.). *Learner Corpora, Second Language Acquisition and Foreign Language Teaching*. International Symposium on Computer. Hong-Kong: The Chinese University of Hong Kong, 1998.

ŠTĚPÁNKOVÁ, K. *Learner Translation Corpus: CELTraC (Czech-English Learner Translation Corpus)*. Dissertação (Mestrado em Língua e Literatura inglesa) – Faculdade de Artes, Universidade de Masaryk, 2014.

STUPIELLO, E. Tecnologias de tradução: implicações éticas para a prática tradutória. *TradTerm*, São Paulo, v. 19, p. 71-91, nov. 2012.

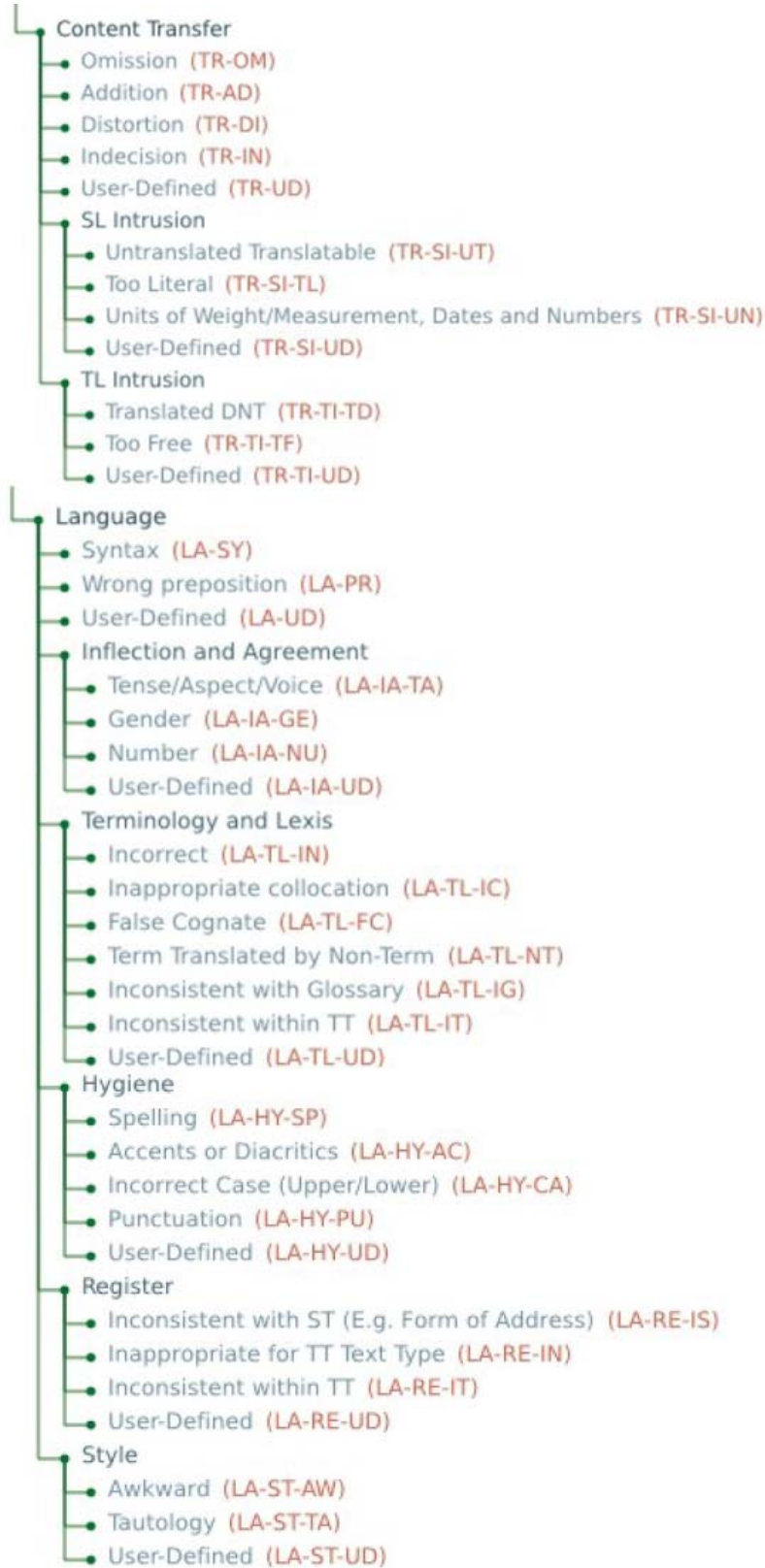
UZAR, R. A corpus methodology for analyzing translation. *Cadernos de Tradução*, Santa Catarina, v. 1, n. 9, p. 235-263, jan./jun. 2008. Disponível em: <<https://periodicos.ufsc.br/index.php/traducao/article/view/5988>>. Acesso em: 11 jan. 2016.

VENUTI, L. How to read a translation. In: _____. *Translation changes everything: theory and practice*. Londres: Routledge, 2013.

- WALSH, B. Uma experiência de planejamento de ensino de tradução: conjugando competência e capacidade de renovação na formação de tradutores autocríticos. *Cadernos de Tradução*, Florianópolis, n. 34, p. 206-228, jul./dez. 2014.
- WURM, A. Eigennamen und Realia in einem Korpus studentischer Übersetzungen (KOPTE). *Trans-kom*, v. 6, n. 2, 2013.
- WYLER, Lia. *Línguas, poetas e bacharéis: uma crônica da tradução no Brasil*. Rio de Janeiro: Rocco, 2003.
- ZANETTIN, F. Corpus linguistics and translation studies. *Translation-driven corpora*. Londres: Routledge, 2014.
- ZANETTIN, F. *Corpora in Translation Practice*. Universidade para Estrangeiros de Perugia. Perugia. 2002.

ANEXO A – Tipologia de erros do MeLLANGE

Figura 1 – Fluxograma da tipologia de erros do MeLLANGE



Fonte: ZANETTIN, 2014, f. 2.4.

ANEXO B – Tipologia de erros do CELTraC

Figura 2 – Fluxograma da tipologia de erros do CELTraC



ANEXO C – Propaganda de uma associação de cardiologiaFigura 3 – Propaganda da *American Heart Association*

**HIT THE
BARS
AFTER WORK.**

Happy hour is happier when the recreation you choose is good for both body and mind. That's why more and more people are devoting that time to biking, walking or working out. It can improve your outlook, increase your energy and decrease your risk of heart disease. And it's guaranteed not to cause a morning hangover. To learn more, visit our web site at www.americanheart.org or call 1-800-AHA-USA1.

American Heart Association
Fighting Heart Disease and Stroke

This space provided as a public service © 1998 American Heart Association

Fonte: ALVES, F.; MAGALHÃES, C.; PAGANO, A., 2000, f. 1.3.