

UNIVERSIDADE FEDERAL DE SANTA CATARINA  
PÓS-GRADUAÇÃO EM ESTUDOS DA TRADUÇÃO

NORMALIZAÇÃO LEXICAL EM TRADUÇÕES DE  
*DOM CASMURRO*, DE MACHADO DE ASSIS:  
UM ESTUDO BASEADO EM CORPUS

MESTRANDO: IURI ABREU

ORIENTADOR: PROF. DR. WERNER HEIDERMANN

CO-ORIENTADOR: PROF. DR. LINCOLN FERNANDES

Florianópolis, setembro de 2007.



“The most important skill is not to be able to program a computer or even to manipulate available software. Rather, it is to be able to ask insightful questions which address real issues and problems in theoretical, descriptive and applied language studies.”

Graeme Kennedy

## AGRADECIMENTOS

Ao Werner e Lincoln, pelos conhecimentos transmitidos.

À Tatiana, por suportar a separação forçada.

A meus pais, que sempre me apoiaram.

## RESUMO

Esta dissertação consiste na investigação da ocorrência do fenômeno da normalização lexical na tradução de itens lexicalmente criativos, representados pelo uso de colocações marcadas. Para tanto, são analisadas três traduções para a língua inglesa do romance *Dom Casmurro*, de Machado de Assis. Por tratar-se de uma abordagem baseada em corpus, foi desenvolvido um corpus paralelo composto do original em português e suas respectivas traduções. Além disso, dois outros corpora são utilizados para referência: o British National Corpus (BNC), para o inglês, e o corpus do NILC, para o português brasileiro. Os resultados apontaram para uma ocorrência significativa de normalização por parte de todos os tradutores, representando um índice de quase 80% sobre o total de traduções avaliadas.

Palavras-chave: Criatividade lexical, normalização, estudo baseado em corpus, Machado de Assis.

## ABSTRACT

This master's thesis aims at investigating the occurrence of a phenomenon called lexical normalization in the translation of lexically creative items, represented by the use of marked collocations. To do so, three translations into English of the novel *Dom Casmurro*, by Machado de Assis, are analyzed. Since it is a corpus-based approach, a parallel corpus was developed, which is composed of the original work in Portuguese and its respective translations. In addition, two other corpora are used as reference: The British National Corpora (BNC), for the English language, and the NILC corpus, for Brazilian Portuguese. The results pointed to a significant occurrence of normalization by all translators under investigation, representing a rate of nearly 80% over the total number of assessed translations.

Keywords: Lexical creativity, normalization, corpus-based study, Machado de Assis.

## SUMÁRIO

RESUMO.....	5
ABSTRACT.....	6
LISTA DE TABELAS.....	8
LISTA DE GRÁFICOS.....	9
1. INTRODUÇÃO.....	10
1.1 CONTEXTUALIZAÇÃO DA PESQUISA.....	10
1.2 TIPO DE PESQUISA.....	11
1.3 ESTUDOS DA TRADUÇÃO BASEADOS EM CORPORA.....	12
2. JUSTIFICATIVA.....	13
3. OBJETIVO.....	16
4. REVISÃO DA LITERATURA – CONCEITOS FUNDAMENTAIS.....	17
4.1 Normalização lexical.....	17
4.2 Princípio idiomático e colocação.....	19
4.3 Colocações marcadas e não-marcadas.....	20
4.4 Machado de Assis.....	21
4.4.1 O autor.....	21
4.4.2 A obra.....	22
4.4.3 Machado em inglês.....	23
5. MÉTODO.....	25
5.1 Digitalização dos textos.....	25
5.2 Corpus paralelo.....	26
5.3 Corpora de referência.....	28
5.3.1 Corpus em Língua Portuguesa.....	28
5.3.2 Corpus em Língua Inglesa.....	30
6. ANÁLISE DOS DADOS.....	32
6.1 Escolha do nóculo.....	32
6.2 O nóculo “olhos”.....	33
6.3 Análise das linhas selecionadas.....	36
7. CONCLUSÃO.....	60
REFERÊNCIAS.....	65
Anexo 1.....	68
Anexo 2.....	72
Anexo 3.....	75

## LISTA DE TABELAS

Tabela 1 – Exemplos de colocações marcadas.....	18
Tabela 2 – Dimensões do corpus de estudo.....	24
Tabela 3 – Lista de palavras-chave de <i>Dom Casmurro</i> no original.....	30
Tabela 4 – Linhas selecionadas da concordância do nóculo <i>olhos</i> .....	33



## LISTA DE GRÁFICOS

Gráfico 1 – Ocorrência de normalização na tradução de colocações marcadas, de acordo com cada tradutor.....	58
Gráfico 2 – Ocorrência de normalização considerando o total de traduções avaliadas.....	59

# 1. INTRODUÇÃO

## 1.1 CONTEXTUALIZAÇÃO DA PESQUISA

Orientando-se pelo mapa do território dos Estudos da Tradução traçado em artigo seminal de James Holmes, em 1972, podemos situar nossa pesquisa sob a grande área de *pure translation studies*. Esta, por sua vez, apresenta dois objetivos: (1) descrever os fenômenos do ato tradutório e da(s) tradução(ões) da forma como eles acontecem no mundo de nossa experiência e; (2) estabelecer princípios gerais pelos quais tais fenômenos podem ser explicados e previstos (HOLMES, 1988, p. 71)<sup>1</sup>. Estamos em sintonia com o primeiro objetivo, o qual Holmes denomina *descriptive translation studies* (DTS). Há ainda uma subdivisão dos estudos descritivos em três tipos de pesquisa: *product-oriented DTS*, *function-oriented DTS* e *process-oriented DTS*. O primeiro deles merece uma melhor definição, considerando que este projeto nele se inclui. Segundo definição do próprio Holmes, trata-se de uma

área de pesquisa que descreve traduções existentes, tradicionalmente tem sido uma área importante de pesquisa acadêmica nos estudos da tradução. O ponto de partida para este tipo de estudo é a descrição de traduções individuais, ou a descrição de traduções com foco no texto. Uma segunda fase é aquela da descrição da tradução comparativa, na qual são feitas análises comparativas de diversas traduções do mesmo texto (ibid., p. 72)<sup>2</sup>.

Pode-se fazer uma relação direta com este estudo, pois o foco está na análise de três traduções de um mesmo texto.

De acordo com as doze áreas de pesquisa listadas por Williams & Chesterman (2002, p. 6), este projeto insere-se na área *comparação de traduções com seus textos-fonte*, a qual, por sua vez, é uma subdivisão da área *análise textual e tradução*. Segundo os autores, este tipo de pesquisa

---

<sup>1</sup> “(1) to describe the phenomena of translating and translation(s) as they manifest themselves in the world of our experience, and (2) to establish general principles by means of which these phenomena can be explained and predicted”. Tradução minha para este e demais trechos subsequentes que ainda não tenham sido traduzidos para o português.

<sup>2</sup> “area of research which describes existing translations, has traditionally been an important area of academic research in translation studies. The starting point for this type of study is the description of individual translations, or text-focused translation description. A second phase is that of comparative translation description, in which comparative analyses are made of various translations of the same text”.

envolve a comparação textual de uma tradução com seu original. Uma comparação de tradução lida com diversas traduções, para a mesma língua ou para línguas diferentes, do mesmo original. Tais tópicos não podem, é claro, lidar com todos os aspectos possíveis dos textos, por isso é preciso escolher o(s) aspecto(s) que se quer enfatizar (loc.cit.)<sup>3</sup>.

Em meu caso, trabalharei com três traduções do mesmo original, e o aspecto escolhido foi a normalização lexical manifestada pelo uso de colocações marcadas (ver Capítulo 4, Seção 3).

## 1.2 TIPO DE PESQUISA

Ainda tomando por base os conceitos estabelecidos na obra *The Map*, de Williams & Chesterman (2002), situaremos nossa pesquisa de acordo com as duas divisões propostas pelos autores – conceitual e empírica. Considerando que a pesquisa empírica é aquela que “(...) busca novos dados, novas informações derivadas da observação de dados e de trabalho experimental; busca evidência que confirme ou descarte hipóteses, ou gere novas hipóteses” (op.cit., p. 58)<sup>4</sup>, pode-se afirmar que este estudo possui caráter predominantemente empírico. Como confirmação, Williams & Chesterman citam, a título de ilustração de métodos de pesquisa empírica, os estudos baseados em corpora (op.cit., p. 66), que é justamente a abordagem adotada nesta dissertação.

Quanto à subdivisão da pesquisa empírica – naturalista (ou observacional) ou experimental –, este trabalho apresenta natureza naturalista, uma vez que se propõe “investigar um fenômeno ou processo como acontece na vida real em seu ambiente natural” (op.cit., p. 62)<sup>5</sup>. Além disso, apenas observações serão feitas, sem qualquer tipo de interferência nos dados, outra característica da pesquisa empírica naturalista. A mera seleção de uma obra e suas traduções (e não uma outra obra, ou outras traduções) não é suficiente para classificar a pesquisa como experimental. Para tanto, seria necessário, por

---

<sup>3</sup> “involves the textual comparison of a translation with its original. A translation comparison deals with several translations, into the same language or into different languages, of the same original. Such topics cannot deal with every possible aspect of the texts, of course, so you have to choose the aspect(s) you want to focus on”.

<sup>4</sup> “(...) seeks new data, new information derived from the observation of data and from experimental work; it seeks evidence which supports or disconfirms hypotheses, or generate new ones”.

<sup>5</sup> “investigate a phenomenon or a process as it takes place in real life in its natural setting”.

exemplo, escolher um trecho da obra sendo investigada, entregá-lo a um ou mais tradutores e dar instruções de como proceder, seja registrando suas soluções, seja preenchendo um questionário, ou qualquer outra forma que contasse com a interferência direta do pesquisador. Isso posto, pode-se afirmar que esta pesquisa tem caráter empírico e naturalista.

### 1.3 ESTUDOS DA TRADUÇÃO BASEADOS EM CORPORA

O arcabouço metodológico que dá sustentação a este estudo tem origem nos estudos da tradução baseados em corpora (CTS, na sigla em inglês para *corpus-based translation studies*). Trata-se de uma área de pesquisa relativamente jovem, somente tornada viável após o desenvolvimento de tecnologias que permitissem analisar grandes quantidades de texto com precisão e velocidade. Naturalmente, isso só foi possível através do uso dos microcomputadores, os quais permitem que o pesquisador manipule seu corpus ou corpora tendo por pré-requisitos somente um PC e algum software que execute algumas operações básicas (WordSmith Tools, por exemplo).

Pode-se afirmar que os CTS são o resultado da combinação da lingüística de corpus com os estudos descritivos da tradução (DTS, derivado do termo em língua inglesa *descriptive translation studies*). Olohan (2004) sustenta essa afirmação ao constatar que “os estudos baseados em corpora estão claramente alinhados com a perspectiva descritiva” (p. 10)<sup>6</sup>. Essa mesma autora declara que podemos considerar o uso de corpora nos estudos da tradução como sendo “a aplicação de técnicas de análise de corpus, tanto quantitativas como qualitativas, ao estudo de aspectos do produto e do processo da tradução” (ibid., p. 9)<sup>7</sup>. No caso da presente pesquisa, uma combinação de análise quantitativa (para verificar o número de ocorrências de determinado nódulo, por exemplo) e qualitativa (interpretação dos resultados, considerando cada caso isoladamente) será aplicada para avaliar aspectos do produto – e não do processo, que não faz parte do objeto de estudo desta pesquisa – da tradução.

---

<sup>6</sup> “Corpus-based studies in translation are clearly aligned with the descriptive perspective”.

<sup>7</sup> “the application of corpus analysis techniques, both quantitative and qualitative, to the study of aspects of the product and process of translation”.

## 2. JUSTIFICATIVA

A presente pesquisa tenta preencher uma lacuna na área dos estudos da tradução. Essa constatação encontra suporte no CD-ROM “Estudos da Tradução no Brasil: Resumos de Teses e Dissertações”, organizado por Pagano et al. (2001), no qual as autoras fazem um levantamento da produção de teses e dissertações sobre tradução por pesquisadores brasileiros. Fazendo uma busca pelas palavras-chave *normalização* e/ou *criatividade lexical* e/ou *colocações marcadas*, não foi encontrado nenhum estudo, o que aponta para a necessidade de se desenvolver pesquisas abrangendo esses aspectos. Mais recentemente, tomei conhecimento de um trabalho de análise de criatividade lexical com abordagem baseada em corpora utilizando um corpus paralelo. Trata-se da dissertação de mestrado de Bueno (2005), da Universidade Federal de Minas Gerais, que difere da presente pesquisa por apoiar-se na gramática sistêmico-funcional de Halliday, além de analisar a transitividade e a coesão no corpus paralelo composto de *Macunaíma*, de Mário de Andrade, e sua tradução para o inglês, de E. A. Goodland.

A fim de corroborar ou refutar os achados e as hipóteses levantadas por Kenny (2001), faz-se necessário que replicações sejam realizadas, de modo a ampliar e consolidar um determinado conhecimento. A própria autora é categórica ao afirmar que “há uma necessidade óbvia (...) de mais pesquisas empíricas utilizando corpora paralelos e comparáveis monolíngües” (Kenny, 1998)<sup>8</sup>. Por tratar-se de uma área de conhecimento que ainda dá seus primeiros passos, os estudos da tradução baseados em corpora se beneficiarão de estudos sobre normalização lexical envolvendo pesquisadores trabalhando com diferentes pares de línguas. Só assim será possível chegar a conclusões mais abrangentes em relação a esse fenômeno tradutório.

Em adição, tenho por objeto de estudo um dos maiores escritores em língua portuguesa, que já foi extensamente analisado por críticos literários. Todavia, há uma escassez de material sobre as traduções de Machado de Assis, e nenhuma que seja de meu conhecimento versando especificamente sobre as traduções de *Dom Casmurro* para língua inglesa. Esta pesquisa representa, portanto, uma contribuição original, sob uma perspectiva inédita, à vasta bibliografia dedicada a um dos mestres de nossa literatura.

---

<sup>8</sup> “there is an obvious need (...) for more empirical research using parallel and monolingual comparable corpora”.

A opção por trabalhar com o romance *Dom Casmurro* e suas traduções para a língua inglesa teve origem no segundo semestre de 2005, quando o Prof. John Gledson esteve na Universidade Federal de Santa Catarina (UFSC), em caráter de professor visitante, para ministrar um curso aos alunos do Programa de Pós-Graduação em Estudos da Tradução (PGET). Ao longo das aulas – que possuíam teor predominantemente prático – pudemos constatar as dificuldades que o texto machadiano, tanto contos quanto romances, apresentava ao tradutor, em função de diversas características que foram sendo evidenciadas a partir do cotejo do original com as traduções existentes (inclusive para idiomas como o francês, o italiano e o espanhol). O estilo conciso de Machado de Assis, a dificuldade de verter os adjetivos, a tentativa de manter o tom do original, entre outros, foram os principais problemas evidenciados pelo tradutor<sup>9</sup>. Em ensaio desenvolvido para concorrer a um prêmio oferecido pelo Ministério da Cultura de nosso país, a ser dado ao melhor ensaio escrito por um estrangeiro sobre o autor de Quincas Borba, Gledson ainda cita a “sutileza e o ritmo narrativo de Machado” (Gledson, 2005, p. 19) como dificuldades tradutórias.

A partir de uma experiência prática de tradução, discutindo em aula as escolhas feitas por Gledson e outros tradutores de Machado para o inglês, percebi que estava diante de um objeto de pesquisa ao mesmo tempo rico e fascinante, além de pouco explorado: a representação de um dos ícones da literatura brasileira através da tradução. Entretanto, por se tratar de um tema aberto a múltiplas abordagens, fez-se necessário delimitar a pesquisa a algum item de especial importância e relevância.

Em razão de uma inclinação pessoal pelos estudos da tradução baseados em corpora, procurei por um aspecto relacionado a essa área, o qual fosse passível de investigação por meio de um corpus paralelo composto da obra *Dom Casmurro* em português e suas traduções para a língua inglesa. Inicialmente, considerei a hipótese de pesquisar a questão dos adjetivos e/ou sintagmas nominais referentes a partes do corpo, motivado pela célebre descrição de Capitu, uma das personagens principais do romance: “olhos de ressaca”. A dificuldade em traduzir esse trecho reside, em grande parte, na associação incomum entre os nódulos que o compõem, o que nos remete ao conceito de colocação, tema central em uma abordagem baseada em corpus. Fazendo uma relação entre

---

<sup>9</sup> Gledson, além de traduzir *Dom Casmurro* para o inglês, também já traduziu alguns contos machadianos para coletâneas de língua inglesa.

“associação incomum” e “colocação”, obtém-se o que Baker (1992) denominou “colocação marcada”<sup>10</sup>.

Até esse momento, portanto, estava definido que a pesquisa se desenvolveria a partir de um corpus paralelo literário contendo a obra *Dom Casmurro* e suas traduções para o inglês, enfatizando a questão das colocações marcadas. Foi então que entrou em cena o estudo *Lexis and creativity in translation: a corpus-based study*, relatado por Kenny (2001). A autora também fez uso de um corpus paralelo literário, embora do par alemão-inglês, com o qual investigou a ocorrência da normalização lexical de itens lexicalmente criativos, entre os quais se incluíam as colocações marcadas. Definiu-se, então, o objetivo principal deste projeto: a investigação da normalização lexical nas traduções para língua inglesa de colocações marcadas em *Dom Casmurro*, de Machado de Assis.

Até o momento, há três traduções do romance *Dom Casmurro* publicadas em inglês, em diferentes épocas e por tradutores de nacionalidades distintas, a saber, Helen Caldwell (norte-americana, em 1953), Scott Buccleuch (escocês, em 1992) e John Gledson (inglês, em 1997). Optou-se por incluir todas as traduções existentes na análise para, além de ampliar o corpus de pesquisa, ter condições de compará-las e verificar as diferentes soluções adotadas por cada um dos tradutores.

---

<sup>10</sup> “marked collocation”.

### 3. OBJETIVO

O objetivo principal deste trabalho é a investigação da ocorrência ou não do fenômeno de uma das hipóteses da tradução, inicialmente propostas por Baker (1996), a saber, a normalização. Para isso, restringirei minha pesquisa ao nível lexical, uma vez que o foco deste estudo recai sobre as colocações marcadas como forma de criatividade lexical, fazendo uso de um corpus paralelo composto por uma das obras de maior relevância na literatura brasileira – *Dom Casmurro*, de Machado de Assis – e três traduções para a língua inglesa. As colocações marcadas exercem importante papel no que tange ao aspecto criativo do romance, pois representam um recurso do qual o autor se utiliza para criar um estilo próprio.

Isso posto, surgem algumas perguntas, para as quais a presente pesquisa busca fornecer respostas: (i) as colocações marcadas encontradas no original foram normalizadas pelos tradutores?; (ii) se o foram, de que forma se deu essa normalização?; (iii) há casos em que a normalização ocorreu com maior intensidade?; (iv) pode-se fazer generalizações a respeito de como os tradutores costumam lidar com a questão das colocações marcadas?; e (v) tais generalizações são válidas simplesmente para o par português-brasileiro/inglês, ou são corroboradas por outros estudos relacionados ao mesmo tema<sup>11</sup>?

A tentativa de determinar o porquê das escolhas adotadas pelos tradutores deverá permear o estudo. No entanto, vale ressaltar que tais julgamentos, muitas vezes, terão caráter especulativo, em função do alto grau de subjetividade envolvido em qualquer análise das escolhas feitas por um tradutor, não devendo ser incluído, portanto, entre os objetivos principais desta pesquisa.

---

<sup>11</sup> Em razão da escassez de trabalhos sobre normalização lexical na literatura, o principal estudo para fins de comparação será aquele desenvolvido por Kenny (2001).



## 4. REVISÃO DA LITERATURA – CONCEITOS FUNDAMENTAIS

Esta seção apresenta os conceitos fundamentais para a compreensão da presente pesquisa, bem como pontos de convergência e divergência com o estudo desenvolvido por Kenny (2001), intitulado *Lexis and creativity in translation: a corpus-based study*.

### 4.1 Normalização lexical

Fazendo uso de um corpus paralelo (GEPCOLT), composto de textos literários do par alemão-inglês, a autora investiga o fenômeno da normalização na tradução de itens lexicalmente criativos. A normalização do ponto de vista do usuário de um corpus paralelo, segundo definição da própria Kenny, ocorre sempre que “os tradutores optam por soluções convencionais na língua-alvo para problemas criados por características incomuns ou criativas do texto-fonte” (op.cit., p. 66)<sup>12</sup>, ou seja, é a tendência que os tradutores apresentam em preferir as normas da língua-alvo, fato que já fora incluído por Baker (1993) entre as hipóteses da tradução<sup>13</sup>. Normalização lexical – principal objeto de estudo de Kenny – pode ser definida, portanto, como “a normalização no nível de palavras individuais e colocações” (loc.cit.)<sup>14</sup>.

O enfoque dado por Kenny recai sobre a criatividade lexical, que a autora busca, no corpus paralelo especificamente desenvolvido para a pesquisa, através de *hapax legomena* (itens lexicais que ocorrem uma única vez no corpus), formas específicas de autor (itens criativos que se repetem apenas na obra de um mesmo autor) e colocações incomuns.

Para desenvolver sua pesquisa, Kenny utiliza dois corpora de referência: um de língua inglesa (BNC), e outro para o alemão (IDS Mannheim Corpus). Uma vez que um dos objetivos é verificar o que acontece, na tradução para o inglês, com as colocações marcadas em textos originalmente escritos em alemão, a autora justifica sensatamente o uso de um corpus de referência na língua-alvo:

---

<sup>12</sup> “translators opt for conventional target language solutions to problems posed by creative or unusual source text features”.

<sup>13</sup> As outras hipóteses da tradução são: simplificação, explicitação e estabilização.

<sup>14</sup> “normalization at the level of individual words and collocations”.

o foco não pode estar em colocações do alemão consideradas marcadas do ponto de vista da língua inglesa (...), mas, pelo contrário, deve estar em colocações da língua-alvo consideradas marcadas do ponto de vista da língua-alvo” (op.cit., p. 84)<sup>15</sup>.

A justificativa para o uso de um corpus de referência em alemão, mesma língua dos textos originais, é verificar até que ponto determinados itens lexicais ou colocações incomuns podem ser considerados criativos ou não.

A abordagem adotada por Kenny é baseada em corpus. Minha pesquisa, seguindo essa mesma linha, apresenta diversos pontos em comum com o trabalho desenvolvido por Kenny (ibid.). À semelhança da pesquisadora da Universidade de Dublin, meu objeto de estudo inicial é um corpus paralelo composto de textos literários. A principal diferença é que estarei trabalhando com apenas uma obra na língua-fonte (*Dom Casmurro*, de Machado de Assis) e três traduções dessa mesma obra para a língua inglesa.

A ênfase dada na criatividade lexical permanece a mesma. Porém, enquanto Kenny investigou três manifestações distintas de itens lexicalmente criativos – *hapax legomena*, formas específicas de autor e colocações incomuns –, limitarei minha investigação a este último item. A opção pela pesquisa de itens criativos entre os *hapax legomena* só seria justificada se o corpus constasse de escritores reconhecidamente inovadores em termos de criação de palavras novas. Machado, apesar de ser um escritor de estilo único, é bastante conservador no que tange aos neologismos, diferentemente do que acontece com os autores modernistas, como Oswald e Mário de Andrade, por exemplo. Uma boa ilustração desse fato pode ser encontrada em Bueno (2005). Ao investigar a criatividade lexical na tradução, embora com base na lingüística sistêmico-funcional, a autora elege a obra *Macunaíma*, de Mário de Andrade, como seu objeto de estudo. Nesse caso, a investigação dos *hapax legomena* revela formas criativas como *esgurejou*, *desesquentou* e *urarizaram* (op.cit., p. 71), entre outras.

Em relação às formas específicas de autor, o uso somente seria válido se o corpus paralelo constasse de, no mínimo, dois autores diferentes. Uma vez que o corpus utilizado nesta pesquisa é composto por um único autor, está descartada a hipótese de investigação desse outro tipo de criatividade lexical.

---

<sup>15</sup> “the focus cannot be on German collocations considered remarkable from an English-language point of view (...), but rather must be on source-language collocations considered remarkable from a source-language point of view”.

Por outro lado, as colocações incomuns – ou marcadas, na definição de Baker (1992) – são as formas de representação de criatividade lexical mais indicadas para investigação no corpus em questão. Um único exemplo serve de confirmação para essa afirmação: “olhos de ressaca”. Provavelmente a colocação mais famosa em *Dom Casmurro*<sup>16</sup>, essa descrição dos olhos de Capitu é um exemplo perfeito de colocação incomum, motivo pelo qual é de extremo interesse do pesquisador investigar se houve normalização lexical nas traduções para o inglês.

#### 4.2 Princípio idiomático e colocação

Há dois princípios interpretativos para explicar as características de organização de um texto, a saber, o princípio de escolha aberta (*open-choice principle*) e o princípio idiomático (*idiom principle*). O primeiro, que diz respeito à visão habitual que se tem da linguagem, sustenta que a produção da linguagem se dá através de uma série de escolhas – o chamado modelo *slot-and-filler* –, onde “a cada ponto que uma unidade se completa (uma palavra, sintagma ou oração), uma ampla gama de opções se abre e a única restrição é a gramaticalidade” (SINCLAIR, 1991, p. 109)<sup>17</sup>. Na prática, entretanto, o que acontece é que as opções, em determinadas situações (quando se faz uma escolha de registro, por exemplo), são enormemente reduzidas. Para resolver esse impasse e dar conta das restrições que o modelo de escolha aberta não consegue explicar, Sinclair propôs um segundo modelo interpretativo da linguagem: o princípio idiomático – ou princípio colocacional, na terminologia de Partington (1998). Neste, “o usuário da língua tem disponível um grande número de frases semi-pré-construídas que constituem escolhas únicas, mesmo que pareçam analisáveis em segmentos” (ibid., p. 110)<sup>18</sup>. O princípio idiomático é uma das principais bases teóricas da lingüística de corpus e, por conseqüência, a base para o trabalho desenvolvido nesta pesquisa. É a partir dele que se torna possível realizar uma investigação de criatividade lexical em colocações, que nada mais são do que uma ilustração do princípio idiomático. Como bem observou Kenny (2001),

---

<sup>16</sup> Tanto que, na tradução francesa, foi incluída como subtítulo do romance: *Dom Casmurro et les yeux de ressac*.

<sup>17</sup> “at each point where a unit is completed (a word or a phrase or a clause), a large range of choice opens up and the only restraint is grammaticalness”.

as colocações usuais são uma manifestação importante do princípio idiomático, mas as colocações incomuns geram mudanças. Elas representam pontos em que interpretações padrão não podem ser aplicadas. Assim, elas exibem um tipo especial de marcação que os tradutores podem ou não procurar recriar” (p. 100)<sup>19</sup>.

O termo “colocação” foi cunhado pelo lingüista Firth (1957) e ilustrado por um famoso dito: “you shall judge a word by the company it keeps”.

Pode-se encontrar diversas definições de colocação na literatura (Leech, 1974; Sinclair, 1991a; Hoey, 1991; Baker, 1992). No entanto, utilizaremos a definição proposta por Kenny (2001): “colocação refere-se à co-ocorrência de itens lexicais não interpretados semanticamente dentro de uma distância especificada entre um e outro em textos naturais” (p. 84)<sup>20</sup>. A palavra-chave nessa definição é “co-ocorrência”. Por exemplo, *idéia* tem uma probabilidade maior de co-ocorrer com *brilhante*, *absurda*, *fixa* e *ter* do que com *geladeira*, *janela* ou *desligar*.

Vale ressaltar que o significado nem sempre é responsável pelo padrão de colocação e que é comum sinônimos ou quase-sinônimos apresentarem diferentes conjuntos de colocados (Baker, 1992, p. 47). Para verificar a veracidade dessas afirmações, basta mencionar que alguns falantes do português praticam *magia negra*, mas, dificilmente, algum deles teme a *magia preta*<sup>21</sup>.

#### 4.3 Colocações marcadas e não-marcadas

Ao analisar as traduções do romance *Dom Casmurro* para a língua inglesa, estaremos direcionando o foco de atenção à questão das colocações marcadas, que servem como evidência de criatividade lexical.

Antes de progredir, faz-se necessário esclarecer a sua aceção, diferenciando-as das colocações não-marcadas. Utilizaremos a definição fornecida por Mona Baker: “uma

---

<sup>18</sup> “a language user has available to him or her a large number of semi-preconstructed phrases that constitute single choices, even though they might appear to be analyzable into segments”.

<sup>19</sup> “habitual collocations are one important manifestation of the idiom principle, but unusual collocations occasion switches. They represent points at which default interpretations cannot apply. They thus exhibit a special type of markedness that translators may or may not seek to recreate”.

<sup>20</sup> “collocation refers to the co-occurrence of semantically uninterpreted lexical items within a specified distance of each other in naturally occurring text”.

<sup>21</sup> No corpus do NILC, há três ocorrências para “magia negra” e nenhuma para “magia preta”.

colocação marcada é uma combinação incomum de palavras, que desafia nossas expectativas enquanto ouvintes ou leitores” (Baker, 1992, p. 51)<sup>22</sup>.

As colocações marcadas são freqüentes em ficção, poesia, humor e propaganda (loc.cit.), pois são uma maneira de usar a língua com criatividade. Alguns exemplos de colocações marcadas podem ser observados na Tabela 1.

1. “Era o homem de maiores ruindades calmas que já se viu.” – Guimarães Rosa
2. “A casa olhava para ele.” – Machado de Assis
3. “O amor bate na aorta.” – Carlos Drummond de Andrade
4. “Quem tem boca vai e arruma.” – Jô Soares
5. “Tupi or not tupi, that is the question.” – Oswald de Andrade

**Tabela 1.** Exemplos de colocações marcadas.

Por contraste, colocação não-marcada é uma combinação de palavras com a qual os falantes nativos de uma língua já estão acostumados. Ao ouvir ou ler uma frase como “Luísa tem **lindos** olhos **azuis**”, não há nada de singular com as co-ocorrências para o nódulo olhos (em negrito). Por outro lado, quando Machado de Assis descreve a personagem Capitu como tendo “olhos de ressaca”, gera um estranhamento pela associação incomum.

#### 4.4 Machado de Assis

##### 4.4.1 O autor

Joaquim Maria Machado de Assis (Rio de Janeiro, 1839-1908) nasceu no Morro do Livramento, filho de um operário mestiço de negro e português. Publica seu primeiro trabalho literário aos 16 anos, no jornal *Marmota Fluminense*. Seu primeiro romance, *Ressurreição*, sai em 1872, época em que já contribuía regularmente para jornais e revistas cariocas. Autodidata, Machado incursionou em diversos gêneros literários – romance, poesia, conto, crônica, teatro e ensaio –, além de crítica literária e teatral e tradução (entre

---

<sup>22</sup> “a marked collocation is an unusual combination of words, one that challenges our expectations as hearers

elas, *Os trabalhadores do mar*, de Victor Hugo, publicada em 1866). Aos 30 anos, casa-se com Carolina Augusta de Novais, moça de origem portuguesa e de família de intelectuais, que falece em 1904 sem lhe deixar filhos. Participa da criação da Academia Brasileira de Letras – por vezes chamada Casa de Machado de Assis –, sendo o principal responsável pela sobrevivência e pelo prestígio do novo instituto e exercendo o cargo de presidente até sua morte, em 1908.

Machado é unanimidade entre os críticos. A bibliografia sobre sua vida e obra é vastíssima e continua a crescer<sup>23</sup>. Os manuais de literatura brasileira têm dificuldade em classificá-lo nos movimentos literários existentes, sendo muitas vezes tratado como um caso à parte, da mesma forma que Augusto dos Anjos, por exemplo. Afrânio Coutinho afirma que Machado “atravessou incólume todos os movimentos e escolas, constituindo um mundo à parte, um estilo composto de técnicas precisas e eficazes” (Coutinho, 1986, p. 152).

#### 4.4.2 A obra

Costuma-se dividir a obra machadiana em duas fases: a primeira encerra-se com o romance *Iaiá Garcia*, de 1878; a segunda começa com *Memórias póstumas de Brás Cubas*, em 1881, e termina com o último romance de Machado, *Memorial de Aires*, de 1908.

A obra utilizada nesta pesquisa faz parte da segunda fase. Publicado em 1900, é o romance mais traduzido de Machado<sup>24</sup>, já sendo referido como o romance mais famoso da literatura brasileira. John Gledson – crítico literário especializado em Machado e um de seus tradutores para o inglês – argumenta que *Dom Casmurro* não está tão longe do Realismo, como sustentam alguns críticos:

(...) se entendermos por realismo a intenção do romancista de revelar, através da ficção, a verdadeira natureza da sociedade que está retratando, *Dom Casmurro* é um romance realista, não apenas em termos genéricos, mas em seus detalhes, tanto na forma como no conteúdo. (Gledson, 1991, p. 13)

---

or readers”.

<sup>23</sup> O lançamento mais recente é a biografia *Machado de Assis, um gênio brasileiro*, de Daniel Piza.

<sup>24</sup> *Dom Casmurro* já foi traduzido para os seguintes idiomas: inglês, alemão, espanhol, francês, italiano, servo-croata, polonês, romeno, sueco, tcheco e estoniano.

Muitas vezes, em resumos e manuais, dá-se uma ênfase exagerada na questão do adultério: Capitu traiu Bentinho ou não? Até o lançamento de *The Brazilian Othello of Machado de Assis*, de Helen Caldwell (uma das tradutoras de *Dom Casmurro* para o inglês), a culpa de Capitu era inquestionável. Sob uma perspectiva diferente, Gledson afirma que “o romance é um estudo da obsessão patológica de Bento e não do adultério de Capitu” (ibid., p. 33). Para Antônio Cândido, “dentro do universo machadiano, não importa muito que a convicção de Bento seja falsa ou verdadeira, porque a conseqüência é exatamente a mesma nos dois casos: imaginária ou real, ela destrói sua casa e a sua vida” (Cândido, 1995). Ou seja, não se deve encarar o livro como uma grande charada que Machado propôs a seus leitores e ficou se deliciando com o fato de ela ser indecifrável.

#### 4.4.3 Machado em inglês

As primeiras obras de Machado traduzidas para a língua inglesa foram três contos (*O Enfermeiro*, *Viver e A Cartomante*), publicados na antologia *Brazilian Tales*, de 1921. A sua recepção não foi das mais entusiasmadas, recebendo críticas negativas de algumas revistas e jornais da época. Esse foi um dos principais motivos para que levasse trinta anos até que os romances de Machado chegassem aos leitores anglo-falantes.

Daphne Patai, em seu ensaio sobre a recepção das obras traduzidas de Machado para o inglês, afirma que a literatura brasileira, à exceção de Jorge Amado, ainda é pouco conhecida nos Estados Unidos, fazendo com que tradutores “profissionais” sejam bem raros, traduzindo mais por amor do que por qualquer outra razão (Patai, 1999, p. 87). Mesmo uma escritora do porte de Clarice Lispector só teve sucesso na divulgação de sua obra entre os estudos feministas.

A sugestão de Patai para tornar Machado um *best-seller* em inglês é utilizar uma estratégia de marketing focando um público-alvo bem específico, promovendo Machado como um “escritor de cor preocupado sobretudo com a questão da raça”, ou apresentar suas personagens femininas como “heroínas profeministas”<sup>25</sup> (ibid., p. 88). É claro que se trata de uma brincadeira por parte da ensaísta, apenas para demonstrar como funciona o mercado editorial norte-americano. Outro fato por ela relatado ilustra o pouco prestígio de nosso

---

<sup>25</sup> “a writer of color concerned above all with the issue of race”; “profeminist heroines”.

escritor no exterior. Na cidade onde Patai mora, uma livraria local vendeu vinte e cinco exemplares de *O alquimista*, de Paulo Coelho, em três meses, enquanto que só havia uma cópia de *Dom Casmurro* e nenhum outro livro de Machado em estoque.

Na década de 1950, aparecem três traduções de Machado em inglês: *Memórias póstumas de Brás Cubas*, *Dom Casmurro* e *Quincas Borba*, de três tradutores diferentes, publicadas pela Noonday Press. Até então, oito dos nove romances de Machado já foram publicados no idioma inglês, alguns deles em mais de uma versão, como é o caso de *Dom Casmurro*. Em ordem cronológica, a primeira tradução é de Helen Caldwell, de 1953. Só quarenta anos mais tarde (1992), sairia a segunda versão, de Robert Scott-Bucleuch. A mais recente, de 1997, é a de John Gledson.

A curiosidade fica por conta da segunda versão, de Bucleuch. O tradutor, sem nenhuma justificativa na introdução, omitiu nada menos que oito capítulos inteiros, mais alguns trechos de dois outros capítulos. As razões que levaram o tradutor a esse tipo de procedimento não estão explicitadas na edição consultada – que conta com uma introdução sua, comentando algumas dificuldades da tradução –, mas Patai aponta alguns possíveis motivos para tal. Após relatar que Gregory Rabassa podou vinte por cento do romance *Tocaia Grande*, de Jorge Amado, Patai sugere que as críticas negativas quando da publicação da tradução de Caldwell podem ter influenciado Bucleuch a eliminar os capítulos mais digressivos do romance, uma vez que a maior crítica recaía sobre o fato de que o romance era um tanto truncado, sobretudo em função das digressões. O tradutor, portanto, teria “sacrificado a narrativa pelo enredo, o qual, Scott-Bucleuch parece pensar, deve progredir com o mínimo de impedimentos possível”<sup>26</sup> (ibid, p. 99).

---

<sup>26</sup>“(…) to sacrifice the narrative to the plot, which, Scott-Bucleuch seems to think, must proceed with as few impediments as possible”.



## 5. MÉTODO

Para facilitar a busca dos itens lexicalmente criativos, optou-se por uma abordagem baseada em corpus. Corpus, neste trabalho, está entendido como “qualquer coleção de textos naturais (em oposição a exemplos/sentenças), em formato eletrônico e passíveis de análise automática ou semi-automática (ao invés de manual)” (BAKER, 1995, p. 226)<sup>27</sup>.

### 5.1 Digitalização dos textos

Para que pudéssemos investigar os textos, era preciso tê-los em formato eletrônico. O original em português foi extraído integralmente do *site* do NUPILL (Núcleo de Pesquisas em Informática, Literatura e Linguística – CCE/UFSC), o qual disponibiliza diversos textos da literatura brasileira que já caíram em domínio público.

Quanto às traduções, foi necessário digitalizá-las, em um processo que teve as seguintes etapas:

- 1) escaneamento do texto – para isso, utilizou-se um *scanner* da marca HP Scanjet 3200c e um *software* de reconhecimento de caracteres (Omnipage Pro 12);
- 2) revisão do texto escaneado – apesar de eficiente, o *software* Omnipage não é livre de erros. Qualquer falha tipográfica no original pode ocasionar uma leitura equivocada por parte do aplicativo, exigindo, portanto, uma revisão do texto já em formato eletrônico;
- 3) etiquetagem dos textos – processo que facilita consultas posteriores. Consiste em inserir as seguintes etiquetas (*tags*): título dos capítulos (exemplo: <title>One Secret for Another</title>), quebra de parágrafo (texto do parágrafo</p>) e número de página (<pb n=21>, para designar a página 21);
- 4) análise dos textos em formato eletrônico – para realizar tal análise, utilizou-se o *software* WordSmith Tools 3.0, desenvolvido por Mike Scott, da University of Liverpool. Esse aplicativo permite gerar listas de palavras (em ordem

---

<sup>27</sup> “any collection of running texts (as opposed to examples/sentences), held in electronic form and analysable automatically or semi-automatically (rather than manually)”.

alfabética ou por frequência), concordância (palavra ou frase no contexto) e palavras-chave (usando outro corpus como referência).

## 5.2 Corpus paralelo

A decisão pela utilização de um corpus paralelo composto de textos exclusivamente literários para trabalhar com questão da criatividade lexical foi tomada seguindo os mesmos critérios de Kenny (2001). A autora argumenta que “[os textos literários] podem conter mais ocorrências de criatividade lexical do que outros textos”<sup>28</sup>, uma vez que eles “liberam todo o potencial criativo da língua, enquanto que outras categorias textuais se baseiam num potencial lingüístico reduzido” (*ibid.*, p. 112)<sup>29</sup>. Mais adiante, a seguinte afirmação reforça a visão da autora: “Se pretendemos (...) revelar evidências de normalização, a tradução de prosa narrativa parece um bom ponto de partida” (*ibid.*, p. 113)<sup>30</sup>.

Considerando o corpus que inclui o romance *Dom Casmurro* em português e três traduções para a língua inglesa, podemos afirmar se tratar de um corpus paralelo, de acordo com a definição de Baker (1995): “um corpus paralelo consiste de textos originais na língua-fonte A e suas versões traduzidas para a língua B”<sup>31</sup>. Tomando por base a classificação proposta por Fernandes (2006, p. 98), o corpus de pesquisa pode ser assim classificado:

- (a) **Número de línguas** – pode-se classificar um corpus paralelo em *bilíngüe*, *trilíngüe* ou *multilíngüe*, quando for composto de mais de três idiomas. O presente estudo utiliza um corpus paralelo bilíngüe: português brasileiro (língua-fonte) e inglês (língua-alvo);
- (b) **Restrição temporal** – há duas maneiras de classificar um corpus paralelo quanto à restrição de períodos de tempo: *sincrônico* (para um período particular no tempo) e *diacrônico* (quando o foco está no desenvolvimento

---

<sup>28</sup> “[literary texts] may contain more instances of lexical creativity than other texts”.

<sup>29</sup> “give full rein to the creative potential of language while other text categories draw on a reduced linguistic potential”.

<sup>30</sup> “If we wish (...) to uncover evidence of normalization, the translation of narrative prose would seem a good place to start”.

<sup>31</sup> “a parallel corpus consists of original, source language-texts in language A and their translated versions in language B”.

histórico do objeto de estudo). Uma vez que as traduções que compõem o corpus foram publicadas em diferentes períodos (1953, 1992 e 1997), pode-se classificar o corpus como diacrônico;

- (c) **Domínio do corpus** – neste aspecto, os corpora podem ser classificados como *gerais* e *especializados*. Um exemplo de corpus *geral* é o BNC, que inclui textos das mais diversas áreas de conhecimento, e serve para o estudo de traduções como um todo. Por outro lado, o corpus *especializado* lida com gêneros ou tipos de texto específicos. Por se tratar de apenas um romance e suas traduções, temos um corpus pertencente ao domínio especializado;
- (d) **Direcionalidade** – entende-se por direcionalidade (*directionality*) a direção de tradução dos textos que compõem um determinado corpus (Zanettin, 2000). Há três classificações possíveis: *unidirecional*, *bidirecional* e *multidirecional*. Se o corpus inclui textos da L1 e sua(s) tradução(ões) para a L2, diz-se que este corpus é unidirecional, ou seja, a direção do texto traduzido é uma só. Se o corpus contar com textos da L1, suas traduções para L2, além de textos na L2 traduzidos na L1, trata-se de um corpus bidirecional. Um exemplo bastante citado na literatura é o English-Norwegian Parallel Corpus, que consiste de textos originais em inglês e suas traduções para o norueguês, e textos originais em norueguês e suas traduções para o inglês (Johansson, 1998, pp. 7-8). Finalmente, corpus multidirecional é aquele que envolve mais de duas línguas e não está centrado na L1. O corpus de estudo pode ser classificado como unidirecional, ou seja, português brasileiro traduzido para o inglês.

Em suma, estamos trabalhando com um corpus paralelo bilíngüe, diacrônico, especializado e unidirecional. A Tabela 2 abaixo traz mais informações a respeito das dimensões do corpus, em número de palavras.

Traduções		Original
Caldwell	77.531	66.537
Buccleuch	70.441	
Gledson	76.928	
Total	224.900	

**Tabela 2.** Dimensões do corpus de estudo.

O programa WordSmith permite definir o hífen como quebra de palavras ou não. Esse ajuste gera um problema em relação às ênclises e mesóclises (*chamam-me, vi-lhe, dir-me-ia*) e aos substantivos compostos (*campos-santos, bem-amados*) na obra original. Se a opção “hyphen breaks words” estiver selecionada, resolve-se a questão das ênclises e mesóclises, mas os substantivos compostos serão computados como duas palavras cada. Se, por outro lado, a opção estiver desmarcada, *chamam-me* representará uma palavra e *me chamam*, duas palavras, o que é uma inconsistência lingüística. Por esse motivo, a opção “hyphen breaks words” foi desativada nos ajustes do programa, e todos os pronomes precedidos de hífen foram calculados separadamente. O resultado final é um total de 66.537 palavras: 65.349 do cálculo inicial somadas a 1.188 casos de ênclises e mesóclises.

### 5.3 Corpora de referência

Para averiguar se determinada colocação é marcada, pode-se utilizar o “instinto” de falante nativo de um idioma, comparando-a com as colocações com as quais se está familiarizado. No entanto, não convém utilizar esse procedimento em uma pesquisa acadêmica, pois impossibilita a replicabilidade do estudo por parte de outros pesquisadores. Além disso, como bem ressaltou Kenny (2001), “usos criativos da língua têm maior destaque quando considerados contra um pano de fundo daquilo que é típico de uma língua” (p. 32)<sup>32</sup>. Por essas razões, optamos pelo uso de dois corpora de referência: um para o português brasileiro, outro para a língua inglesa. Por corpus de referência, entende-se, aqui, um corpus utilizado para propósitos comparativos, previamente desenvolvido para fornecer informações abrangentes sobre determinada língua. Os corpora de referência visam representar o maior número possível de variedades da língua, para que possam, assim, ser utilizados como fonte de referência para gramáticas, dicionários, thesauri, além de pesquisas como a proposta neste projeto.

#### 5.3.1 Corpus em Língua Portuguesa

Para referência, usaremos o corpus de português brasileiro compilado pelo NILC (Núcleo Interinstitucional de Linguística Computacional). Conforme divulgado pelos próprios pesquisadores, o corpus é composto de aproximadamente 35 milhões de palavras, consistindo de textos em prosa, os quais foram divididos em *textos corrigidos*, *textos não corrigidos* e *textos semicorrigidos*.

Entre os textos denominados *corrigidos* (32.590.000 palavras), estão aqueles publicados para grande número de leitores e que supostamente passaram por uma correção por especialistas em revisão de textos. É composto de textos de diversas áreas (literatura, química, biologia, arte, entre outros), oriundos de livros, revistas, jornais, constituição brasileira e textos jurídicos.

*Textos não corrigidos* (738.000 palavras) são aqueles escritos por pessoas de nível médio de escolaridade (2º grau) e universitários. Alguns exemplos incluem redações, monografias e textos de publicidade. São classificados como não corrigidos porque, por não terem sido publicados em grande escala, é muito provável que não tenham passado por uma revisão profissional.

Finalmente, temos os textos semicorrigidos (1.150.000 palavras), que consistem de textos publicados para um pequeno número de leitores, ou não publicados, que até podem ser corrigidos, mas não por especialistas da área de revisão de textos. Contratos, relatórios, dissertações acadêmicas, entre outros, incluem-se nesta classificação.

Uma das limitações do corpus do NILC, especificamente para a pesquisa a que se propõe este estudo, é a reduzida quantidade de textos literários. De fato, uma pesquisa utilizando a função WordList do programa WordSmith aponta um total de pouco mais de 2 milhões de palavras referentes à parte literária (representada por antologias e textos da literatura brasileira), o que representa menos de 10% do total de palavras do corpus.

Para sanar esse problema, fez-se necessário aumentar consideravelmente o volume de obras literárias, a fim de dar maior credibilidade a um estudo que pretende justamente avaliar itens lexicalmente criativos em um clássico da literatura brasileira, bem como a forma como eles foram lidados pelos tradutores nas respectivas traduções.

Na Internet, é possível encontrar obras que já caíram em domínio público, digitalizadas por voluntários e/ou grupos de pesquisa. Dois *sites* serviram de base para a

---

<sup>32</sup> “creative uses of language are brought into greater relief when one sees them against a backdrop of what is typical for a language”.

extração dos textos: o NUPILL (v. seção 5.1) e o *Virtual Books Online*, hospedado no portal Terra. Além de clássicos da literatura brasileira, também optamos por incluir obras de autores portugueses. Da mesma forma que o corpus de referência em língua inglesa contém autores ingleses, irlandeses e estadunidenses, não vimos impedimento para incluir obras da literatura luso-brasileira em um corpus em língua portuguesa. Entretanto, foram incluídos apenas contos e romances, excluindo poesia e teatro, visto que o corpus de estudo é composto de um romance e suas traduções. Aproximadamente um quarto do total de palavras é referente a obras do próprio Machado, o que possibilitará verificar se determinado item considerado criativo é recorrente na escrita machadiana. A lista completa dos autores e obras incluídos neste subcorpus encontra-se no Anexo 2.

O resultado final deste corpus de referência consiste de mais de 42 milhões de palavras, sendo que cerca de 20% delas representam obras literárias. Por meio desse artifício, acreditamos tornar o trabalho mais confiável em relação aos resultados obtidos. Para evitar confusão, este novo corpus será referido posteriormente como CRLP (Corpus de Referência em Língua Portuguesa).

### 5.3.2 Corpus em Língua Inglesa

De acordo com informações divulgadas no próprio *site*, o British National Corpus (BNC) possui 100 milhões de palavras, incluindo linguagem escrita e falada de fontes diversas, projetado para ser uma amostra representativa do inglês britânico da última parte do século XX, tanto escrito quanto falado.

A parte escrita do BNC – que representa 90% do corpus – inclui trechos de jornais, periódicos especializados e revistas para todas as idades e interesses, livros acadêmicos e ficção popular, cartas publicadas e não-publicadas, ensaios colegiais e universitários, entre outros.

Os 10% restantes provêm de fontes de língua falada. Aqui estão incluídas conversações informais gravadas de voluntários selecionados de diferentes idades, regiões e classes sociais, além de outras fontes de coleta de dados, como reuniões governamentais e de negócios e programas de rádio.

A maior proporção de textos escritos é favorável a esta pesquisa, uma vez que o principal corpus de estudo é composto de textos literários. Por outro lado, similarmente ao corpus do NILC, há uma deficiência de obras da literatura em língua inglesa. Para corrigir essa falha, utilizamos o mesmo procedimento em relação ao corpus de referência em língua portuguesa: acrescentamos um volume significativo de contos e romances, moldando o corpus para que se tornasse mais adequado à presente pesquisa. Os textos foram extraídos do *site* denominado Project Gutenberg ([www.gutenberg.org](http://www.gutenberg.org)), que hospeda diversos livros eletrônicos que não são mais protegidos por direitos autorais. O Anexo 3 traz a relação completa dos autores e obras incluídos neste subcorpus.

Aproximadamente 11 milhões de palavras foram adicionadas ao já extenso corpus do BNC, ou seja, o corpus de referência sofreu um acréscimo de textos literários de mais de 10%, o que servirá para aumentar a credibilidade da análise de itens lexicalmente criativos provenientes das traduções que compõem o corpus de estudo. Este novo corpus será denominado, em referências posteriores, de CRLI (Corpus de Referência em Língua Inglesa).

## 6. ANÁLISE DOS DADOS

### 6.1 Escolha do nóculo

Antes de analisar a ocorrência de colocações marcadas no corpus paralelo, faz-se necessário selecionar um nóculo para investigação. Kenny (2001) sugere duas abordagens, uma das quais será adotada nesta pesquisa. A primeira delas compreende

a seleção de um nóculo que seja relativamente comum no corpus sob investigação (...) e, a seguir, a análise de uma concordância deste nóculo gerada automaticamente para verificar se contém alguma colocação criativa (p. 129)<sup>33</sup>.

A segunda abordagem proposta pela autora refere-se às colocações usadas por um mesmo autor, as quais poderiam ser consideradas criativas. Esta última só faria sentido se o corpus paralelo em questão contivesse duas ou mais obras. Por se tratar de um corpus com enfoque em uma única obra, adotaremos, naturalmente, a primeira abordagem.

A título de exemplificação, é válido mencionar os critérios utilizados por Kenny (2001) e Bueno (2005) na escolha de seus respectivos nósculos para análise de criatividade lexical em colocações marcadas. Esta seleciona os nósculos “brincar” e “rir” na obra *Macunaíma*, de Mário de Andrade, “por estarem, ambos, na lista gerada pela ferramenta *KeyWords* (o que aponta para sua representatividade no corpus); por serem dois processos; e por estarem semanticamente relacionados em sua concepção convencional” (BUENO, 2005, p. 83). Por outro lado, Kenny prefere apenas um nóculo (o lema *Auge*, “olho”) e aponta diversas razões para sua escolha. Entre elas, (i) por ocorrer com frequência regular em todos os textos do GEPCOLT, (ii) em função de observações interessantes feitas anteriormente por outros lingüistas e (iii) pelo fato de que o nóculo “olho” tem participação em diversas frases fixas ou semi-fixas (ibid., p.135-136).

Vale ressaltar, ainda, que a escolha do nóculo conta, em grande parte, com a intuição do pesquisador. É a própria Kenny (ibid., p. 135) quem comenta sobre o fato: “(...)

---

<sup>33</sup> “selecting a node that is relatively common in the corpus under investigation (...) and then examining an automatically generated concordance of this node to see if it contains any creative collocations”.



há um inevitável papel da intuição (do próprio pesquisador ou de outros linguistas), e até mesmo da sorte, na seleção de nódulos que mereçam análise na lingüística de corpus”<sup>34</sup>.

## 6.2 O nódulo “olhos”

Coincidentemente, um dos nódulos escolhidos para investigação de criatividade lexical neste estudo foi o mesmo selecionado por Kenny (2001). Os motivos que me levaram a essa opção, entretanto, divergem dos da autora supracitada, conforme explicitado abaixo.

A primeira questão a ser elucidada refere-se à representatividade do nódulo no corpus. Para isso, é preciso utilizar a ferramenta KeyWords, do programa WordSmith, a qual “permite uma seleção de itens de uma lista de palavras (ou mais) por meio da comparação de suas frequências com uma lista de referências” (Sardinha, 2004, p. 96). Uma vez que o objetivo é analisar as traduções *para* o inglês, o que interessa à pesquisa são as palavras-chave da obra original. A Tabela 3 traz uma lista das primeiras vinte palavras-chave do romance *Dom Casmurro* em português (corpus de estudo), usando o corpus do NILC como referência.

N	WORD	FREQ.	DOM.TXT %	FREQ.	NILC.LST %	KEYNESS
1	ME	1.035	1,55	31.047	0,08	4.047,8
2	CAPITU	337	0,50	381	< 0,001	3.264,4
3	LHE	483	0,72	20.711	0,06	1.567,4
4	NÃO	1.526	2,28	291.668	0,79	1.245,7
5	ERA	553	0,83	42.277	0,11	1.228,1
6	EU	533	0,80	40.145	0,11	1.197,2
7	MINHA	342	0,51	13.304	0,04	1.170,1
8	MÃE	229	0,34	6.304	0,02	929,7
9	QUE	2.687	4,02	782.774	2,13	909,7
10	CAPÍTULO	186	0,28	4.470	0,01	802,3
11	ESCOBAR	110	0,16	568	< 0,001	789,6
12	COUSA	98	0,15	696	< 0,001	646,3
13	MAS	605	0,90	97.240	0,26	632,9
14	MIM	162	0,24	5.198	0,01	611,4
15	PADRE	118	0,18	1.971	< 0,001	589,5
16	OLHOS	164	0,25	5.843	0,02	587,2

<sup>34</sup> “(...) there is an inevitable role for intuition (one’s own, or that of other linguists), and even good fortune, in the selection of nodes worth analyzing in corpus linguistics”.

17	JUSTINA	55	0,08	69	< 0,001	524,3
18	NEM	243	0,36	19.604	0,05	516,6
19	BENTINHO	56	0,08	149	< 0,001	467,1
20	SEMINÁRIO	85	0,13	1.251	< 0,001	444,8

**Tabela 3.** Lista de palavras-chave de *Dom Casmurro* no original.

Para uma melhor análise dos dados acima, é preciso compreender o significado de cada coluna:

1. posição da palavra-chave, por ordem crescente;
2. palavra-chave;
3. frequência no corpus de estudo;
4. porcentagem em relação ao total do corpus de estudo;
5. frequência da palavra-chave no corpus de referência;
6. porcentagem em relação ao total do corpus de referência. Os casos em que há espaços em branco representam valores menores do que 0,001;
7. chavicidade<sup>35</sup>, que traz uma estatística de comparação calculada com base no teste do qui-quadrado ou *log-likelihood*.

Como pode ser observado, o nóculo *olhos* encontra-se na 16ª posição, com 164 ocorrências no corpus de estudo. Esse nóculo representa 0,25% do total do corpus de estudo, como indicado pela quarta coluna. O valor de p menor do que 0,000001 significa que existe a probabilidade de uma em 1 milhão de tal resultado ser fruto de uma mera coincidência. A ocorrência do nóculo *olhos* entre as palavras-chave do corpus de estudo é, portanto, o primeiro motivo que justifica sua escolha para posterior análise de concordância.

A segunda razão está diretamente relacionada às descrições de Capitu, uma das personagens femininas mais famosas da literatura brasileira. Segundo Bentinho, narrador do romance, Capitu possuía “olhos de ressaca”, expressão que, pela complexidade que apresenta, já merece um estudo mais aprofundado; além disso, José Dias, o agregado da família Santiago, descreve Capitu como tendo “olhos de cigana oblíqua e dissimulada”. Esta última constatação, inclusive, serve para aumentar o ciúme e a desconfiança de

<sup>35</sup> Esta é a tradução utilizada por Sardinha (2004) para o termo *keyness*.

Bentinho em relação à sua esposa. Em suma, o papel de relevância que o nódulo *olhos* parece desempenhar em *Dom Casmurro* é mais um motivo para justificar sua seleção para análise.

O próximo passo é realizar uma concordância utilizando como palavra de busca o nódulo *olhos*. A relação completa de todas as ocorrências (164) encontra-se no Anexo 1. Cabe ao pesquisador interpretar os resultados obtidos por meio do auxílio do software, procurando padrões recorrentes ou, no caso da presente pesquisa, colocações potencialmente marcadas. Na suspeita de uso de criatividade lexical, utiliza-se o corpus de referência (NILC) para confirmação, verificando se determinada colocação pode ser considerada criativa, ou se é simplesmente um caso de uso habitual na língua. Se a suspeita for confirmada, passa-se, a seguir, a uma análise das opções feitas pelos três tradutores, fazendo uma consulta no corpus de referência da língua inglesa (BNC), a fim de determinar se houve ou não a ocorrência de normalização na tradução das colocações marcadas.

Após uma primeira análise das ocorrências, as colocações mais habituais foram eliminadas, restando as seguintes concordâncias para que uma investigação mais aprofundada fosse realizada:

1. sida às saias de minha mãe, não atendia aos **olhos** ansiosos que eu lhe mandava; também
2. ra quadragenária, magra e pálida, boca fina e **olhos** curiosos. Vivia conosco por favor de mi
3. ade do estilo, o que eles foram e me fizeram. **Olhos** de ressaca? Vá, de ressaca. É o que
4. o e comborço Escobar. Cheiram também aos **olhos** de ressaca de Capitu. Assim, posto se
5. seria o discurso. CAPÍTULO CXXIII / **OLHOS DE RESSACA** Enfim, chegou
6. m adágio dacolá... CAPÍTULO XXXII: **OLHOS DE RESSACA** Tudo era mate
7. eu coração? Talvez porque nenhuma tinha os **olhos** de ressaca, nem os de cigana oblíqua e
8. epente, cessando a reflexão, fitou em mim os **olhos** de ressaca, e perguntou-me se tinha m
9. mitológico. Ainda há pouco, falando dos seus **olhos** de ressaca, cheguei a escrever Tétis; ri
10. ais me envolviam no seu mistério, sem que os **olhos** de ressaca de Capitu deixassem de crê
11. e compridas eram agora breves e alegres; os **olhos** desaprenderam de chorar, se porventur
12. gordo e pesado, tinha a respiração curta e os **olhos** dorminhocos. Uma das minhas recorda
13. is de algum tempo de silêncio, recolhendo os **olhos** e um suspiro! --Tem agradado muito est
14. randura, levantava-me o queixo e espetava os **olhos** em mim, ansioso também, como a prim
15. sto, a reflexão; íamos dar com ele, muita vez, **olhos** enfiados em si, cogitando. Respondia-n
16. outras andorinhas estavam trepadas no ar, os **olhos** enfiados nos olhos, mas tão cautelosos
17. ; achou até natural e espetou-me outra vez os **olhos**. Então contei-lhe por alto o que podia,

18. aram cair de cansadas ou de esquecidas. Os **olhos** fitavam-se e desfitavam-se, e depois de  
19. xto para mirá-los mais de perto, com os meus **olhos** longos, constantes, enfiados neles, e a  
20. pitu, a princípio, não disse nada. Recolheu os **olhos**, meteu-os em si e deixou-se estar com  
21. E, após alguma reflexão, fitando em mim uns **olhos** murchos e teimosos, perguntou-me: --C  
22. mbrado a definição que José Dias dera deles, "**olhos** de cigana oblíqua e dissimulada." Eu n  
23. e depressa e não achei compostura; metia os **olhos** pelas cadeiras. Ao contrário, Capitu erg  
24. ter, e não vi que essa menina travessa e já de **olhos** pensativos era a flor caprichosa de um f  
25. o Escobar era um tanto metediço e tinha uns **olhos** policiais a que não escapava nada. --Sã  
26. depois, como se faz a um filho de verdade; os **olhos** que ele me deitou foram ternos e agrad  
27. quecer inteiramente a mão de Sancha nem os **olhos** que trocamos. Agora achava-lhes isto,  
28. razão particular. Tudo isso me empanava os **olhos**... Vieram os jornais do dia: davam notíc

**Tabela 4.** Linhas selecionadas da concordância do nóculo *olhos*.

### 6.3 Análise das linhas selecionadas

Linha 1: “Capitu, cosida às saias de minha mãe, não atendia aos **olhos ansiosos** que eu lhe mandava.”

#### Traduções:

- Capitú, hugging my mother's skirts, paid no heed to the **anxious looks** I sent her. (Caldwell)
- Capitu, who remained glued to my mother's side, paid no attention to the **anxious looks** I cast at her. (Bucclench)
- Capitu, who stayed close to my mother, paid no attention to the **anxious looks** I directed at her (Gledson)

No CRLP, há cinco ocorrências para *olhos + ansiosos*<sup>36</sup>: no próprio *Dom Casmurro*; em *Inocência*, de Visconde de Taunay; em *A mortalha de Alzira*, de Aluísio de Azevedo; em *A alma encantadora das ruas*, de João do Rio; e em uma tradução de *O*

<sup>36</sup> O nóculo principal na busca foi *olhos*, com um horizonte de pesquisa para *ansiosos* estabelecido em cinco itens à esquerda e cinco à direita.

*chamado da floresta*, de Jack London, ou seja, apenas três ocorrências – à exceção de *Dom Casmurro* – em textos originalmente escritos em português.

Uma busca pelos colocados do nóculo *olhos* no corpus de referência revela que as colocações mais usuais são:

- **cores** – azuis, negros, verdes, pretos;
- **descrição física** – grandes, pequenos, miúdos;
- **estado atual** – abertos, fechados, arregalados, desvairados.

Não é comum, portanto, atribuir estados de espírito a esse nóculo. Por outro lado, o nóculo *olhar* apresenta comportamento distinto. Uma pesquisa por seus colocados mais frequentes demonstra que a sua associação com adjetivos referentes a estados de espírito não é incomum. Alguns exemplos são: *agitado*, *colérico*, *comovido*, *desanimado*, *desvairado*, *inquieta*, etc. Ainda no CRLP, há uma ocorrência para “olhar ansioso” (curiosamente, em outro romance de Machado, *Helena*) e outra para “olhares ansiosos”. Pode-se concluir, a partir da análise qualitativa desses dados, que Machado fez uso de uma colocação incomum, ao associar “olhos” com “ansiosos”, quando o esperado seria “olhares”.

Para fins de comparação, pesquisamos, no CRLI, ocorrências para “anxious looks” e para “anxious eyes”. A primeira colocação foi encontrada em apenas uma ocorrência, em uma obra sobre poesia feminina do século XVIII:

With anxious Looks my Countenance is clad,

O trecho é, muito provavelmente, a citação de um verso de alguma poeta do século sendo estudado (em função dos substantivos em maiúsculas). Para a segunda expressão, os resultados são diversos: “anxious eyes” apresenta 18 ocorrências no CRLI:

1. Leicester’s tacklers, unwisely, were keeping **anxious eyes** on Carling and
2. She drove slowly and carefully, her **anxious eyes** peering ahead...
3. ...watching her new friend with sorrowful **anxious eyes**.
4. She fixed him with **anxious eyes**.

5. ... the **anxious eyes** of a child peered at her from a white mask.
6. Before their **anxious eyes**, these innocent girls see the much...
7. Bradley's dark and **anxious eyes** narrowed with suspicion.
8. ...scanning the tables with **anxious eyes** and making agitated...
9. Alighting hesitantly, she cast **anxious eyes** around her, wondering...
10. I saw Benjamin's **anxious eyes** but ignored his warning look.
11. ...Chuck met the **anxious eyes** of the others grouped around him.
12. ...occasionally looking at her man with those **anxious eyes**, to which...
13. opened it to discover her standing without, with **anxious eyes**...
14. But, though I looked all round with **anxious eyes**, I could see...
15. His mother's **anxious eyes** were fixed on his face. Even now he...
16. ...the first time something in her **anxious eyes**, in the way she...
17. Her **anxious eyes**, her charming lips, her slip of a figure, were...
18. ... voice that trembled. Her **eyes** were **anxious**, agonized. The...

Neste exemplo, não se confirma a hipótese de normalização, uma vez que a colocação “anxious eyes”, evitada pelos tradutores, mostrou ser mais usual do que “anxious looks”. Conscientemente ou não, todos os tradutores utilizaram a colocação menos comum, da mesma forma que o autor do original.

A mesma linha 1 permite um outro tipo de análise, desta vez investigando o verbo colocado com o nódulo *olhos*. A concordância é repetida abaixo, com os itens em negrito indicando a colocação sendo investigada.

“Capitu, cosida às saias de minha mãe, não atendia aos **olhos** ansiosos que eu lhe **mandava**.”

Traduções:

- Capítú, hugging my mother's skirts, paid no heed to the anxious **looks I sent** her. (Caldwell)
- Capitu, who remained glued to my mother's side, paid no attention to the anxious **looks I cast** at her. (Bucleuch)
- Capitu, who stayed close to my mother, paid no attention to the anxious **looks I directed** at her (Gledson)

As três opções diferentes nas traduções sugerem que a tradução do verbo não é tão simples como poderia, em princípio, parecer.

Uma consulta no CRLP não traz nenhuma ocorrência para qualquer forma do verbo *mandar* + *olhos*. Por outro lado, o exemplo abaixo, extraído de uma busca pela associação de “mandar” e suas variações com o nódulo “olhar”, é a indicação de uma colocação mais comum:

Não viu, não adivinhou sequer que Pedro naturalmente pararia um instante, para voltar a cara e **mandar** um derradeiro **olhar** à moça enterrada.

Quanto aos verbos que colocam com “um olhar”, incluindo o objeto da ação (por exemplo, “lançou-me um olhar”, “fitou nele um olhar”), encontramos *lançar*, *dardejar*, *derramar*, *despedir*, *embeber*, *fitar*, *atirar*, *cravar*, *deitar*, *correr*, *fixar*, *rolar* e *volver*.

Daí podemos concluir que “mandar um olhar” é a colocação usual, enquanto que “mandar olhos” é um exemplo de colocação criativa de Machado.

Quanto às escolhas feitas pelos tradutores, é preciso analisar cada caso individualmente. Buccleuch utiliza o verbo “to cast” para acompanhar “looks”. Vejamos as ocorrências para essa colocação no CRLI:

1. ...I **cast** envious **looks** at yet another Great Southern Hotel, and dragged...
2. ...Lachlan **cast** glowering **looks** at both Elizabeth and Marion Aluinn.
3. ...they had **cast** us unrestrained **looks** which said: “At last, I’m all yours...”
4. Several times Floy **cast** sidelong **looks** at Balor, trying to decide how...
5. ...the busts of Roman emperors **cast** classically guarded **looks** at his...
6. ...while young Lady C **cast** meaningful **looks** at sturdy gamekeeper...
7. ... the odd **looks** they had **cast** at one another were repeated several...

Os exemplos indicam tratar-se de uma colocação adequada à estrutura da língua fonte. O comportamento desse verbo, em todas as ocorrências, pode ser expresso da seguinte forma: *to cast* + adjetivo + objeto. Embora inverta a posição do verbo (como no exemplo 7), o tradutor optou por uma colocação usual em língua inglesa.

Caldwell, por sua vez, prefere o verbo *to send*. No CRLI, não há ocorrências de sua associação com o substantivo *looks*. O que parece ter acontecido foi uma tentativa de reproduzir com máxima fidelidade o verbo usado no original, numa tradução literal de “mandava” por “sent”, corroborando a afirmação de Patai de que “[Caldwell] reconheceu que seu trabalho era uma tradução de um texto brasileiro, e manteve traços daquele texto (...) quando não havia alternativas aceitáveis em inglês” (1999, p. 91)<sup>37</sup>. Finalmente, Gledson opta por um terceiro verbo: *to direct*. Uma análise das ocorrências no CRLI revela uma única ocorrência:

1. ...and many sobre **looks** were **directed** at Doctor Agrippa, Ruthven...

Ou seja, ainda que não tenha apresentado um número elevado de ocorrências, não se pode afirmar que a colocação “direct looks at” seja uma criação do tradutor, comprovando a hipótese de normalização.

Linha 2: “Era quadragenária, magra e pálida, boca fina e **olhos curiosos**.”

Traduções:

- She was a woman of forty, thin, pale, with a haughty mouth and **inquisitive eyes**. (Caldwell)
- She was in her forties, pale and thin with a small mouth and **inquisitive eyes**. (Buckleuch)
- She was in her forties, lean and pale, with a thin mouth and **inquisitive eyes**. (Gledson)

A colocação destacada no exemplo não parece, após análise do corpus de referência, ser marcada, uma vez que há 18 ocorrências em diversas obras literárias (por exemplo, em Adolfo Caminha, José de Alencar e Lima Barreto). Além disso, essa colocação foi encontrada em textos jornalísticos, mais especificamente em um artigo no caderno de

---

<sup>37</sup> “[Caldwell] acknowledged that her work was a translation of a Brazilian text, and she retained traces of that text (...) when no acceptable alternatives in English presented themselves”.



turismo de um jornal brasileiro, o que serve testemunhar a favor de seu uso menos criativo que rotineiro.

A linha 2, portanto, está descartada como exemplo de colocação marcada.

Linha 3: “Não me acode imagem capaz de dizer, sem quebra da dignidade do estilo, o que eles foram e me fizeram. **Olhos de ressaca**? Vá, de ressaca.”

Traduções:

- I can find no image – without breaking the dignity of my style – to convey what they were and what they did to me. **Eyes like the tide**? Yes, like the tide. (Caldwell)
- I can think of no comparison which, without lowering the dignity of the style, can convey what they were and the effect they produced on me. **Whirlpool eyes**? Very well then, whirlpool eyes be it. (Bucleuch)
- No image comes to mind that doesn't offend against the rules of good style, to say what they were and what they did to me. **Undertow eyes**? Why not? Undertow. (Gledson)

Neste caso, a seguinte análise engloba as linhas 3 a 10, todas contendo a mesma, e muito provavelmente a mais criativa, colocação no corpus de estudo (“olhos de ressaca”). Uma vez que todos os tradutores mantêm a mesma escolha ao longo do texto, um único exemplo basta para os fins desta análise.

A comprovação da originalidade desta colocação está no fato de que todas as ocorrências de “olhos de ressaca”, quando não são do próprio *Dom Casmurro*, fazem referência a ele (em críticas literárias, por exemplo).

Talvez a maior dificuldade de tradução recaia sobre o sentido da palavra “ressaca”. Segundo definição de Houaiss, o termo refere-se ao “forte movimento das ondas sobre si mesmas, resultante de mar muito agitado, quando se chocam contra obstáculos no litoral.” Aquele que olhar nos olhos da personagem poderia se sentir sugado, assim como a corrente submarina faz com os banhistas.

A primeira tradutora, Caldwell, optou por comparar os olhos de Capitu à maré, fenômeno que está mais relacionado à elevação e abaixamento das águas do mar. Visto que o objetivo deste estudo não é avaliar a qualidade – ou adequação – da tradução, é preciso ater-se à evidência ou não de normalização. No CRLI, não há uma única ocorrência de “eyes like the tide”. Por outro lado, em uma pesquisa secundária realizada no site de busca Google<sup>38</sup>, encontramos uma ocorrência dessa colocação na canção *The man in the moon*, de Adrian Belew. Não se pode afirmar que a opção de Caldwell foi equivalente à do original, considerando que a construção *eyes + like + (adjetivo) substantivo* é muito freqüente no corpus de referência. Alguns exemplos são “eyes like furnaces”, “eyes like pale sapphires”, “eyes like red fire”, entre outros. Trata-se, neste caso, de normalização no texto traduzido.

Buckleuch opta pelo termo *whirlpool*, que pode ser traduzido por remoinho, um movimento de rotação em espiral que passa a idéia de algo que engole o que lhe passa por perto, como os olhos de Capitu. A colocação não é usual na língua inglesa, embora haja algumas poucas ocorrências. Não se pode considerar este um caso de normalização; a questão principal aqui seria avaliar qual a escolha mais adequada para traduzir o termo “ressaca”: *tide*, *whirlpool* ou *undertow*, sendo que o último foi utilizado por Gledson, a ser analisado a seguir.

*Undertow* parece ser o termo que mais se aproxima do original. Sua definição, “the seaward pull of receding waves after they break on a shore”, de acordo com o American Heritage Dictionary of the English Language, assemelha-se à definição de “ressaca” dada por Houaiss. As escassas ocorrências desta colocação aparecem em artigos acadêmicos sobre o romance em questão, o que torna evidente o fato de que os pesquisadores utilizaram a versão de Gledson para consulta, talvez por ser a mais recente.

Entre as três escolhas, pode-se afirmar que Buckleuch e Gledson fizeram uso de colocações criativas, ao passo que Caldwell incorreu em normalização.

Linha 11: “(...) os **olhos desaprenderam** de chorar, se porventura choravam antes.”

---

<sup>38</sup> Apesar de haver restrições ao uso da Internet como um todo para a pesquisa acadêmica, principalmente em função da falta de confiabilidade dos resultados, eventualmente este recurso será utilizado para adicionar dados relevantes à análise.

Traduções:

- (...) his **eyes forgot** how to cry if indeed they had cried before. (Caldwell)
- (...) and his **tears**, if he shed any before, **no longer flowed**. (Bucclench)
- (...) his **eyes forgot** how to cry, if indeed they wept before. (Gledson)

Para a combinação *olhos + desaprender*, há uma única ocorrência no CRLP, além do próprio Machado. Trata-se da obra *O arco de Sant'ana*, do escritor português Almeida Garrett. Colocações mais comuns são: “olhos vermelhos de chorar”, “olhos inchados de chorar” e “olhos cansados de chorar”. Por esse motivo, será considerada colocação marcada.

Caldwell e Gledson fazem a mesma escolha. Apesar da existência da palavra *unlearn* no idioma inglês, ambos preferiram evitar o estranhamento que a colocação *his eyes unlearned to cry* poderia causar em anglófonos, optando pelo mais usual *forget + how + to + verbo*. Essa estrutura está presente em dezenas de ocorrências no CRLI, conforme ilustrado pelos exemplos abaixo:

1. If you don't read, then eventually you will **forget how to** read.
2. I know how long it is, how deep, and one does not **forget how to** swim.
3. ... that for the moment she quite **forgot how to speak** good English...
4. I must be precise here: the woman **forgot how to hate** him...
5. Eventually she'll **forget how to feed** herself, how to go toilet bedridden...

Bucclench, por sua vez, evita o uso do substantivo *eyes*, alternando para *tears*, mas sem criar nenhuma colocação que possa ser considerada marcada em inglês. Há diversos exemplos da associação do verbo *flow* com *tears* no CRLI.

O resultado é que houve presença de normalização em todos os casos, uma vez que as escolhas feitas pelos tradutores não mantiveram a mesma colocação incomum do original.

Linha 12: “Era gordo e pesado, tinha a respiração curta e os **olhos dorminhocos**.”

Traduções:

- He was a fat, heavy man, short of breath and **sleepy-eyed**. (Caldwell)
- He was fat and heavy, short of breath and with **sleepy eyes**. (Bucclench)
- He was fat and heavy, short of breath and with **sleepy eyes**. (Gledson)

Em todo o CRLP, não há uma única ocorrência da colocação “olhos dorminhocos”. Pesquisando o Google, e excluindo todos os sites contendo ou fazendo referência a *Dom Casmurro*, há menos de uma dezena de ocorrências. Uma delas numa tradução de *O Mágico de Oz* e as demais em *blogs* variados. Trata-se de mais um exemplo de como Machado atribuía características humanas aos olhos de suas personagens (na linha 1, há outro exemplo deste uso).

De certa forma, todos os tradutores optaram pelo adjetivo *sleepy*, equivalente mais próximo ao original “dorminhoco”. Por meio de uma busca pela colocação “sleepy eyes” no corpus de referência, foram encontradas diversas ocorrências, principalmente em obras literárias, conforme os exemplos abaixo:

1. ...little ones kneeling round, their **sleepy eyes** blinking and red... (*Tess of the D'urbervilles*, Thomas Hardy)
2. ...dark bushy hair, and dull **sleepy blue eyes** (*Walden*, Henry David Thoreau)
3. ...apple-faced old boy, with **sleepy eyes** and a grey moustache... (*The man who was Thursday*, G. K. Chesterton)
4. ...of his neck and smiled at Stephen with **sleepy eyes**... (*A portrait of the artist as a young man*, James Joyce)

Os dados indicam que este caso pode ser explicado pelo fato de que a colocação, em português, é marcada, mas não no idioma inglês. Não se pode afirmar que se trata de uma ocorrência do fenômeno da normalização, pois os tradutores simplesmente utilizaram a tradução padrão para “dorminhoco”. Mesmo Caldwell opta por uma forma que está dicionarizada, representando uma colocação não-marcada.

Linha 13: “(...) depois de algum tempo de silêncio, **recolhendo os olhos** e um suspiro.”

Traduções:

- (...) after some silence, as he **withdrew his eyes** and sighed. (Caldwell)
- Buccleuch eliminou este trecho de sua tradução (v. 4.4.3).
- (...) after staying silent for a while, he **brought his eyes back into focus** and sighed. (Gledson)

Machado faz com que um ex-colega de seminário de Bentinho *recolha* os olhos, como se estes fossem algo externo a ele. De fato, as colocações mais usuais como objeto direto do verbo “recolher” são coisas que se pode pegar e mudar de lugar, como roupa, tropa, armas, lixo, material reciclável, etc.

Na busca por ocorrências semelhantes no CRLP, há apenas uma outra associando olhos e o verbo recolher, a saber, *Esau e Jacó*, do próprio Machado. Isso serve para corroborar a relevância que o autor dá aos olhos de suas personagens, tornando-os quase algo autônomo, com vida própria. Pelo resultado da busca, não há dúvidas de que este exemplo possa ser considerado colocação marcada.

Os dois tradutores que preferiram não eliminar este trecho fizeram uma interpretação distinta do verbo “recolher”. Caldwell parece o ter entendido no sentido de “puxar para si, retrain-se”, utilizando, em sua tradução, uma colocação usual. *Withdraw + eyes* conta com um número elevado de ocorrências no corpus de referência (em Dickens, Joyce, Hardy, Hawthorne, Austen, Conrad e Sir Walter Scott, somente na parte literária do corpus). Gledson, por outro lado, entende o verbo como “voltar à posição inicial”. Naquele excerto, Bentinho encontra um ex-colega de seminário que lhe mostra um panegírico que havia escrito há 26 anos. Por cortesia, Bentinho lê alguns trechos e comenta ao amigo que o panegírico realmente é muito bonito. Ao fim, o outro, que escutava com “os olhos no ar”, parece voltar a si de um estado de quase delírio, fica algum tempo em silêncio, recolhe os olhos, suspira e exclama: “Tem agradado muito este meu *Panegírico!*”. A opção de Gledson contrasta com a expressão “olhos no ar”, fazendo a personagem “trazer seus olhos

de volta ao foco”. A estrutura *bring + eyes + (into) focus* apresenta uma série de ocorrências no CRLI, ou seja, não representa colocação marcada.

Em ambos os casos, portanto, houve normalização, independente da interpretação dada ao sentido deste trecho.

Linha 14: “(...) levantava-me o queixo e **espetava os olhos** em mim, ansioso também, como a prima Justina na véspera.”

Traduções:

- (...) raised my chin and **fixed his eyes** on me, anxiously, just as Cousin Justina had on the evening before. (Caldwell)
- (...) lifting my chin to **gaze at me**, as curious as Cousin Justina had been the day before. (Bucleuch)
- (...) lifted my chin, and **gazed at me** anxiously, like cousin Justina the evening before. (Gledson)

Antes da análise das traduções, é preciso certificar-se da existência de colocação marcada no original. No CRLP, todas as ocorrências, à exceção de uma (*A Cidade e as Serras*, de Eça de Queirós) são em textos machadianos: nos romances *Memórias Póstumas de Brás Cubas* e *Esau e Jacó*, na novela *Casa Velha* e no conto *Um Errádio*. Esses dados nos permitem afirmar que a colocação *espetar + olhos* é uma forma específica de autor, funcionando como uma marca registrada de Machado.

“Espetar os olhos” é o mesmo que “fixar os olhos, a vista”, o que justifica a escolha dos tradutores pelos verbos *fix* e *gaze*. Não há, no entanto, nenhuma novidade em tal associação, o que é comprovado por dezenas de ocorrências de ambas as opções. Este caso, por esse motivo, pode ser contabilizado como mais um exemplo de normalização.

Linha 15: “Íamos dar com ele, muita vez, **olhos enfiados em si**, cogitando.”

Traduções:

- Many times we would come upon him, **withdrawn within himself**, thinking. (Caldwell)
- We would often come upon him, **his eyes pensive**, lost in thought. (Buckleuch)
- Often we would find him, with **his eyes turned inwards**, thinking. (Gledson)

*Enfiar + olhos* não parece ser colocação exclusiva de Machado. No CRLP, foram encontradas ocorrências desta mesma estrutura em outros autores (Eça de Queirós, José de Alencar e Inglês de Sousa) contemporâneos de Machado. O estranhamento causado neste pesquisador talvez seja devido ao fato de que se trata de uma colocação datada, pouco utilizada por autores modernos. Entretanto, ser datada não significa ser marcada e, portanto, não será considerada para os fins deste estudo.

Linha 16: “Dentro, as minhas outras andorinhas estavam trepadas no ar, os **olhos enfiados nos olhos (...)**”

Traduções:

- Within, my swallows trod air, **eyes tangled in eyes (...)** (Caldwell)
- In the house my two other swallows were floating in the air, **gazing into each other’s eyes (...)** (Buckleuch)
- Inside the house, my two other swallows were floating on air, **their eyes deep in each other’s eyes (...)** (Gledson)

Este caso difere do acima na medida em que o autor altera a tradicional locução “olhos nos olhos”, inserindo ali um verbo inesperado. De fato, uma busca no CRLP não revelou nenhuma outra ocorrência além do próprio *Dom Casmurro*.

Caldwell cria uma colocação que, a julgar pela falta de ocorrências tanto no CRLI como na Internet como um todo, pode ser classificada de marcada, criando uma imagem inusitada, de dois pares de olhos que estão como que embaralhados uns nos outros.

Buccleuch e Gledson, pelo contrário, optam por colocações mais usuais. A escolha deste, apesar de apresentar um número menor de ocorrências, ainda assim não pode ser considerada marcada.

Linha 17: “Nem cuides que pasmou de me ver namorado; achou até natural e **espetou-me** outra vez **os olhos**. ”

Traduções:

- And do not imagine he was amazed to discover that I was in love. He found it quite natural, and once more **his eyes bored into mine**. (Caldwell)
- And you mustn't imagine that he was shocked to learn that I was in love; he found it quite natural, and **gazed at me intently** once more. (Buccleuch)
- And don't think that he was astonished to find me in love; he even found it natural and **his eyes fixed me** again. (Gledson)

Na linha 14, já foi demonstrado que a associação do verbo *espetar* com o substantivo *olhos* é presença constante em diversas obras de Machado. Este caso foi mantido para análise devido às diferentes soluções encontradas pelos tradutores.

Caldwell utiliza o verbo *to bore* para o *espetar* no original. Este verbo, segundo o dicionário Houaiss (2001), quer dizer “fazer um buraco em algo utilizando uma ferramenta”, mas, embora possa parecer incomum, o corpus de referência mostra um número elevado de ocorrências, conforme exemplo abaixo:

1. His **eyes bored** into hers, unblinking.
2. His **eyes bored** into hers like gimlets, cool and beautiful and unforgiving.
3. ... and his dark brown **eyes bored** into hers, as if demanding an answer.
4. His **eyes bored** into hers for another nerve-stretching minute...



Bucclench, que já havia utilizado o verbo *to gaze* em sua tradução para o trecho da linha 14, acrescenta aqui o advérbio *intently*. Mais uma vez, uma busca no CRLI comprova o uso freqüente desta colocação, com mais de duas dezenas de ocorrências.

A análise para o verbo *to fix*, escolhido por Gledson, já fora realizada na linha 14, dispensando nova busca no corpus de referência. Neste exemplo, pode-se afirmar que todos os tradutores demonstraram o fenômeno da normalização lexical em suas opções.

Linha 18: “Os olhos fitavam-se e desfitavam-se (...)”

Traduções:

- **Our eyes met, then looked away (...)**(Caldwell)
- **Our eyes met, then strayed away (...)** (Bucclench)
- **Our eyes stared into one another, then looked away (...)** (Gledson)

Neste caso, a colocação marcada está menos em “fitavam-se” do que em “desfitavam-se”. De fato, há outras ocorrências de *olhos + fitar-se* no CRLP, mas nenhuma outra de *olhos + desfitar-se*. Machado vale-se do recurso de utilizar o verbo “fitar” seguido deste mesmo verbo com um prefixo de negação, o que lhe garante um interessante efeito literário.

Nenhum dos tradutores consegue reproduzir o mesmo efeito. Embora tenham optado por diferentes combinações, nenhuma representa uma colocação marcada. *Eyes + met* aparece no CRLI com mais de 300 ocorrências, e *eyes + stared + into* também está significativamente representado no corpus de referência. Entretanto, essas escolhas referem-se a “fitavam-se”, que não é a principal colocação sob investigação. Para “desfitavam-se”, Caldwell e Gledson usaram *looked away*, ao passo que Bucclench variou o verbo, preferindo *strayed away*. No primeiro caso, o resultado da busca excede as 500 ocorrências, além de estar devidamente dicionarizado, por exemplo, no *Oxford Dictionary of Phrasal Verbs*. A escolha de Bucclench, apesar de não constar do dicionário, também ocorre com uma freqüência que não nos permitiria classificá-la de colocação marcada.

Linha 19: “Imaginou que era um pretexto para mirá-los mais de perto, com os meus **olhos longos, constantes, enfiados neles** (...)”

Traduções:

- She imagined that it was a pretext to look closer, with my **own long, unflickering eyes enmeshed in hers**. (Caldwell)
- She imagined that it was a pretext to look at them much closer, with my **longing eyes steadily fixed on hers** (...) (Buccluch)
- She thought it was a pretext to look at them more closely, with my **eyes longingly fixed on hers** (...) (Gledson)

Neste trecho, das três colocações suspeitas de serem marcadas, pode-se logo descartar *enfiar + olhos*, pelos motivos explicitados na análise da linha 15.

Para *olhos + longos*, uma busca no CRLP revelou que todas as ocorrências são em textos literários; uma em *O Gaúcho*, de José de Alencar, e as demais em obras machadianas:

1. Não teve seguramente os olhos longos e loquazes, como dantes... (*Quincas Borba*)
2. E ficou a rir e a olhar com longos olhos ávidos e felizes... (*Lágrimas de Xerxes*)
3. Como não, se Camila tinha uns longos olhos mortais? (*Pílades e Orestes*)
4. Trazia na lembrança os longos olhos dominadores que ela evitava afrontar... (*Uma Partida*)

Pela recorrência, parece válido analisar este caso em maior profundidade. Machado não parece estar se referindo ao comprimento físico dos olhos da personagem, mas sim ao tipo de olhar, demorado, que permanece um longo tempo fixado em determinado objeto ou pessoa. Talvez o mais esperado fosse a associação de *olhar + longo*. Confirmando esta hipótese, uma busca específica no subcorpus literário revela 26

ocorrências, cinco vezes mais do que *olhos + longos*. Por esse motivo, será considerada colocação marcada.

Na língua inglesa, os dados mostram que a combinação *long + look* é muito mais freqüente do que *long + eyes*, quando se quer referir a um “olhar demorado”. Caldwell, conscientemente ou não, conseguiu manter a mesma singularidade do original, não representando um caso de normalização.

Buckleuch e Gledson utilizam o adjetivo *longing* e o advérbio *longingly*, respectivamente. Se a personagem olha com *longing eyes*, então esses olhos estão desejosos de algo, ansiando pela coisa ou pessoa a quem dirigem o olhar. De fato, o trecho refere-se ao momento em que Bentinho pede para olhar nos olhos de Capitu, a fim de confirmar ou não a afirmação de José Dias, de que a futura esposa tem “olhos de cigana oblíqua e dissimulada”. Apesar de adequadas, ambas as traduções incorrem em normalização lexical, uma vez que os tradutores se valeram de associações comuns na língua inglesa.

Linha 20: “**Recolheu os olhos**, meteu-os em si e deixou-se estar com as pupilas vagas e surdas (...)”

Traduções:

- She **withdrew her eyes**, turned them inward, and let them remain with the pupils vague and unseeing. (Caldwell)
- She **withdrew her gaze** from me and remained still, her eyes expressionless (...) (Buckleuch)
- She **withdrew her eyes**, turned them inwards, and stayed that way, with her pupils vague and sightless (...) (Gledson)

O mesmo raciocínio desenvolvido para a linha 13 pode ser aplicado neste caso. A mudança de *eyes* para *gaze*, em Buckleuch, não altera o resultado. Embora haja menos ocorrências para esta colocação, ainda são em número elevado, não configurando uma colocação marcada.

Linha 21: “E, após alguma reflexão, fitando em mim uns **olhos murchos e teimosos**, perguntou-me (...)”

Traduções:

- And after some reflection, he peered into my face with **faded, insistent eyes**, and asked (...) (Caldwell)
- Buccleuch eliminou este trecho de sua tradução (v. 4.4.3). (Buccleuch)
- Then, after some reflection, fixing me with **faded, insistent eyes**, he asked me (...) (Gledson)

A colocação *olhos + murchos* ainda se pode encontrar em *Helena* e *Iaiá Garcia*, ou seja, mais dois romances machadianos.

1. ...ao passo que os **olhos** doridos e **murchos** pareciam reviver de um resto de luz. (*Helena*)
2. ...Estela tinha o rosto desfeito e **murchos** os belos **olhos**. (*Iaiá Garcia*)

Este caso assemelha-se ao da linha 14, em que a colocação investigada aparece em outros textos de Machado. Quanto à escolha de Caldwell e Gledson, *faded + eyes* apresenta um número elevado de ocorrências, não se configurando em colocação marcada. Talvez a colocação *withered + eyes* (utilizando a primeira tradução para *murcho* do dicionário de Taylor) conseguiria evitar a normalização, uma vez que sua ocorrência é 90% menor no CRLI.

Mais uma vez, uma busca pela associação de *olhos + teimosos* indica ser esta uma preferência de Machado. Além de *Dom Casmurro*, está presente nos contos *A Cartomante* e *D. Paula*:

1. Agora a ação da pessoa, os **olhos teimosos** de Rita, que procuravam muita vez os dele... (*A Cartomante*)
2. ...e surgia outra vez com os seus grandes **olhos** sagazes e **teimosos**. (*D. Paula*)

A frequência na escolha dos tradutores indica tratar-se de uma colocação usual, dada sua alta ocorrência no corpus de estudo. A normalização, portanto, está presente para os dois adjetivos associados a “olhos” neste caso.

Linha 22: “Tinha-me lembrado a definição que José Dias dera deles, ‘**olhos de cigana oblíqua e dissimulada**’.”

Traduções:

- I had remembered the definition that José Dias had given of them, “**gypsy’s eyes, oblique and sly.**” (Caldwell)
- I had remembered José Dias’s description of them, ‘**sly and cunning like a gypsy’s.**’ (Buckleuch)
- I had remembered the definition José Dias had given them; “**a gypsy’s eyes, oblique and sly.**” (Gledson)

Junto com “olhos de ressaca”, esta é a colocação mais famosa do romance. Sua originalidade é inegável, e a busca no CRLP só confirma este fato. Há duas ocorrências, sendo que uma delas é no próprio *Dom Casmurro*, e a outra é uma referência intertextual no conto *Grogue*, de Toni Brandão:

1. Ela desliga a câmera. Liga um sorriso. Que combina com o olhar misterioso. Dissimulado. De cigana oblíqua. De ressaca. De Capitu.

Em Machado, os adjetivos “oblíqua” e “dissimulada” referem-se à “cigana”, e não a “olhos”. Nas traduções, por outro lado, *oblique*, *sly* e *cunning* estão relacionados à palavra *eyes*. Portanto, no corpus de referência, a busca deve ser feita associando [*oblique*, *sly*, *cunning*] + *eyes*.

No CRLI, não há uma única ocorrência para *oblique* + *eyes*. Quando pesquisada a Internet como um todo, nota-se que muitas das ocorrências estão em textos da área médica, mais especificamente da oftalmologia, o que não serve de comparação com o uso metafórico deste adjetivo no texto original. Não há aqui, portanto, a ocorrência de

normalização em Caldwell (a tradução de Gledson foi publicada 40 anos após a de Caldwell, e o próprio autor, em comunicação pessoal, admitiu ter se valido de boas escolhas feitas pela tradutora norte-americana).

As associações [*sly, cunning*] + *eyes*, por outro lado, apresentam um número de ocorrências que não nos permitiriam classificá-las de marcadas. O fenômeno da normalização está ausente somente na primeira análise.

Linha 23: “Eu levantei-me depressa e não achei compostura; **metia os olhos pelas** cadeiras.”

Traduções:

- I rose quickly. I lost my composure and **kept my eyes on** the furniture. (Caldwell)
- I jumped to my feet hastily, in such confusion that I **could not meet his eye**. (Buckleuch)
- I got up quickly and could not keep my composure: I **fixed my eyes on** the chair. (Gledson)

Além de em *Dom Casmurro*, *meter* + *olhos* está presente em outras três obras machadianas:

1. ...os **olhos metidos** para dentro viam pensar o cérebro. (*Quincas Borba*)
2. **Meti os olhos** no chão, e fui andando. (*O Enfermeiro*)
3. Joanhinha meteu os olhos no regaço, vexada do que ele iria dizer. (*O Diplomático*)

No CRLP, aparece ainda em *A Conquista*, de Coelho Neto (...**metia os olhos** indiscretos por todas as janelas...). À exceção desses exemplos, não há mais ocorrências dessa colocação. A recorrência em outros textos de Machado indica tratar-se de colocação marcada.

A estrutura *keep* + pronome possessivo + *on* ocorre em dezenas de exemplos no CRLI. Não há, portanto, nada de marcado na colocação escolhida por Caldwell.

Bucclench parece optar por uma saída mais fácil, sem se dar ao trabalho de procurar por um equivalente em sua língua materna. Além de excluir o complemento (“pelas cadeiras”), o tradutor escocês utiliza uma colocação não-marcada (*meet + eyes*).

Ao trocar o verbo *keep*, escolhido por Caldwell, for *fix*, Gledson também não consegue evitar a normalização. Pelo contrário, sua opção é a mais freqüente no corpus de referência.

Linha 24: “(...) e não vi que essa menina travessa e já de **olhos pensativos** era a flor caprichosa de um fruto sadio e doce...”

Traduções:

- (...) and did not see that this mischievous little girl with the **pensive eyes** was the capricious flower of a sweet, wholesome fruit... (Caldwell)
- (...) not realising that the mischievous girl with the **thoughtful eyes** was the capricious blossom of a sweet, wholesome fruit. (Bucclench)
- (...) and didn't see that that mischievous girl, already with her **thoughtful eyes**, was the capricious bloom which would produce such a sweet, wholesome fruit... (Gledson)

A colocação acima, inicialmente intuída como sendo marcada, revelou-se não-marcada na busca ao CRLP. Apesar de Machado associar *olhos + pensativos* em outras obras (*A Mão e a Luva*, *A Chinela Turca* e *Um Errádio*), essa colocação aparece em outros textos do corpus de referência, todos pertencentes ao subcorpus literário. Isso indica que, apesar de incomum em, digamos, textos jornalísticos, não pode ser considerada uma colocação marcada quando a referência é um corpus exclusivamente literário.

Quanto às traduções, todas mantêm o caráter de colocação não-marcada. Apesar de *pensive + eyes* ter menos ocorrências do que as escolhas de Bucclench e Gledson, ainda assim não constituem exemplos de colocações marcadas na língua inglesa.

Linha 25: “(...) o meu amigo Escobar era um tanto metedido e tinha uns **olhos policiais** a que não escapava nada.”

Traduções:

- (...) my friend Escobar was a bit meddlesome and had **policeman’s eyes** that missed nothing. (Caldwell)
- (...) my friend Escobar was a little too noseey and had **eyes like a detective’s**, which missed nothing. (Bucclench)
- (...) my friend Escobar was a little inquisitive and that he had **eyes like a policeman**, that took everything in. (Gledson)

Foi encontrada apenas uma segunda ocorrência de *olhos + policiais*, em um poema de Carlos Drummond de Andrade intitulado *Elegia*:

“E vou me recolher  
ao cofre de fantasmas, que a notícia  
de perdidos lá não chegue nem açule  
os **olhos policiais** do amor-vigia.  
Não me procurem que me perdi eu mesmo  
como os homens se matam, e as enguias  
à loca se recolhem, na água fria.”

Justamente por ter ocorrido em um poema, e em um autor sabidamente inventivo, esta colocação será tratada como marcada na seguinte análise.

Deve-se ressaltar que “policial”, em português, pode ser tanto substantivo como adjetivo. Tanto Machado quanto Drummond fazem uso do adjetivo, e não do substantivo. Isso equivale a dizer que os olhos de Escobar se comportam à maneira de um policial, “a que não escapam nada”. Em inglês, o termo “police” refere-se ao substantivo polícia (a organização), ou ao verbo policiar.

Bucclench e Gledson escapam a essa dificuldade comparando os olhos da personagem com os de um “detective” ou “policeman”. No entanto, este é um claro



exemplo de normalização, uma vez que *eyes + like + a/an* é uma colocação habitual da língua inglesa, com dezenas de exemplo no corpus de referência. Em Caldwell, também não há evidências de qualquer tentativa de manter a colocação marcada do original no texto de chegada. Mais uma vez, comprova-se o fenômeno da normalização nas três traduções.

Linha 26: “(...) os **olhos** que ele **me deitou** foram ternos e agradecidos.”

Traduções:

- (...) his **answering look** was tender and grateful. (Caldwell)
- (...) the **look he gave me** was one of tenderness and gratitude. (Buccluch)
- (...) the **look he gave me** was tender and grateful. (Gledson)

*Deitar + olhos* é uma colocação recorrente na obra machadiana. À exceção de *Dom Casmurro*, aparece em outras dezessete obras. No entanto, esta alta prevalência não é privilégio de Machado. Há mais cinco dezenas de exemplos no CRLP. Cabe salientar que todas as ocorrências se encontram no subcorpus literário, basicamente composto de textos mais antigos, que já caíram em domínio público. Em textos mais recentes (jornalísticos, didáticos, acadêmicos), a colocação acima está ausente. Com isso, pode-se concluir que “deitar olhos” costumava ser muito usado na literatura contemporânea de Machado, mas caiu em desuso, e sua ocorrência é quase nula hoje em dia.

Por esse motivo, não pode ser classificada de colocação marcada e, portanto, não serve para os fins desta pesquisa.

Linha 27: “Não havia meio de esquecer inteiramente a mão de Sancha nem **os olhos que trocamos**.”

Traduções:

- There was no forgetting the hand of Sancha, nor **the looks we had traded**. (Caldwell)

- I could not entirely forget Sancha’s hand, nor **the looks we had exchanged**. (Buckleuch)
- There was no way of completely forgetting Sancha’s hand, nor **the look we exchanged**. (Gledson)

Nenhuma outra ocorrência para *trocar + olhos*, mas diversos exemplos para *trocar + olhar(es)*, o que indica que esta, e não aquela, é a colocação usual. Machado é criativo ao transformar os olhos de suas personagens (na passagem sob análise, os olhos de Sancha e Bentinho) em objetos que parecem ser exteriores ao próprio corpo, sendo passíveis de “troca”.

Uma breve análise das traduções indica que todos os tradutores evitaram utilizar o substantivo *eyes*. Resta saber se, na língua inglesa, *exchange/trade + look(s)* é mais usual do que *exchange/trade + eyes*.

De fato, essa colocação comporta-se de forma similar no CRLI. *Look(s)* é o objeto direto mais utilizado, ao passo que *eyes* não ocorre uma única vez em associação com *exchange/trade*. Para evitar a normalização, os tradutores deveriam ter optado por “exchanged/traded eyes”, talvez causando estranhamento ao leitor, mas mantendo o mesmo grau de criatividade do original.

Linha 28: “Tudo isso **me empanava os olhos**.”

Traduções:

- All this **made my eyes grow dim**. (Caldwell)
- **My eyes misted over**. (Buckleuch)
- All this **made my eyes mist over**. (Gledson)

Ao contrário da intuição deste pesquisador, Machado não detém a exclusividade da colocação *empanar + olhos*. Diversos outros autores a utilizam, embora alguns com pequenas alterações, conforme os exemplos abaixo:

1. Ia-se-lhe empanando o brilho dos olhos. (*Inocência*, Visconde de Taunay)
2. (...) como se já estivesse de todo passada a nuvem que empanava o sol dos seus olhos. (*Triste Fim de Policarpo Quaresma*, Lima Barreto)
3. Um lampejo de alegria raiou nos olhos empanados do moribundo (...) (*O Garimpeiro*, Bernardo Guimarães)
4. Súbito abriu os olhos empanados. (*Noite na Taverna*, Álvares de Azevedo)

Se não há colocação marcada no original, não há motivo para investigar as soluções adotadas pelos tradutores. Trata-se de uma colocação não-marcada, que não serve para os fins deste estudo.

## 7. CONCLUSÃO

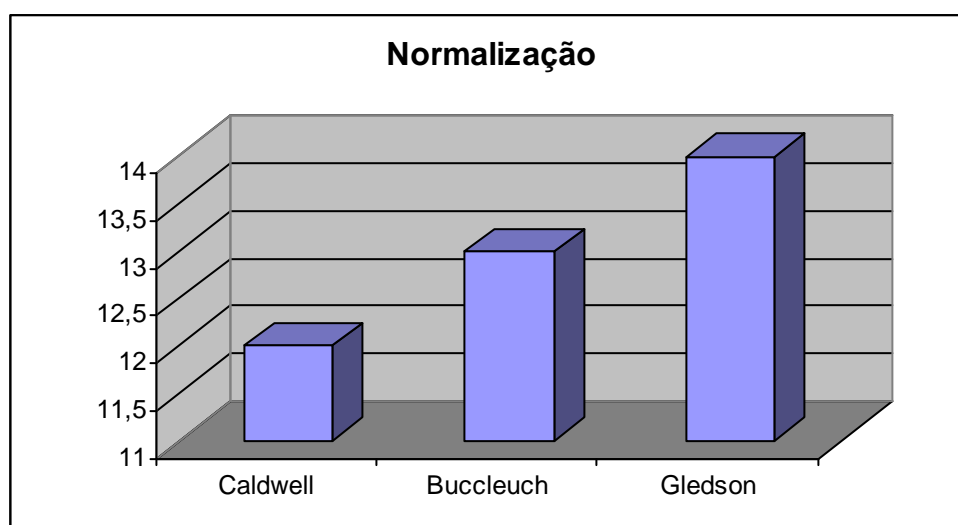
A presente pesquisa, que tinha por uma de suas motivações iniciais preencher uma lacuna na área dos estudos da tradução no Brasil, cumpriu com o que se propôs, revelando ser a investigação da ocorrência de normalização em traduções a partir de colocações marcadas utilizadas nos textos originais um campo de pesquisa fértil e promissor.

Este trabalho teve por objetivo investigar a ocorrência de normalização, tendo por objeto de estudo um corpus paralelo composto pelo romance *Dom Casmurro*, de Machado de Assis, e três traduções (por diferentes tradutores em diferentes épocas) para a língua inglesa. A normalização foi pesquisada em nível lexical, através de colocações marcadas, que, por sua vez, representam uma forma de criatividade lexical. Para que os itens lexicalmente criativos fossem encontrados com maior facilidade, a abordagem adotada foi a baseada em corpus. Para isso, a referida obra machadiana e suas respectivas traduções foram digitalizadas a fim de serem passíveis de análise por meio digital. Os corpora de referência – um para o português brasileiro, outro para a língua inglesa – serviram para comprovar ou não a intuição inicial do pesquisador. É possível supor de antemão que determinada colocação é marcada, a partir do conhecimento que temos da língua. No entanto, os dados apresentados pelos corpora, por vezes, comprovaram tratar-se de suposição equivocada, oferecendo maior fidedignidade aos resultados. Por esse motivo, foram excluídas da análise algumas linhas pré-selecionadas com base somente em julgamento intuitivo (linhas 2, 15, 26 e 28).

No capítulo de análise de dados, houve a necessidade de escolher um nóculo para aprofundar a busca de colocações marcadas. Pelos motivos explicitados nesse capítulo, o nóculo “olhos” foi escolhido, mostrando ter um tratamento especial por parte de Machado, que diversas vezes, e em diferentes obras, confere características humanas aos olhos de suas personagens (v. linha 12 da análise dos dados). Após uma busca de todas as suas ocorrências no corpus de estudo com o auxílio de programas de computador, foram encontradas 164 ocorrências. Destas, através de uma pré-seleção manual, foram extraídas 28 para uma investigação mais aprofundada com o suporte dos corpora de referência. Passemos, agora, a uma discussão dos resultados da análise dos dados, mas não sem antes revisitar as perguntas de pesquisa. Pretendia-se averiguar o seguinte: (i) se houve normalização por parte dos tradutores quando da tradução de colocações marcadas do

original; (ii) de que forma aconteceram; (iii) se ocorreram com maior intensidade em alguns casos; (iv) se é possível generalizar a questão da tradução de colocações marcadas; e (v) se as generalizações são válidas somente para o par português-brasileiro/inglês, ou são corroboradas por outros estudos.

Das 28 ocorrências do nóculo “olhos” com colocações possivelmente marcadas, 18 foram efetivamente analisadas, após a exclusão de repetições (linhas 4 a 10) e colocações não-marcadas (linhas 2, 15, 26 e 28). O gráfico 1 mostra o número de casos de normalização separados por tradutor.



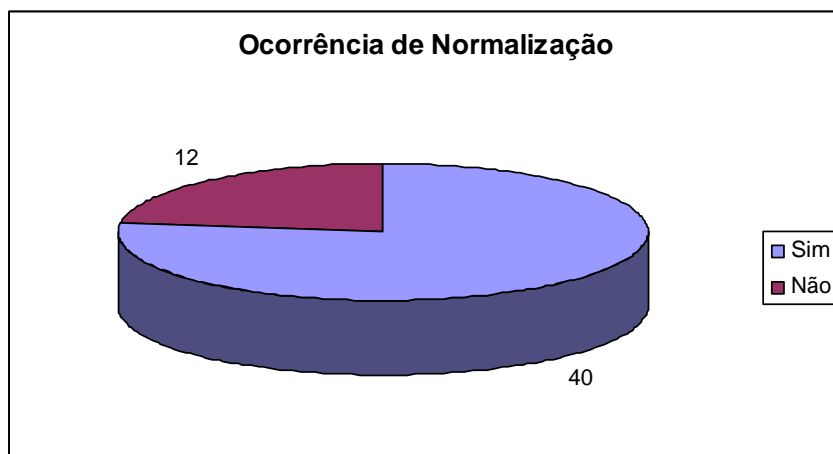
**Gráfico 1.** Ocorrência de normalização na tradução de colocações marcadas, de acordo com cada tradutor.

Caldwell, com o menor índice de todos os tradutores, ainda assim teve um percentual de 66,6% de normalizações (12/18). Buccleuch, com 13 ocorrências, aparece com 72,2%, considerando que este tradutor não pôde ser avaliado em dois casos (linhas 13 e 21), pois omitiu o trecho selecionado para análise (v. 4.4.3). Gledson, por sua vez, apresentou o maior índice de normalizações, com 14 casos, equivalente a 77,77% de ocorrência de normalização.

O fato de Caldwell aparecer como a tradutora com menor prevalência de normalização não surpreende. Conforme mencionado na segunda parte da análise da linha 1, a crítica literária Daphne Patai já havia identificado a característica desta tradutora norte-

americana em manter os traços da obra original quando não havia opções aceitáveis na sua língua materna.

O gráfico 2 mostra o resultado do número total de normalizações (três traduções em cada linha selecionada, menos duas traduções omitidas por Bucleuch) sem distinguir por tradutor.



**Gráfico 2.** Ocorrência de normalização considerando o total de traduções avaliadas.

Conforme se pode observar, houve uma prevalência alta do fenômeno da normalização. De 52 casos investigados, 40 tiveram suas traduções normalizadas, eliminando o caráter de colocação marcada existente no trecho original. Em termos percentuais, pode-se afirmar que a normalização esteve presente em aproximadamente 77% dos casos, valor este que significativo e que serve para corroborar a hipótese de que há uma tendência por parte dos tradutores em normalizar colocações marcadas. Isso responde à primeira pergunta levantada por esta pesquisa: as colocações marcadas encontradas no original foram, de fato, normalizadas pelos tradutores.

Partiremos agora para a segunda questão, referente à forma como tal processo ocorreu. Considerando os casos de normalização identificados na análise dos dados, o que se pode constatar é uma tendência que os tradutores apresentaram em evitar a estranheza da colocação ao traduzir para suas línguas maternas. Para isso, optaram por uma colocação não-marcada em língua inglesa em substituição a colocações marcadas do original. Talvez por consciência de que suas traduções passariam por uma revisão editorial (vale lembrar que estamos aqui no território de suposições, pois tal hipótese seria impossível de se

comprovar), os tradutores tenham preferido, de certa forma, “facilitar” a tradução para os futuros leitores, evitando arriscar uma tradução que emulasse a colocação marcada da obra original para não serem acusados de uma tradução truncada, de difícil compreensão.

A pergunta seguinte trata da maior ou menor intensidade em alguns casos de normalização. Não foi possível localizar um padrão neste sentido, a não ser pelo fato de que houve uma frequência elevada de normalização na seguinte combinação de colocação: [olhos] + verbo e verbo + [olhos]. Por outro lado, houve casos de colocação [olhos] + substantivo que também foram normalizados, por isso não podemos chegar a nenhuma conclusão específica quanto a esta questão.

A amostra pequena contida neste estudo (apenas uma obra e três traduções) torna perigoso qualquer tipo de generalização. Por esse motivo, é preciso ter cuidado ao tentar responder à questão de número quatro, referente à possibilidade de generalizar o modo como os tradutores lidam com colocações marcadas no original. Entretanto, o que se pode afirmar é que alguns tradutores tendem a normalizar mais do que outros e que a maior causa de colocações marcadas no original foi a característica muito peculiar de Machado em atribuir uma vida quase própria aos olhos de suas personagens (*olhos ansiosos, teimosos, curiosos*, ou então *mandar olhos, recolher os olhos*, etc.). Essa constatação nos impede de arriscar qualquer generalização, pois bastaria que outro autor estivesse sob investigação para que resultados distintos fossem encontrados.

Finalmente, quanto à possibilidade aventada pela quinta pergunta de pesquisa, relativa à validade das generalizações para qualquer par de idiomas, faz-se necessária uma comparação com o único estudo encontrado na literatura sobre normalização lexical, a saber, Kenny (2001). Essa autora chega à conclusão de que a “evidência disponível era esparsa” (op. cit., p. 210)<sup>39</sup>, justificando que, para uma conclusão mais consistente seria necessário ter um corpus paralelo mais representativo, com maior número de obras e autores. O mesmo vale para a presente pesquisa, que contou apenas com uma obra em língua portuguesa e três traduções para o inglês. A outra conclusão a que chegou Kenny não é válida para o par português brasileiro/inglês. Trata-se de uma alta prevalência de normalização em características lexicais do texto-fonte referentes à formação de palavras na língua alemã (derivação e conversão). Por se tratarem de processos distintos, não

---

<sup>39</sup> “available evidence was sparse”.

podemos afirmar que o mesmo seja válido para o português, o que é corroborado pela não ocorrência de nenhum caso de neologismo entre as colocações marcadas analisadas.

Uma das limitações do presente estudo, conforme já mencionado anteriormente, é o tamanho amostral do corpus de estudo. Para se obter resultados mais precisos, seria necessário contar com um número mais representativo de autores e obras no corpus paralelo, com originais e suas respectivas traduções. A falta de material digitalizado disponível é um dos maiores empecilhos para o desenvolvimento de pesquisas desta natureza. Além disso, há a questão dos direitos autorais, impedindo com que muitos textos sejam incluídos em um corpus de estudo para posterior investigação. Quanto a outras pesquisas nessa área, fica a sugestão de mais estudos relacionados ao mesmo tema – normalização lexical em colocações marcadas – utilizando o mesmo par de idiomas, mas incluindo outros autores e obras, para que se possa aumentar o volume de dados e possibilitar generalizações mais abrangentes e precisas a partir dos resultados encontrados.



## REFERÊNCIAS

Assis, Machado de. *Obras Completas*. 3 volumes. Rio de Janeiro, Nova Aguilar, 1959.

\_\_\_\_\_. *Dom Casmurro*. Trad. de Helen Caldwell. New York: The Noonday Press, 1991.

\_\_\_\_\_. *Dom Casmurro*. Trad. de Robert Scott-Bucleuch. New York: Penguin Books, 1994.

\_\_\_\_\_. *Dom Casmurro*. Trad. de John Gledson. Oxford: Oxford University Press, 1997.

BAKER, Mona. *In other words: a coursebook on translation*. London and New York: Routledge, 1992.

\_\_\_\_\_. "Corpus linguistics and translation studies: implications and applications", in M. Baker, G. Francis, & E. Tognini-Bonelli (orgs.) *Text and technology: in honour of John Sinclair*. Philadelphia/Amsterdam: John Benjamins, pp. 233-250, 1993.

\_\_\_\_\_. "Corpora in translation studies: an overview and some suggestions for future research". *Target 7*: 223-43, 1995.

BRITISH NATIONAL CORPUS. Disponível em: <http://www.natcorp.ox.ac.uk>. Acesso em junho de 2006.

BUENO, Letícia Taitson. *Transitividade, coesão e criatividade lexical no corpus paralelo Macunaíma, de Andrade e Macunaíma, de Goodland*. Belo Horizonte: Universidade Federal de Minas Gerais, 2005.

CÂNDIDO, Antônio. *Vários escritos*. 3 ed. São Paulo: Duas Cidades, 1995.

COUTINHO, Afrânio. *A literatura no Brasil*. 3 ed. Rio de Janeiro: José Olympio; Niterói: UFF-Universidade Federal Fluminense, 1986.

FERNANDES, Lincoln. *Corpora in translation studies: revisiting Baker's typology*. Revista Fragmentos, 2006 (no prelo).

FIRTH, J. R. *Papers in Linguistics – 1934-1951*. Oxford: Oxford University Press, 1957.

GLEDSON, John. *Machado de Assis: impostura e realismo: uma reinterpretação de Dom Casmurro*. Tradução de Fernando Py. São Paulo: Companhia das Letras, 1991.

\_\_\_\_\_. *Traduzindo Machado de Assis*. Tradução de Luana Freitas (no prelo).

HOLMES, James. *The name and nature of translation studies*. Translated! Papers on literary translation and translation studies. Amsterdam: Rodopi, 1988.

HOUAISS, Antonio. *Dicionário Eletrônico da Língua Portuguesa*. São Paulo: Objetiva, 2001.

JOHANSSON, S. "On the role of corpora in cross-linguistic research", in S. Johansson and S. Oksefjell (Eds.) *Corpora and cross-linguistic research*. Amsterdam: Rodopi, pp. 3-24, 1998.

KENNY, Dorothy. *Lexis and creativity in translation: a corpus-based study*. Manchester: St. Jerome, 2001.

\_\_\_\_\_. *Creatures of Habit? What translators usually do with words*. *Meta*, XLIII, 4, 1998. Disponível em: <http://www.erudit.org/erudit/meta/v43n04/index.html>. Acesso em junho de 2006.

OLOHAN, Maeve. *Introducing corpora in translation studies*. London and New York: Routledge, 2004.

PAGANO et al. (orgs.) *Estudos da Tradução no Brasil: Resumos de teses e dissertações* (CD-ROM). Belo Horizonte: B B Martins, 2001.

PARTINGTON, Alan. *Patterns and meanings: using corpora for English language research and teaching*. Amsterdam and Philadelphia: John Benjamins, 1998.

PATAI, Daphne. *Machado in English*. in: Machado de Assis: reflections on a Brazilian master writer. Austin: University of Texas Press, 1999.

SARDINHA, Tony Berber. *Linguística de corpus*. São Paulo: Manole, 2004.

SINCLAIR, John. *Corpus, concordance, collocation*. Oxford: Oxford University Press, 1991.

\_\_\_\_\_. *Council of Europe Multilingual Lexicography Project*. Informe não publicado apresentado ao Conselho Europeu sob contrato nº 57/89, 1991b.

TAYLOR, James L. *Portuguese/English Dictionary*. 14<sup>a</sup> ed., Rio de Janeiro: Record, 2001.

WILLIAMS, J. & CHESTERMAN, A. *The map. A beginner's guide to doing research in translation studies*. Manchester: St. Jerome, 2002.

ZANETTIN, F. "Parallel corpora in translation studies: issues in corpus design and analysis". In M. Olohan (Ed.), *Intercultural faultlines. Research models in translation studies I: textual and cognitive aspects*. Manchester, UK: St. Jerome, 2000.

Anexo 1 – Lista de ocorrências para o nódulo *olhos* no corpus de estudo, reordenada de acordo com o primeiro item à direita do nódulo principal.

1 dissimulado; levantou o olhar, sem levantar os **olhos**. A voz, um tanto sumida, perguntou-me  
2 a do pai, não é? --Sim, tem alguma cousa, os **olhos**, a disposição do rosto. É o pai, um pou  
3 o caso de Capitu. Cuidei de recompor-lhe os **olhos**, a posição em que a vi, o ajuntamento d  
4 éis. Escobar pegou o papel, passou-os pelos **olhos** a fim de os decorar. e enquanto eu fitav  
5 a com o espírito no menino; em casa, com os **olhos** a observá-lo, a mirá-lo, a perguntar-lhe  
6 não ser que fosse para certificar aos próprios **olhos** a realidade que o coração lhe dizia... F  
7 oficiais do homem é fechar e apertar muito os **olhos** a ver se continua pela noite velha o son  
8 sida às saias de minha mãe, não atendia aos **olhos** ansiosos que eu lhe mandava; também  
9 S E MIL AVE-MARIAS Levantei os **olhos** ao céu, que começava a embruscar-se,  
10 e a cara entre as mãos e chegando muito os **olhos** aos meus. Então eu esperei tantos ano  
11 rgia e obstinação que empreguei em fechar os **olhos**, apertá-los bem, esquecer tudo para dor  
12 chamando-me seu filho. Estava queimando os **olhos** ardiem nos meus, toda ela parecia con  
13 lguma cousa, um soneto. A insônia, musa de **olhos** arregalados, não me deixou dormir uma  
14 ira para o lado de papai, não precisa revirar os **olhos**, assim, assim... Era depois de jantar, e  
15 nadar. Quando saímos, tornei a falar com os **olhos** à dona da casa. A mão dela apertou m  
16 uação é infinita. A recordação de uns simples **olhos** basta para fixar outros que os recordem  
17 pernas como armas de fuga ou de defesa. Os **olhos** bastavam ao primeiro efeito. Um mover  
18 inho? Para não fitá-lo, deixei cair os olhos. Os **olhos**, caindo, viram que uma das presilhas d  
19 achei até um jeito dos pés de Escobar e dos **olhos**... Capitu deixou-se estar pensando e ol  
20 le há de falar. --Você jura? --Juro. Deixe ver os **olhos**, Capitu. Tinha-me lembrado a definição  
21 do tempo, desciam-lhe pelas costas. Morena, **olhos** claros e grandes, nariz reto e comprido,  
22 ia, como a tia, a mesma cabeça, os mesmos **olhos** claros, a mulher é que lhe disse que o  
23 cinco anos, um rapagão bonito, com os seus **olhos** claros, já inquietos, como se quisesse  
24 uiel de Sousa Escobar era um rapaz esbelto, **olhos** claros, um pouco fugitivos, como as mã  
25 e calor que as senti estremecer. Enxugou os **olhos** com os dedos, eu os bejei de novo, por  
26 é indispensável, e estou às suas ordens. Os **olhos** com que me disse isto eram embuçado  
27 se jantar, a fim de se lhe fazer uma saúde, os **olhos** com que aceitou seriam de protonotário  
28 ras uma só, mm uma só criatura seráfica. Os **olhos** continuaram a dizer cousas infinitas, as  
29 ode. Papai quer ver? E séria, fitou em mim os **olhos**, convidando-me ao jogo. O susto é natu  
30 e chorou de veras, ou se somente enxugou os **olhos**; cuidou que os enxugou somente. Vendo  
31 ra quadragenária, magra e pálida, boca fina e **olhos** curiosos. Vivia conosco por favor de mi  
32 que os não esganei um dia. quando desviei os **olhos** da rua onde estavam duas andorinhas tr  
33 om certeza, pela boca fora. Não podia tirar os **olhos** daquela criatura de quatorze anos, alta,  
34 DIA Assim chorem por mim todos os **olhos** de amigos e amigas que deixo neste m  
35 a retinha também. Momento houve em que os **olhos** de Capitu fitaram o defunto, quais os da  
36 de algumas hesitações, resolveu dar-lha. Os **olhos** de Capitu, quando recebeu o mimo, não  
37 em da história! A pérola de César acendia os **olhos** de Capitu. Foi nessa ocasião que ela p  
38 xata e poética para dizer o que foram aqueles **olhos** de Capitu. Não me acode imagem capa  
39 dispôs-se a correr; o menino, sem tirar-lhe os **olhos** de cima, fez-nos outro sinal de silêncio;  
40 e acabou com a mesma tristeza da outra. Os **olhos**, de costume fugidios, quase me comer  
41 equiel, achei que Capitu tinha razão; eram os **olhos** de Escobar, mas não me pareceram es  
42 sse-me: --Vê-se que era um coração puro! Os **olhos** de Escobar, claros como já disse, eram  
43 esa; mas, dizendo-lhe eu que, na beleza, os **olhos** de Ezequiel saíam aos da mãe, Capitu

44 capítulo passado e dos futuros; fiquemos nos **olhos** de Ezequiel. CAPÍTULO CX  
45 te ponto,--lembra-me como se fosse hoje,--os **olhos** de José Dias fulguraram tão intensame  
46 eu digo. Vem amanhã. Sancha não tirava os **olhos** de nós durante a conversa, ao canto da  
47 pensei nisso na cama. Só então senti que os **olhos** de prima Justina, quando eu falava, par  
48 ais me envolviam no seu mistério, sem que os **olhos** de ressaca de Capitu deixassem de cre  
49 ada? Capitu fez um gesto de impaciência. Os **olhos** de ressaca não se mexiam e pareciam  
50 o e comborço Escobar. Cheiram também aos **olhos** de ressaca de Capitu. Assim, posto se  
51 l seria o discurso. CAPÍTULO CXXIII / **OLHOS DE RESSACA** Enfim, chegou  
52 m adágio dacolá... CAPÍTULO XXXII: **OLHOS DE RESSACA** Tudo era maté  
53 eu coração? Talvez porque nenhuma tinha os **olhos** de ressaca, nem os de cigana oblíqua e  
54 epente, cessando a reflexão, fitou em mim os **olhos** de ressaca, e perguntou-me se tinha m  
55 mitológico. Ainda há pouco, falando dos seus **olhos** de ressaca, cheguei a escrever Tétis; ri  
56 ade do estilo, o que eles foram e me fizeram. **Olhos** de ressaca? Vá, de ressaca. É o que  
57 o se me daria beijá-la na testa. Entretanto, os **olhos** de Sancha não co  
58 tivo. A noite era clara. Dali mesmo busquei os **olhos** de Sancha, ao pé do piano; encontrei-o  
59 deza e de espiritualidade que falava aos meus **olhos** de seminarista. Era não menos que isto  
60 pois sobre ela rosto a rosto, mas trocados, os **olhos** de uma na linha da boca do outro. Pedi-  
61 fazendo vir do credor a relevação da dívida. Os **olhos** dela brilharam, mas a boca disse que n  
62 os que o Diabo lhe deu... Você já reparou nos **olhos** dela? São assim de cigana oblíqua e di  
63 s policiais a que não escapava nada. --São os **olhos** dele, expliquei. --Nem eu digo que seja  
64 e compridas eram agora breves e alegres; os **olhos** desaprenderam de chorar, se porventur  
65 são: --Conto com o senhor para salvar-me. Os **olhos** do agregado escancararam-se, as sobr  
66 dre, e que reparasse no padre, não tirasse os **olhos** do padre. Em casa, brincava de missa,-  
67 gordo e pesado, tinha a respiração curta e os **olhos** dorminhocos. Uma das minhas recorda  
68 fechar os olhos. Quando novamente abria os **olhos** e a carta, a letra era clara e a notícia cl  
69 que nada, mas não achei língua. Todo eu era **olhos** e coração, um coração que desta vez ia  
70 grimas. Capitu limitou-se a arregalar muito os **olhos**, e acabou por dizer: --Padre é bom, não  
71 ha mãe apalpava-lhe o coração, revolvía-lhe os **olhos**, e o meu nome era entre ambas como  
72 . O que faço, em chegando ao fim, é cerrar os **olhos** e evocar todas as cousas que não ache  
73 em... Seja homem, ande. Pádua enxugou os **olhos** e foi para casa, onde viveu prostrado alg  
74 tava só tinha o pai ao pé de si, enxugando os **olhos** e mirando um triste bilhete de loteria. N  
75 nares padre, vens morar comigo. Enxuguei os **olhos** e o nariz. Ela afagou-me, depois quis re  
76 is de algum tempo de silêncio, recolhendo os **olhos** e um suspiro! --Tem agradado muito est  
77 randura, levantava-me o queixo e espetava os **olhos** em mim, ansioso também, como a prim  
78 o, é a crise que me tomou quando vi todos os **olhos** em mim, os pés quietos, as orelhas ate  
79 la de visitas, falávamos do seminário. Com os **olhos** em mim, Capitu queria saber que notíci  
80 sto, a reflexão; íamos dar com ele, muita vez, **olhos** enfiados em si, cogitando. Respondia-n  
81 outras andorinhas estavam trepadas no ar, os **olhos** enfiados nos olhos, mas tão cautelosos  
82 ; achou até natural e espetou-me outra vez os **olhos**. Então contei-lhe por alto o que podia,  
83 e com uma espécie de vertigem, sem fala, os **olhos** escuros. Quando eles me clarearam vi  
84 es do homem que recebera, defunto, aqueles **olhos**... É impossível que algum Homero não t  
85 os e moços, sedas e chitas, e provavelmente **olhos** feios e belos, mas eu não vi uns nem o  
86 aram cair de cansadas ou de esquecidas. Os **olhos** fitavam-se e desfitavam-se, e depois de  
87 tantes, até que novamente se ergueram, e os **olhos** fixaram-se na parede do pátio, como qu  
88 leos da teologia... Desta vez a fulguração dos **olhos** foi menor, as pálpebras não lhe caíram  
89 ssim perto, me envolvia, me encarava com os **olhos** furados e escuros. Quanto mais andava  
90 u favor, mas a fé velava com os seus grandes **olhos** ingênuos. Minha mãe faria, st pudesse,  
91 a do seu tempo) esquecia o adágio: longe dos **olhos**, longe do coração. Nós não podíamos t

92 xto para mirá-los mais de perto, com os meus **olhos** longos, constantes, enfiados neles, e a  
93 l é ele? perguntou escrevendo a resposta nos **olhos**. --Mamãe, acudi eu. José Dias apertou-  
94 e ia para a Europa. Dito isto, espreitou-me os **olhos**, mas creio que eles não lhe disseram n  
95 stavam trepadas no ar, os olhos enfiados nos **olhos**, mas tão cautelosos que se desenfia  
96 vez, ao dar com ela, não sei se era dos meus **olhos**, mas Capitu pareceu-me lívida. Seguiu-  
97 UXO E O COLORIDO Nem só os **olhos**, mas as restantes feições, a cara, o cor  
98 o que, para desnortear a justiça, os mesmos **olhos** matadores seriam olhos piedosos, e co  
99 se lendo de verdade; creio que era quando os **olhos** me caíam na palavra do fim da página,  
100 uela maneira os seus olhos teimosos. Outros **olhos** me procuravam também, não muitos, e  
101 pitu, a princípio, não disse nada. Recolheu os **olhos**, meteu-os em si e deixou-se estar com  
102 to da moça redobrou tanto que senti os meus **olhos** molhados e fugi. Vim para perto de uma  
103 Mas nada de melancolias; não quero falar dos **olhos** molhados, à entrada e à saída. Pouco e  
104 e era muito chorão por esse tempo, sentia os **olhos** molhados... Era amor puro, era efeito d  
105 E, após alguma reflexão, fitando em mim uns **olhos** murchos e teimosos, perguntou-me: --C  
106 já namorava a outro, acompanhá-lo-ia com os **olhos** na rua, falar-lhe-ia à janela, às ave-mari  
107 ele meu vizinho, fiquei apavorado e desviei os **olhos**. Não sei que mão oculta me compeliu a  
108 dos, e as lágrimas faziam-lhe encarquilhar os **olhos**. Não obstante o total falava e cativava o  
109 mais assídua e terna, vivia ao pé dela, com os **olhos** nela. Minha mãe era de natural simpáti  
110 mente. Calou se alguns instantes; eu tinha os **olhos** nele, ele voltara os seus para o lado da  
111 va o chapéu em cortejar a vizinhança, risonho, **olhos** no ar, antes mesmo da administração i  
112 adre Lopes, oh! o Padre Lopes... Ele, com os **olhos** no ar, devia estar ouvindo, e naturalmen  
113 PITOLINA Voltei-me para ela; Capitu tinha os **olhos** no chão. Ergueu-os logo, devagar, e fic  
114 Como me achasse estirado na cama, com os **olhos** no tecto, lembrou-me a recomendação  
115 rdo; afinal deixei-me estar de costas, com os **olhos** no tecto, mas nem assim. vinha mais n  
116 o unidas. De quando em quando enxugava os **olhos**. O cocheiro aventurou duas ou três perg  
117 mbrado a definição que José Dias dera deles, "**olhos** de cigana oblíqua e dissimulada." Eu n  
118 ue é, Bentinho? Para não fitá-lo, deixei cair os **olhos**. Os olhos, caindo, viram que uma das p  
119 ão tênues que se esgarçam ao menor abrir de **olhos** ou voltar de corpo, e não continuam mai  
120 a abaixou também a cabeça, mas voltando os **olhos** para cima a fim de ver os meus. Fiz-me  
121 do a loteria e Pandora me aborrecem, ergo os **olhos** para eles, e esqueço os bilhetes branco  
122 gostava da ressurreição. Às vezes, fechava os **olhos** para não ver gestos nem nada, mas o d  
123 a a casa achava-a mais alta e mais cheia; os **olhos** pareciam ter outra reflexão, e a boca ou  
124 conheciam-se as ocasiões pelo apertado dos **olhos**. Pediu-me algumas circunstâncias mai  
125 te, Capitu inclinou-se para fora, eu relancei os **olhos** pela rua, estava deserta. Peguei-lhe na  
126 e depressa e não achei compostura; metia os **olhos** pelas cadeiras. Ao contrário, Capitu erg  
127 ão; naturalmente pasmava do intruso. Corri os **olhos** pelo ar, buscando algum pensamento q  
128 ter, e não vi que essa menina travessa e já de **olhos** pensativos era a flor caprichosa de um f  
129 uidar de nós mesmos, cada um com os seus **olhos** perdidos, provavelmente. Os dele estav  
130 justiça, os mesmos olhos matadores seriam **olhos** piedosos, e correriam a chorar a vítima.  
131 o Escobar era um tanto metediço e tinha uns **olhos** policiais a que não escapava nada. --Sã  
132 dedos, não; onde está o lenço? Enxuguei os **olhos**, posto que de todas as palavras de Jos  
133 e, não abria as vidraças, chegava a fechar os **olhos**. Quando novamente abria os olhos e a  
134 ram semelhantes, a testa principalmente e os **olhos**. Quanto ao gênio, era um, pareciam irm  
135 . Se me não engano, chegou a dizê-lo com os **olhos**. Quanto ao marido, tocava agora com  
136 ambém a alguém há de você sair, com esses **olhos** que Deus lhe deu; são exatamente os d  
137 depois, como se faz a um filho de verdade; os **olhos** que ele me deitou foram ternos e agrad  
138 que me dizia, os gestos de lenço, os próprios **olhos** que enxugava eram tais que me comovi  
139 ças o fazem. O que nem todas fazem é ter os **olhos** que esta tinha. Em nenhuma vi as ansi

140 m... Deixa de manha, Bentinho. Creio que os **olhos** que lhe deitei foram tão queixosos que  
141 ferro nem corda, pistola nem punhal- mas os **olhos** que lhe deitei, se pudessem matar, teri  
142 ua não é de todo má. Capitu, apesar daqueles **olhos** que o Diabo lhe deu... Você já reparou  
143 ela? murmurou Capitu. E acrescentou com os **olhos**, que brilhavam extraordinariamente -- S  
144 mal de morte, e Deus pode tudo. Enxugue os **olhos**, que é feio um mocinho da sua idade an  
145 quecer inteiramente a mão de Sancha nem os **olhos** que trocamos. Agora achava-lhes isto,  
146 usava cabeleira grande; o retrato mostra uns **olhos** redondos, que me acompanham para to  
147 quei. --Nem eu digo que sejam de outro. --São **olhos** refletidos, opinou tio Cosme. --Seguram  
148 nguíssimas a idéias brevíssimas. Enxuguei os **olhos**, repito, e fui andando, ansioso agora po  
149 tá comigo. Perfeição não era; ao contrário, os **olhos** saíram esbugalhados, e os cabelos era  
150 e levantei, nem sei se iria. Capitu fitou-me uns **olhos** tão ternos, e a posição os fazia tão súp  
151 tos que me explicou daquela maneira os seus **olhos** teimosos. Outros olhos me procuravam  
152 ativo, a tal ponto ás bochechas, os dentes, os **olhos**, toda a cara, toda a pessoa, todo o mu  
153 . Enquanto ele falava, Capitu dava-me com os **olhos** todas as sortes grandes e pequenas. A  
154 or único. José Dias ficou aturdido. Quando os **olhos** tornaram às dimensões ordinárias: --Ma  
155 rém, que, como eu estava cansado, fechei os **olhos** três ou quatro vezes; tanto bastou para  
156 om de todo. --Estou. Ouvia, espetando-me os **olhos**. Três dias depois disse que me estava  
157 zer a palavra última; nós a líamos, porém, nos **olhos** um do outro, vibrante e decisiva, e sem  
158 1872. --Você já reparou que Ezequiel tem nos **olhos** uma expressão esquisita? perguntou-m  
159 um abraço! outro! mais outro! adeus! Tinha os **olhos** úmidos deveras; levava a cara dos dese  
160 e a voz lhe tremia, e pareceu-me que tinha os **olhos** úmidos. Disse-lhe que também sentia a  
161 ncia na xícara, e comecei a mexer o café, os **olhos** vagos, a memória em Desdêmona inoc  
162 a muito amiga de si. Quando tornou, trazia os **olhos** vermelhos; disse-nos que, ao mirar o fil  
163 razão particular. Tudo isso me empanava os **olhos**... Vieram os jornais do dia: davam notíc  
164 era para ele como nada: cerrava um pouco os **olhos**, voltados para cima, e sussurrava as de

Anexo 2 – Subcorpus literário em português

	Arquivo	Nº palavras
<b>Adolfo Caminha</b>		
Bom-Crioulo	ac01.txt	36.238
A Normalista	ac02.txt	61.100
Tentação	ac03.txt	40.721
<b>Aluísio de Azevedo</b>		
Casa de Pensão	aa01.txt	95.807
O Cortiço	aa02.txt	81.734
Livro de uma Sogra	aa03.txt	53.904
Mattos, Malta ou Matta?	aa04.txt	19.252
A Mortalha de Alzira	aa05.txt	51.439
O Mulato	aa06.txt	98.409
<b>Bernardo Guimarães</b>		
A Escrava Isaura	bg01.txt	53.312
O Garimpeiro	bg02.txt	43.816
<b>Coelho Neto</b>		
A Conquista	cn01.txt	90.636
<b>Euclides da Cunha</b>		
Os Sertões	ec01.txt	158.552
<b>Franklin Távora</b>		
O Cabeleira	ft01.txt	51.891
O Matuto	ft02.txt	77.372
O Sacrifício	ft03.txt	46.466
<b>João do Rio</b>		
A Alma Encantadora das Ruas	jo01.txt	62.541
<b>Joaquim Manuel de Macedo</b>		
A Luneta Mágica	jm01.txt	55.664
A Moreninha	jm02.txt	47.938
As Mulheres de Mantilha	jm03.txt	69.287



As Vítimas-Algozes	jm04.txt	98.867
--------------------	----------	--------

<b>José de Alencar</b>
------------------------

Cinco Minutos	ja01.txt	14.799
Diva	ja02.txt	29.346
Encarnação	ja03.txt	30.174
O Garatuja	ja04.txt	33.047
O Gaúcho	ja05.txt	71.604
O Guarani	ja06.txt	109.050
Iracema	ja07.txt	25.019
Lucíola	ja08.txt	46.019
A Pata da Gazela	ja09.txt	35.200
Senhora	ja10.txt	77.921
Sonhos D'Oro	ja11.txt	73.337
Til	ja12.txt	70.638
Ubirajara	ja13.txt	23.038
A Alma do Lázaro	ja14.txt	12.169
Guerra dos Mascates	ja15.txt	79.011

<b>Júlio Diniz</b>
--------------------

As Pupilas do Senhor Reitor	jd01.txt	95.428
-----------------------------	----------	--------

<b>Júlio Ribeiro</b>
----------------------

A Carne	jr01.txt	49.410
---------	----------	--------

<b>Lima Barreto</b>
---------------------

Os Bruzundangas	lb01.txt	42.184
Clara dos Anjos	lb02.txt	50.030
O Cemitério dos Vivos	lb03.txt	28.144
Histórias e Sonhos	lb04.txt	60.727
O Homem que Sabia Javanês e outros contos	lb05.txt	30.362
Marginália	lb06.txt	66.672
A Nova Califórnia	lb07.txt	39.757
Recordações do Escrivão Isaías Caminha	lb08.txt	61.944
O Subterrâneo do Morro do Castelo	lb09.txt	18.069
Triste Fim de Policarpo Quaresma	lb10.txt	66.737
Aventuras do Dr. Bogoloff	lb11.txt	25.925
Vida Urbana	lb12.txt	68.289

<b>Manuel Antônio de Almeida</b>
----------------------------------

Memórias de um Sargento de Milícias	ma01.txt	60.774
-------------------------------------	----------	--------

<b>Pe. Antônio Vieira</b>
---------------------------

Maria Rosa Mística	pa01.txt	13.827
Sermão II	pa02.txt	17.645
Sermão III	pa03.txt	14.225
Sermão pelo Bom Sucesso das Armas de Portugal contra as de Holanda	pa04.txt	9.120
Sermão dos Bons Anos	pa05.txt	8.839
Sermão de Santo Antônio	pa06.txt	12.135
Sermão da Sexagésima	pa07.txt	11.435
Sermão da Quinta Domingo da Quaresma	pa08.txt	8.744

#### **Raul Pompéia**

O Ateneu	rp01.txt	62.187
As Jóias da Coroa	rp02.txt	28.284
14 de Julho na Roça	rp03.txt	62.102

#### **Simões Lopes Neto**

Lendas do Sul	sn01.txt	18.811
---------------	----------	--------

#### **Visconde de Taunay**

Ao Entardecer	vt01.txt	32.429
Inocência	vt02.txt	51.581

### Anexo 3 – Subcorpus literário em inglês

	Arquivo	Nº palavras
<b>Aldous Huxley</b>		
Crome Yellow	ah01.txt	57.723
Brave New World	ah02.txt	66.294
<b>Bram Stoker</b>		
Dracula	bs01.txt	160.975
<b>Charles Dickens</b>		
A Christmas Carol	cd01.txt	28.850
David Copperfield	cd02.txt	359.124
Great Expectations	cd03.txt	186.603
Hard Times	cd04.txt	104.049
Oliver Twist	cd05.txt	159.017
The Pickwick Papers	cd06.txt	304.535
A Tale of Two Cities	cd07.txt	137.386
<b>Charles Lamb</b>		
Tales from Shakespeare	cl01.txt	100.618
The Adventures of Ulysses	cl02.txt	33.643
<b>Charlotte Brönte</b>		
Jane Eyre	cb01.txt	188.220
The Professor	cb02.txt	89.385
<b>Daniel Defoe</b>		
Robinson Crusoe	dd01.txt	233.022
<b>Edgar Allan Poe</b>		
The Works of Edgar Allan Poe - Volume I	ep01.txt	83.154
The Works of Edgar Allan Poe - Volume II	ep02.txt	96.994
The Works of Edgar Allan Poe - Volume III	ep03.txt	101.212
The Works of Edgar Allan Poe - Volume IV	ep04.txt	81.070
The Works of Edgar Allan Poe - Volume V	ep05.txt	75.025
<b>E.M. Forster</b>		

Howards End	ef01.txt	110.208
The Longest Journey	ef02.txt	95.223
A Room with a View	ef03.txt	67.514
Where Angels Fear to Tread	ef04.txt	50.136

### Emily Brönte

Wuthering Heights	eb01.txt	117.117
-------------------	----------	---------

### George Orwell

Animal Farm	go01.txt	30.227
1984	go02.txt	104.450

### G.K. Chesterton

The Man Who Knew Too Much	gc01.txt	59.893
The Man Who Was Thursday	gc02.txt	58.476

### Henry David Thoreau

Walden	ht01.txt	116.901
On the Duty of Civil Disobedience	ht02.txt	9.408

### Henry James

The Ambassadors	hj01.txt	165.024
The Beast in the Jungle	hj02.txt	18.792
The Turn of the Screw	hj03.txt	42.831
The Portrait of a Lady	hj04.txt	232.881
Roderick Hudson	hj05.txt	132.828
The American	hj06.txt	135.185
Washington Square	hj07.txt	64.422
What Maisie Knew	hj08.txt	95.997
The Golden Bowl	hj09.txt	207.667
The Aspern Papers	hj10.txt	39.151
Daisy Miller	hj11.txt	21.465
The Pupil	hj12.txt	17.569
The Jolly Corner	hj13.txt	14.434
The Europeans	hj14.txt	60.034
Confidence	hj15.txt	76.681
The Reverberator	hj16.txt	55.176
The Awkward Age	hj17.txt	138.857
A Passionate Pilgrim	hj18.txt	27.756
Madame de Mauves	hj19.txt	32.122
A Bundle of Letters	hj20.txt	14.098
The Author of Beltraffio	hj21.txt	19.476
The Real Thing	hj22.txt	10.618

The Altar of the Dead	hj23.txt	14.798
In the Cage	hj24.txt	35.391

<b>Herman Melville</b>
------------------------

Moby Dick	hm01.txt	216.106
Bartleby, the Scrivener	hm02.txt	14.573

<b>H.G. Wells</b>
-------------------

The Invisible Man	hg01.txt	49.166
The Island of Doctor Moreau	hg02.txt	44.248
The Time Machine	hg03.txt	32.654
The War of the Worlds	hg04.txt	60.519

<b>Jack London</b>
--------------------

The Call of the Wild	jl01.txt	32.125
White Fang	jl02.txt	72.926

<b>James Joyce</b>
--------------------

A Portrait of the Artist as a Young Man	jj01.txt	85.155
Dubliners	jj02.txt	67.983
Ulysses	jj03.txt	264.715

<b>Jane Austen</b>
--------------------

Emma	ja01.txt	160.997
Mansfield Park	ja02.txt	160.461
Persuasion	ja03.txt	83.659
Pride and Prejudice	ja04.txt	122.194
Sense and Sensibility	ja05.txt	119.965

<b>Jonathan Swift</b>
-----------------------

Gulliver's Travels	js01.txt	104.636
A Modest Proposal	js02.txt	3.418

<b>Joseph Conrad</b>
----------------------

Heart of Darkness	jc01.txt	38.765
Lord Jim	jc02.txt	131.098
Nostromo	jc03.txt	171.960

<b>Lewis Carroll</b>
----------------------

Alice's Adventures in Wonderland	lc01.txt	26.688
----------------------------------	----------	--------

Through the Looking-Glass	lc02.txt	29.662
---------------------------	----------	--------

<b>Mark Twain</b>
-------------------

Huckleberry Finn	mt01.txt	112.150
The Adventures of Tom Sawyer	mt02.txt	71.359
A Connecticut Yankee in King Arthur's Court	mt03.txt	119.241
The Prince and the Pauper	mt04.txt	70.625

<b>Mary Shelley</b>
---------------------

Frankenstein	ms01.txt	75.176
--------------	----------	--------

<b>Nathaniel Hawthorne</b>
----------------------------

Twice-Told Tales	nh01.txt	148.333
The Scarlet Letter	nh02.txt	85.055
The House of the Seven Gables	nh03.txt	106.943
The Blithedale Romance	nh04.txt	76.885
The Marble Faun	nh05.txt	140.704

<b>Oscar Wilde</b>
--------------------

The Canterville Ghost	ow01.txt	11.743
De Profundis	ow02.txt	17.783
The Happy Prince and Other Tales	ow03.txt	16.321
The Picture of Dorian Gray	ow04.txt	79.695

<b>Robert Louis Stevenson</b>
-------------------------------

The Strange Case of Dr. Jekyll and Mr. Hyde	rs01.txt	25.906
Treasure Island	rs02.txt	69.128

<b>Rudyard Kipling</b>
------------------------

The Jungle Book	rk01.txt	51.742
The Second Jungle Book	rk02.txt	65.276

<b>Scott Fitzgerald</b>
-------------------------

The Beautiful and Damned	sf01.txt	125.962
This Side of Paradise	sf02.txt	82.613
The Great Gatsby	sf03.txt	50.139

<b>Sir Walter Scott</b>
-------------------------

Ivanhoe	ws01.txt	194.763
---------	----------	---------

**Thomas Hardy**

Far from the Madding Crowd	th01.txt	138.031
Tess of the D'Urbervilles	th02.txt	151.163
Under the Greenwood Tree	th03.txt	58.812
The Return of the Native	th04.txt	143.653
The Mayor of Casterbridge	th05.txt	117.671
The Woodlanders	th06.txt	137.199
Life's Little Ironies	th07.txt	78.764
Jude the Obscure	th08.txt	146.489
A Group of Noble Dames	th09.txt	69.703
Wessex Tales	th10.txt	81.922
A Pair of Blue Eyes	th11.txt	132.071
The Trumpet-Major	th12.txt	115.638
Two on a Tower	th13.txt	95.261

**Virginia Woolf**

Jacob's Room	vw01.txt	55.430
Night and Day	vw02.txt	167.737
The Voyage Out	vw03.txt	137.514
Mrs. Dalloway	vw04.txt	64.769
To the Lighthouse	vw05.txt	69.858
Flush: a Biography	vw06.txt	34.692
The Waves	vw07.txt	78.223

This document was created with Win2PDF available at <http://www.win2pdf.com>.  
The unregistered version of Win2PDF is for evaluation or non-commercial use only.  
This page will not be added after purchasing Win2PDF.