

# ANÁLISE COMPARATIVA DAS EDIÇÕES PORTUGUESA E BRASILEIRA DA OBRA OS LIVROS QUE DEVORARAM O MEU PAI, DE AFONSO CRUZ

Ida Rebelo-Arnold<sup>1</sup>; Anabela Barreiro<sup>2</sup>; Tanara Zingano Kuhn<sup>3</sup>;

Isabel Garcez<sup>4</sup>; Fernando Batista<sup>5</sup>

## RESUMO

Neste trabalho apresentamos uma análise comparativa das edições portuguesa (versão original) e brasileira (versão adaptada) da obra de literatura infantojuvenil *Os Livros que devoraram o meu pai*, do autor português Afonso Cruz. Esta análise tem como objetivo contribuir para a otimização dos processos editoriais necessariamente presentes na adaptação de textos, mas que se adequam a qualquer tipo de processo editorial. Para tal, partimos de alinhamentos ao nível da frase da obra completa para realizar alinhamentos ao nível da unidade lexical multipalavra ou da expressão usando a ferramenta CLUE-Aligner, que permite registar numa base de dados todos os pares de unidades parafrásticas resultantes da tarefa de alinhamento. Focamo-nos essencialmente na comparação de construções com função adjetival e esta análise comparativa pretende verificar que tipos de alterações foram realizadas no processo de adaptação. A partir do estudo dos resultados contrastivos baseados nos pares alinhados, que, na sua maioria, correspondem a unidades parafrásticas, discutimos as implicações das modificações de ordem linguística na constituição do novo texto, em termos semânticos, pontualmente analisados também do ponto de vista literário e/ou cultural. Como forma de manter a qualidade de recepção do texto de chegada, propomos uma tomada de consciência face aos limites impostos por um texto literário, pois é ténue a fronteira entre a adaptação indispensável e a intervenção excessiva. Este estudo apresenta uma base científica para trabalhos futuros na área da edição, revisão e conversão de texto literário de e para qualquer variedade do português.

## INTRODUÇÃO

Os processos de revisão editorial, nos últimos anos, têm vindo a beneficiar de reflexões e orientações da linguística, enquanto ciência da linguagem, mas também de ferramentas de

---

<sup>1</sup> Universidad de Valladolid – [imdarnot@funge.uva.es](mailto:imdarnot@funge.uva.es).

<sup>2</sup> INESC-ID Lisboa – [anabela.barreiro@inesc-id.pt](mailto:anabela.barreiro@inesc-id.pt).

<sup>3</sup> Universidade de Coimbra – [tanarazingano@uc.pt](mailto:tanarazingano@uc.pt).

<sup>4</sup> Universidade de Lisboa – [isabelgarcez@campus.ul.pt](mailto:isabelgarcez@campus.ul.pt).

<sup>5</sup> Instituto Universitário de Lisboa (ISCTE-IUL) e INESC-ID Lisboa – [fernando.batista@inesc-id.pt](mailto:fernando.batista@inesc-id.pt).

processamento de linguagem natural ou linguística computacional, que podem servir para desenvolver tarefas de análise de corpora, geração e sumarização de textos, tradução, parafraseamento, entre outros.

Na tentativa de complementar ferramentas que auxiliam na correção ortográfica e gramatical e até em sistemas de escrita, Barreiro (2009) criou um sistema de geração de paráfrases, o eSPERTo<sup>6</sup>, que visa auxiliar em tarefas variadas, como a sumarização e a resposta a perguntas (cf. MOTA et al., 2016), o ensino da língua, o apoio à tradução automática (BARREIRO, 2009); a escrita técnica controlada, a edição e revisão de textos e a adaptação estilística (BARREIRO et al., 2018), entre outras. Neste caso de estudo, o uso da ferramenta permitiu-nos analisar as unidades parafrásticas de uma obra literária completa nas variedades do português europeu (PE) e do português do Brasil (PB), no seguimento de alguns trabalhos experimentais anteriormente realizados de adaptação entre estas variantes, dos quais destacamos os mais recentes (BARREIRO e MOTA, 2018; REBELO-ARNOLD et al., 2018) e que traçam um rumo mais global, dinâmico e de respeito pelas divergências e convergências que fazem da língua portuguesa uma das línguas mais ricas do planeta (SANTOS, 2016).

O presente estudo enquadra-se num projeto mais amplo, que pretende propor linhas de análise comparativa de edições brasileiras de obras literárias de autores portugueses e de edições portuguesas de obras literárias de autores brasileiros. Defendemos que o desenvolvimento do estudo comparativo destas edições poderá desempenhar um papel importante numa melhor e mais profícua relação transatlântica no que concerne à criação literária em língua portuguesa. Num futuro cada vez mais próximo, será pertinente fomentar o desenvolvimento de ferramentas que possam vir a auxiliar editores, revisores e professores, assim como estudantes e leitores dos dois países, seja do ponto de vista da mediação ou no da recepção literárias.

O principal objetivo deste trabalho é analisar a equivalência dos pares parafrásticos com funções adjetivais das edições PE e PB da obra de literatura infantojuvenil *Os Livros que devoraram o meu pai*, do autor português Afonso Cruz. Tal como defendido por Garcez (2017), a edição literária tem responsabilidades acrescidas que se substanciam no equilíbrio entre o respeito pela intenção autoral e o respeito pela qualidade de recepção que o texto do autor possibilita.

---

<sup>6</sup> eSPERTo é um acrónimo de Sistema de Parafraseamento para Edição e Revisão de Texto, que funciona tanto em português quanto em inglês, System for Paraphrasing in Editing and Revision of Text.

Do ponto de vista linguístico, o objetivo específico é identificar, analisar e discutir uma seleção de unidades lexicais simples, multpalavras e estruturas fráscas com função adjetival, que estão em alternância nas duas variedades – ou seja, que apresentam diferenças entre as escolhas do autor na variedade de PE e as correspondentes soluções adotadas na conversão para PB. A metodologia escolhida centra-se na análise linguística contrastiva com base na Gramática Funcionalista (NEVES, 1997), posta em prática com o auxílio de ferramentas computacionais de tratamento de corpora baseadas no projeto eSPERTO e com recurso a alinhamentos semiautomáticos usando o alinhador CLUE-Aligner (BARREIRO et al., 2016).<sup>7</sup> Os dois subcorpora do estudo são as edições portuguesa e brasileira da obra.

Do ponto de vista literário, e partindo da mesma seleção linguística, analisaremos a validade das escolhas editoriais patentes na edição em PB em duas frentes: 1) fundamentação das escolhas do autor presentes na edição em PE segundo a sua polissemia, estranheza intencional, riqueza rítmica e rede de relações de significado no contexto em que surgem e com a obra no seu todo; 2) avaliação da equivalência qualitativa na recepção de ambas as edições quanto aos mesmos parâmetros enunciados em 1), tendo em consideração que, tanto a edição em PE como a edição em PB são passíveis de serem trabalhadas em contexto escolar por um público-alvo pertencente ao mesmo escalão etário.

Partindo dos resultados alcançados, pretendemos ainda a) incentivar a investigação de recursos linguísticos para os propósitos de mediação da literatura em Português, língua materna e/ou língua estrangeira; b) questionar a importância da mediação literária, seja ela personalizada pelo editor, o revisor ou o professor, no que respeita à recepção literária; c) apresentar um leque de alternativas editoriais e de mediação da leitura que permitam preservar o máximo do texto da edição original e, desta forma, d) aumentar e enriquecer o nível de conhecimento linguístico, literário e cultural entre os leitores de ambas as variedades do Português aqui em estudo.

## **CORPUS DE TRABALHO**

O corpus de trabalho é composto pelo texto integral de duas edições de *Os Livros que devoraram o meu pai*, nomeadamente, a original em PE e sua versão em PB. As razões que levaram à escolha deste corpus prendem-se com: 1) tratar-se de uma obra dirigida ao público

---

<sup>7</sup>Outra ferramenta que podemos aqui destacar para análise de corpora é o Sketch Engine (KILGARRIFF et al., 2004). Trata-se de um recurso online que permite, entre outras coisas, a realização de análises léxico-gramaticais avançadas em corpora paralelos. Em alguns casos, esta ferramenta também foi utilizada como ferramenta de apoio à análise de unidades lexicais em contexto porque é uma ferramenta onde facilmente podemos emparelhar frases e que é de utilização pública e gratuita.

infantojuvenil, e por essa razão poder-se-á defender que a adaptação do texto obriga a um tratamento mais incisivo, pois o entendimento do texto, na versão adaptada, deve ser equivalente ao do público-alvo do texto de partida; 2) esta obra ter recebido o prémio literário Maria Rosa Colaço 2009 (Portugal); 3) estar incluída nas listas do Plano Nacional de Leitura português, recomendada para leitura autónoma do 3.º ciclo (12-14 anos); 4) integrar as listas do Programa Nacional Biblioteca na Escola 2013 do Brasil, para os anos finais do Ensino Fundamental; 5) o texto da edição brasileira apresentar diferenças significativas em relação ao da edição portuguesa.

## METODOLOGIA

O corpus paralelo correspondente à obra analisada contém 1260 pares de frases. As frases são apresentadas no CLUE-Aligner em forma de grelha, em que a frase em PE aparece na vertical formando várias colunas e a frase em PB aparece na horizontal formando várias linhas. O cruzamento das palavras nas colunas e nas linhas forma uma matriz em que cada palavra corresponde a uma célula. A tarefa de alinhamento corresponde ao delineamento das fronteiras de mapeamento, que respeitam critérios de análise linguística. A Figura 1 apresenta o mapeamento que corresponde ao alinhamento das unidades parafrásticas das frases no exemplo (1).

- (1) **O meu pai** só pensava em livros (livros e mais livros!), mas a vida não era da mesma opinião, **a vida dele pensava noutras coisas**, andava distraída, e **ele teve de se empregar**.  
**Meu pai** só pensava em livros (livros e mais livros!), mas a vida não era da mesma opinião, **a dele pensava em outras coisas**, andava distraída, e **ele teve de arrumar um emprego**.

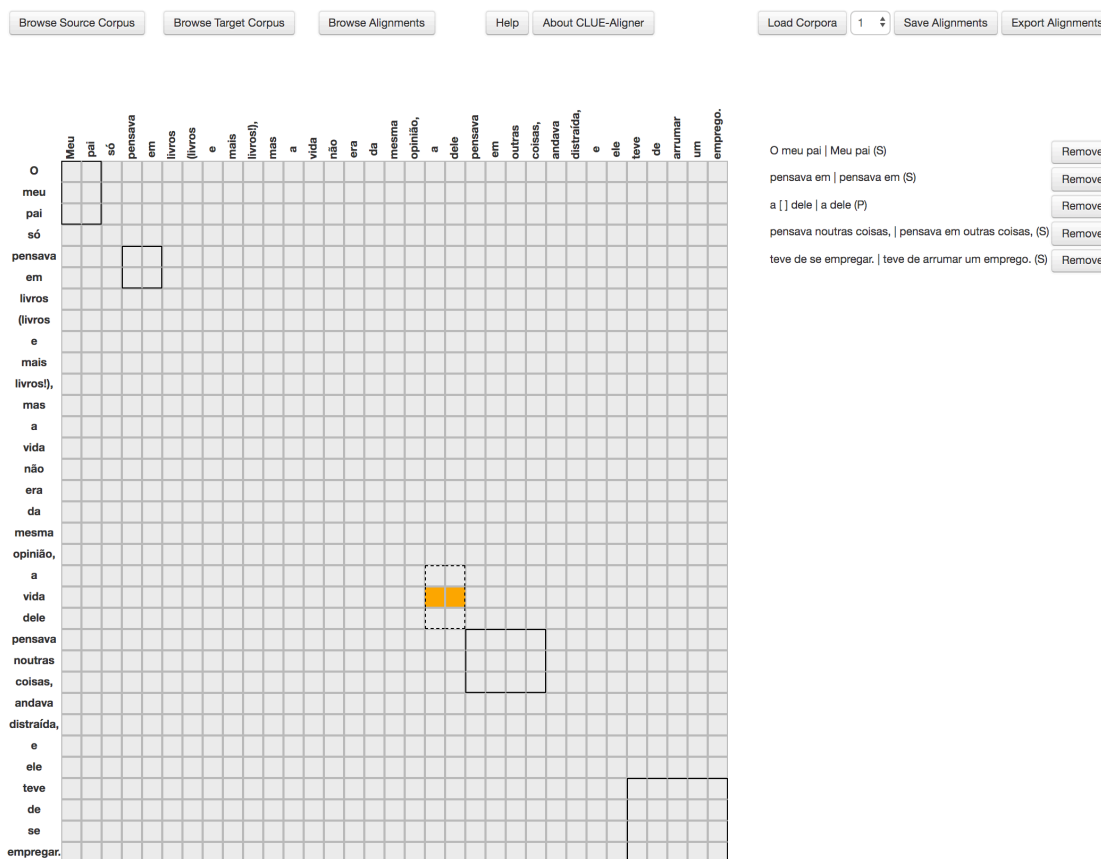
No alinhamento das frases representadas na Figura 1, interessa destacar que o alinhador CLUE-Aligner registou os alinhamentos feitos manualmente pelos anotadores linguísticos. Encontramos 5 alinhamentos, mas destes apenas 4 diferem entre a versão em PE e a versão em PB. A razão pela qual o anotador decidiu anotar um alinhamento de unidades lexicais multipalavra fundamenta-se pura e simplesmente na utilidade que esse par possa ter em trabalho futuro<sup>8</sup>, não sendo propriamente relevante para o presente estudo. As 4 unidades contrastivas apresentam: (i) **diferenças varietais**, tais como o uso de artigo definido antes do pronome possessivo em PE e a ausência desse artigo em PB (*O meu pai* > *Meu pai*), um

---

<sup>8</sup> Interessa colecionar um conjunto de unidades lexicais multipalavra em que surjam relações de dependência ou seleção de preposições, redes argumentais, etc. Dessa forma, o exemplo “pensava em” é um exemplo que interessa registar por indicar que o verbo “pensar” seleciona a preposição “em”, tanto em PE quanto em PB.

fenómeno recorrente ao longo de toda a obra e reconhecido há muito pelos linguistas teóricos como um fator comum de distinção entre o PE e o PB; ou a tendência de contrair preposições com artigos em PE e não em PB, como no caso de “*pensava noutras coisas*” versus “*pensava em outras coisas*”; e o uso de expressões que são comuns numa variedade do português e não na outra, como acontece no par de unidades parafrásticas “*teve de se empregar*” e “*teve de arrumar um emprego*”, que na variedade do PE não se usa (a correspondente mais direta seria “*teve de arranjar um emprego*”); (2) **diferenças estilísticas**, tal como a opção na versão PB de retirar a repetição da palavra “*vida*” na expressão “*a vida dele*” presente na versão PE, o que oblitera um recurso expressivo aparentemente deliberado de dar relevo a duas acepções de “*vida*”, a geral e a da personagem, e/ou para produzir um determinado efeito sonoro que cause impacto no leitor.

**Figura 1:** Alinhamentos de unidades parafrásticas internas das frases do Exemplo (1)



Apesar de destacarmos, neste artigo, a análise das estruturas adjetivais encontradas nos pares parafrásticos, todas as diferenças foram registadas nos alinhamentos (diferenças estilísticas, varietais, etc.), como pode ser visto na Tabela 1 da Secção **Resultados**

**Quantitativos.** Na secção que se segue, discutiremos alguns dos pares que se apresentam mais pertinentes.

## ANÁLISE E CATEGORIZAÇÃO DAS PARÁFRASES

Nesta Secção, apresentamos e discutimos brevemente uma seleção de unidades lexicais contrastivas, ou seja, as escolhas do autor na edição original em PE contrastadas com as substituições na versão em PB. As unidades escolhidas, dentre os vários fenómenos verificados na comparação das paráfrases, compõem grupos mais ou menos homogêneos, de aspetos distintivos entre as duas variedades do Português. Usamos termos vindos da gramática funcional e da gramática tradicional para formalizar uma tentativa de categorias:

### 1) Alternância de preposições em locuções adjetivais – SUBST/ADJ + PREP VAR

- (2) bom a decorar | bom em decorar
- (3) cadeira às riscas | poltrona de riscas
- (4) cheiro a canos, a esgoto | cheiro de cano, de esgoto
- (5) mundo a duas dimensões | mundo em duas dimensões

Neste grupo de que fazem parte os exemplos (2)-(5), alterna-se a preposição de uma variedade para a outra. Em (2), existe uma alternância contrastiva regular. Em PE a preposição tem de ser obrigatoriamente “a” na expressão “*ser bom a [fazer alguma coisa]*” (e.g., *ser bom a decorar, ser bom a memorizar, ser bom a nadar*, etc., mas, o uso de um nome predicativo, em vez de um verbo, obrigaria a que a preposição passasse a ser “em”, tal como “*é bom na nataçãõ*” (quando se tratar de transformar um verbo intransitivo) ou “*é bom na memorizaçãõ de datas históricas*”, em que a ocorrência de um complemento do predicado nominal faz com que a expressão seja igualmente correta como em “*é bom a memorizar datas históricas*”. Em PB, não se seleciona a preposição “a” antes do infinitivo. E a expressão em uso é mesmo “*bom em*”, ainda que possa haver variação regional com outras preposições como “*de*” e “*para*”.

### 2) Sintagma de núcleo verbal ou adverbial contrastando com sintagma de núcleo adjetival – SV/SADV vs SAdj

- (6) vê-se | está aparente
- (7) que se aconselhe | aconselhável
- (8) estava [ ] cheio de pó | estava [ ] empoeirado
- (9) está bem para | é adequado para
- (10) tem de estar a horas na mesa | precisa ser pontual à mesa

No grupo representado pelos exemplos (6)-(10), o verbo envolvido é provavelmente determinante para a seleção de uma ou outra classe gramatical. O que o caracteriza, portanto,

é a mudança de classe gramatical que ocorre com a seleção de outro item lexical na paráfrase correspondente. O sentido mantém-se pela seleção de um elemento com a mesma etimologia ou originado dentro do mesmo campo semântico. Em alguns casos, basta uma única palavra para contrastar com toda a expressão, como no exemplo (7) (*que se aconselhe* | *aconselhável*).

### 3) Alternância de elementos lexicais

A alternância ocorre por razões de frequência de uso, preferência, ou outras razões. Distribuímos as alternâncias encontradas em grupos de regularidades, baseados na classificação estabelecida por Wittmann et al. (1995).

#### A. Usos criativos

Dentre os vários traços semânticos que compõem um item lexical e que este partilha com o significado da construção analisada, o autor do original em PE escolhe o menos frequente dos adjetivos, como em (11).

(11) palavras bicudas | palavras ásperas

Há estranhamento por parte do leitor na medida em que dizemos que uma *pessoa é/está bicuda*, ou *um problema é bicudo*. Ao colocar este adjetivo junto do substantivo “*palavras*”, mantém-se o traço semântico que indica dificuldade de aproximação, desagradável, complicado, mas o estranhamento surge da sua original colocação junto ao substantivo “*palavras*”, quando o adjetivo comumente utilizado para “*palavras*”, com idêntico traço semântico seria “*ásperas*”, tal como consta na versão PB. Por outro lado, há ainda a considerar a riqueza de sentido que o adjetivo “*bicudas*” acrescenta ao vocábulo “*palavras*”, dado que, em toda a obra, as palavras, assim como os livros são frequentemente antropomorfizados (conforme, aliás, o próprio título da obra ilustra), o que faz do adjetivo “*bicudas*” uma escolha coerente e aparentemente intencional.

(12) voz amarrotada | voz enrouquecida

O exemplo (12) apresenta outra colocação pouco frequente em Português que gera estranhamento e surpresa, ainda que não impeça a conexão imediata com a expansão do significado. Os termos escolhidos em cada variedade partilham traços de significado com “*cansada*”, vulgarmente usada. Em PB, temos um sentido literal, afastando-se do uso lúdico das palavras que caracteriza a literatura infantojuvenil e é bastante explorado pelo autor do original em PE. Há ainda a mencionar que “*voz amarrotada*” faz parte do título do respetivo capítulo – “*Por vezes a voz dela fica um pouco amarrotada*” (PE) vs. “*Por vezes, a voz dela*

*fica um pouco enrouquecida*” (PB). Esta escolha é condicente com outras escolhas aproximadas para caracterizar a avó do protagonista narrador, e, no mesmo parágrafo (PE, p. 18), temos, por exemplo, a expressão “*mão engelhada*”. Significativamente, temos ainda o termo “*amarrotada*” na frase “*Por vezes a voz dela fica um pouco amarrotada, quando se senta [...]*” (PE, p. 18), como se a voz fosse uma peça de roupa que se amarrota quando nos sentamos. Este sentido perde-se na frase correspondente da edição PB – “*Por vezes, a voz dela fica um pouco enrouquecida, quando se senta [...]*” (PB, p. 16).

(13) Olhos esquecidos | Olhos absortos

No exemplo (13), a entidade nomeada que costuma ter a qualidade atribuída pelo adjetivo ou exercer a ação contida no correspondente verbo “*esquecer*” tem traço +animada, ou se trata de um objeto que sofre a ação do esquecimento. A entidade “*olhos*” não costuma deter essa qualidade e sim o dono dos olhos. O olhar (de alguém) pode denotar esquecimento, mas não são os olhos que são esquecidos. Neste caso, em PB, essa entidade pode ser objeto de metonímia e representar o dono desse olhar, sendo este, sim, “*absorto*”. A construção em PB opta, portanto, por uma colocação bastante frequente, deixando de lado o aspeto lúdico do uso do adjetivo na versão original da obra. Neste caso, devemos novamente apelar ao contexto em que esta construção existe, como forma de compreender o uso intencional de “*esquecidos*” na versão em PE, pois, no mesmo parágrafo, é mencionado que a avó “*com os seus olhos esquecidos*” (PE, p. 18) se esqueceu do presente de aniversário do neto; e, com a utilização de “*olhos absortos*” (PB, p. 16), esta reiteração de sentido está ausente.

## B. Construções com verbo suporte

Há construções onde a definição de verbo suporte parece aplicar-se somente à paráfrase de uma das variedades, como é o caso representado em (14). Isto ocorre, também, devido às escolhas criativas do autor.

(14) ficar amarrotada | ficar enrouquecida

Ou seja, o que caracteriza as construções com verbo suporte é a possibilidade de se proceder à transformação da estrutura em um verbo, como em “*ficar enrouquecida | enrouquecer*” e, no caso da expressão usada em PE, isso não é possível (cf. (15)).

(15) A voz ficou amarrotada | \*A voz amarrotou

## C. Desvios semânticos ou lexicais

### C1. Alteração de sentido



A alteração de sentido por meio de desvios semânticos ou lexicais trata de uma adaptação que modifica o significado da frase e na qual se perde o sentido do texto original, como ilustra o Exemplo (16).

- (16) Na verdade, uma árvore é um pássaro falhado.  
Na verdade, uma árvore é um pássaro defeituoso.

O que se entende em PE por *um pássaro falhado* não é o mesmo que *um pássaro defeituoso*. Diríamos, antes, em PB, *um pássaro que não deu certo*. Em PE, a alternativa comum poderia ser “*ato falhado*”, o que é menos apelativo do que “*pássaro falhado*”. De qualquer modo, a solução “*defeituoso*”, encontrada para a versão do PB, é menos estimulante e inibe o jogo de palavras suscitado pela versão do PE, nomeadamente ao fazer a associação entre a folhagem de uma árvore e as asas de um pássaro ou entre a sedentarização forçada de um e a liberdade de movimentos de outro.

## C.2 Variação estilística

A variação estilística é um recurso expressivo nem sempre necessário no processo de adaptação. No exemplo (17), ambas as construções (tanto a da versão original, como a da versão adaptada) são usadas em PB sem que haja uma discrepância de uso que justifique a substituição da construção existente no original.

- (17) capaz de ser | com capacidade para ser

No entanto, na versão de PB houve a decisão deliberada de substituir a construção adjetival “capaz de” pela construção nominal “com capacidade para”. Das várias ocorrências da expressão “capaz de”, a substituição da forma adjetival pela forma substantiva só é realizada uma vez em que o sujeito é não humano, neste caso, “cão”, o que pode sugerir que seja essa a razão da substituição de uma expressão por outra ainda que ambas sejam amplamente usadas em PB, mas este sugestionamento é impossível de ser comprovado.

## C.3 Adaptação com fins elucidativos

As adaptações com fins elucidativas são as que podem fazer mais sentido, pois a versão em PE se torna hermética para o leitor em PB, como no exemplo (18).

- (18) não era nenhum cão | não era nenhuma maravilha

Todavia, há a considerar que esta frase é igualmente hermética para um leitor em PE e é apenas justificada porque, poucos parágrafos antes (PE, p. 72) é-nos dito que “*A minha memória não é propriamente fiel*” (“*fiel como um cão*”, poder-se-á extrapolar). Mas esta

extrapolação não seria possível na versão do PB porque também esta frase foi alterada, ficando “*Minha memória não é propriamente confiável*” (PB, p. 63).

Neste caso, temos, também, substituições que abandonam a escolha do autor do original pelo estranhamento que suscitam, no original, a suposta eliminação de termos integrantes de uma expressão idiomática ou a introdução de termos não frequentes em expressões semifixas. Na adaptação em PB, verifica-se a opção por uma expressão fixa ou por uma colocação frequente facilmente reconhecíveis pelo leitor, como ilustram os exemplos (19) e (20).

- (19) personagens de carne | personagens de carne e osso
- (20) consciência lavada | consciência limpa

Sendo, todavia, de realçar que também em PE as expressões “normais” seriam as que constam na edição em PB, o que torna as escolhas autorais conscientes da opacidade que estão a introduzir ao alterá-las, mesmo que não haja um motivo óbvio para tal.

#### C.4 Adaptação varietal *stricto sensu*

Classificamos como adaptação varietal as ocorrências em que um termo pouco frequente em PB é substituído por outro de uso corrente, ou em que o léxico é mesmo diferente numa e noutra variedade (e.g. *assistente de bordo* vs *aeromoça*). No exemplo (21), encontramos um verbo frequente em PE, como é o caso do verbo *pôr*, mas pouco usado em PB. A forma de particípio passado do verbo, *postas*, que é também um adjetivo, é igualmente inusual em PB e encontra-se substituída na versão adaptada pelo adjetivo participial do verbo mais frequente nessa variedade.

- (21) crônicas postas em papel | crônicas colocadas em papel

Nos exemplos (22)-(24), a seguir, encontramos gradações no contraste varietal. Em (a), ambos os termos existem em PB, mas o termo usado em PE tem um uso distinto e mais restrito em PB; em (b), o termo selecionado em PE está em desuso em PB e, em (c), o termo selecionado em PE tem frequência baixíssima em PB.

- (22) era mais ou menos vulgar<sup>7</sup>. | era mais ou menos comum.
- (23) A tua mãe já deve estar arrelhada | Sua mãe já deve estar zangada
- (24) Ela olhou para mim como se eu fosse parvo. | Ela olhou para mim como se eu fosse idiota.

## RESULTADOS QUANTITATIVOS

O trabalho realizado envolveu a análise contrastiva de 1261 frases, tendo sido encontradas diferenças entre PE e PB em cerca de 54% das frases (685). Cada frase pode, no entanto, conter vários alinhamentos, os quais foram individualmente contabilizados e analisados,

resultando em 1199 alinhamentos diferentes e ainda assim 280 alinhamentos iguais.<sup>9</sup> Alguns destes alinhamentos correspondem a unidades lexicais não contíguas (BARREIRO e BATISTA, 2016), ilustrados na Tabela 2. Por exemplo, o par parafrástico *nada [ ] correu bem | nada [ ] deu certo* foi alinhado sem as inserções. Esta opção, além de lógica/coerente do ponto de vista linguístico— i.e., registámos apenas a expressão –, tem vantagens para a futura utilização dos alinhamentos na criação de regras transformacionais ou gramáticas de geração de paráfrases em que a mesma gramática dê cobertura a todos os casos em que haja alteração do tempo e número do verbo, por exemplo, e em que as inserções sejam palavras ou expressões diferentes (e.g., *nada [do que eles fazem] corre bem > nada [do que eles fazem/produzem/etc.] dá certo*). Os alinhamentos de expressões não contíguas servem também para registar os casos em que houve uma alteração sintática na versão adaptada, como, por exemplo, em *a [ ] dele > a dele*, em que a versão PB não tem nenhuma palavra entre o determinante e o pronome possessivo porque se decidiu retirar a segunda ocorrência da palavra “*vida*” (cf. exemplo 1).

**Tabela 1:** Exemplos de unidades lexicais não contíguas

PE	PB
a [ ] dele	a dele
consideração [ ] por	consideração [ ] por
mais [ ] do que	mais [ ] do que
É [ ] tão simples quanto	É [ ] tão simples quanto
Soube pela [ ] avó	Soube pela [ ] avó
voz [ ] amarrotada	voz [ ] enrouquecida
nada [ ] correu bem	nada [ ] deu certo
Começou a aplicar [ ] tempo a	Começou a dedicar [ ] tempo a
atirou [ ] para	atirou [ ] na

Com base nos alinhamentos finais foram identificados os principais padrões, que se apresentam na Tabela 1, em que a letra “X” corresponde a uma sequência de palavras igual que ocorre antes da diferente (assinalada) e a letra “Y” corresponde a uma sequência de palavras igual que ocorre depois da diferente (assinalada).

<sup>9</sup> Os alinhamentos iguais podem ter emergido por vários motivos, sendo um deles uma característica da própria ferramenta de alinhamento, CLUE-Aligner, que regista automaticamente as inserções das unidades lexicais multipalavra como pares na base de dados dos alinhamentos parafrásticos, i.e., sempre que um anotador acha que uma inserção deve ser retirada de um determinado alinhamento, essa inserção fica registada e daí o motivo de haver um número considerável de inserções iguais em PE e PB. Outro motivo foi o interesse que a unidade lexical suscitou do ponto de vista linguístico, ou por ser uma palavra composta (mais ou menos usada numa ou noutra variedade do português), como por exemplo a palavra composta “*pastéis de nata*”, ou por corresponder a um nome ou verbo preposicional que possa ser útil guardar, etc.

**Tabela 2:** Padrões de alinhamentos mais frequentes na obra completa

Freq.	PE	PB	Exemplo
87	o Y	Y	<b>o meu</b> coração => <b>meu</b> coração
47	a Y	Y	<b>a sua</b> sabedoria => <b>sua</b> sabedoria
18	X da Y	X de Y	<b>da</b> minha avó => <b>de</b> minha avó
14	as Y	Y	<b>As nossas</b> memórias => <b>Nossas</b> memórias
13	os Y	Y	<b>Os seus</b> cabelos negros => <b>Seus</b> cabelos negros
12	dum Y	de um Y	<b>dum</b> mundo => <b>de um</b> mundo
9	Mr Y	Sr Y	<b>Mr</b> Prendick => <b>Sr</b> Prendick
8	X o Y	X Y	para jantar => para <b>o</b> jantar
6	X de Y	X Y	Preciso <b>de</b> saber => Preciso saber riram <b>de</b> contentes => riram contentes
6	num Y	em um Y	<b>num</b> ambiente acolhedor => <b>em um</b> ambiente acolhedor
5	do Y	de Y	<b>do</b> seu reduto literário => <b>de</b> seu reduto literário
4	X Y	X de Y	me chama monstro => me chama <b>de</b> monstro à espera que => à espera <b>de</b> que
4	X das Y	X de Y	o fim <b>das</b> suas vidas => o fim <b>de</b> suas vidas
4	umas Y	algumas Y	<b>umas</b> semanas antes de partir => <b>algumas</b> semanas antes de partir
4	numa Y	em uma Y	<b>numa</b> ponta => <b>em uma</b> ponta
4	das	de	<b>das</b> suas preocupações => <b>de</b> suas preocupações
4	chamo-me	meu nome é	<b>Chamo-me</b> => <b>Meu nome é</b>
3	X é que	X	Onde <b>é que</b> estão => Onde estão
3	X livro	X exemplar	Folheei o <b>livro</b> => Folheei o <b>exemplar</b>
3	X dum	X de um	morreu <b>dum</b> enfarte => morreu <b>de</b> enfarte
3	X atrás de Y	X após Y	letras <b>atrás de</b> letras => letras <b>após</b> letras
3	X a Y	X Y	tenho <b>a</b> certeza disso => tenho certeza disso
3	um Y	o Y	<b>um</b> cheiro de móveis velhos => <b>o</b> cheiro de móveis velhos um tal => o tal
3	ladra	latia	<b>ladra-lhes</b> => <b>latia</b> para elas
3	dos	de	<b>dos</b> seus => <b>de</b> seus
3	disse eu	eu disse	<b>disse eu</b> => <b>eu disse</b>
3	das Y	de Y	<b>das</b> suas rugas => <b>de</b> suas rugas
3	a tua Y	sua Y	<b>A tua</b> mãe => <b>Sua</b> mãe

Nos padrões exibidos na Tabela 2, predominam os que dizem respeito ao contraste entre PE e PB relativamente ao uso do pronome possessivo seguido de um substantivo. Predominantemente, em PE o pronome é acompanhado (precedido) de um artigo definido ou outro determinante e em PB raramente há ocorrências com um determinante antes do pronome possessivo. Este fenómeno predomina também quando o sintagma nominal é precedido por uma preposição (e.g., *das suas vidas* > *de suas vidas*). Outro fenómeno também recorrente é o da contração da preposição com o artigo em PE e a não contração dos dois elementos em PB (e.g., *num ambiente* > *em um ambiente*). Podemos dizer que estes casos são verdadeiros contrastes varietais. Outros, porém, não passam de simples opções estilísticas (*livro* > *exemplar*; *ladra* > *latia*).

Além dos padrões apresentados na Tabela 2, surgem também exemplos mais esporádicos que, na maioria dos casos, correspondem também a opções estilísticas (*as caras* >

os rostos) ou diferenças varietais que podem ser puramente lexicais (*casa de banho* > *banheiro*) ou sintáticas (*a regredir* => *regredindo*), usadas na conversão para o PB.

## CONCLUSÕES

Este artigo teve como objetivo o alinhamento e análise de unidades parafrásticas em texto literário, nomeadamente em toda a obra de Afonso Cruz, *Os Livros que devoraram o meu pai*. O alinhamento foi feito semimanualmente por linguistas portuguesas e brasileiras. Desta tarefa de alinhamento resultaram pares de unidades parafrásticas que classificamos de acordo com a sua natureza sintático-semântica, estilística ou varietal. Os resultados são de um interesse tão vasto que poderíamos ter descrito mais tipos de contrastes. Por razões de tempo e espaço, não o pudemos fazer neste artigo. Decidimos selecionar um fenómeno linguístico mais restrito, unidades lexicais simples, multipalavras e estruturas frásicas com função adjetival, analisá-lo e discuti-lo do ponto de vista linguístico e com a perspectiva de adaptação de texto literário em mente. Reunimos um conjunto de pares de unidades parafrásticas que podem servir para enriquecer ferramentas de parafraseamento incluindo conversão de texto entres as variedades de PE e de PB. Os problemas diagnosticados são muito semelhantes aos que ocorrem nas traduções, onde frequentemente os editores exigem um tipo de escrita convencional e pouco criativo na tradução, mesmo quando os autores escrevem de forma fora do normal e não convencional, perdendo-se assim a riqueza linguística e literária da obra original. Assim sendo, muitas opções literárias, que foram intencionalmente escolhidas para despertar estranheza ou por razões lúdicas, acabam por se perder ou são explicitadas de tal forma que acabam esvaziadas da sua riqueza inicial. Da mesma forma, certos recursos expressivos acabam por se perder, como é o caso das repetições nesta obra, que a edição em PB evitou, mas que constam da edição original em PE, naturalmente mais próxima das escolhas do autor.

## REFERÊNCIAS BIBLIOGRÁFICAS

BARREIRO, Anabela. **Make it Simple with Paraphrases: Automated Paraphrases for Authoring Aids and Machine Translation**. Dissertação de Doutoramento, Faculdade de Letras da Universidade do Porto, 2009.

\_\_\_\_\_; BATISTA, Fernando. Machine translation of non-contiguous multiword units. In: WORKSHOP ON DISCONTINUOUS STRUCTURES IN NATURAL LANGUAGE PROCESSING (DiscoNLP 2016), 2016, San Diego, California, USA. **Proceedings of DiscoNLP 2016**. Stroudsburg: Association for Computational Linguistics, 2016. pp. 22-30.

\_\_\_\_\_; MOTA, Cristina. Paraphrastic variance between European and Brazilian Portuguese. Workshop on NLP for Similar Languages, Varieties and Dialects (VarDial 2018), 5., 2018, Santa Fe, New Mexico, USA. **Proceedings of the Fifth Workshop on NLP for Similar Languages, Varieties and Dialects (VarDial 2018)**. Stroudsburg: Association for Computational Linguistics, 2018. pp. 111-121.

\_\_\_\_\_; RAPOSO, Francisco Afonso; LUÍS, Tiago. CLUE-Aligner: an alignment tool to annotate pairs of paraphrastic and translation units. In: LREC 2016 WORKSHOP TRANSLATION EVALUATION: FROM FRAGMENTED TOOLS AND DATA SETS TO AN INTEGRATED ECOSYSTEM, 2016, Portorož, Eslovénia. **Proceedings of the LREC 2016 Workshop Translation Evaluation: From Fragmented Tools and Data Sets to an Integrated Ecosystem**. Paris: European Language Resources Association (ELRA), 2016. pp. 7-13.

\_\_\_\_\_; REBELO-ARNOLD, Ida; BAPTISTA, Jorge; MOTA, Cristina. Parafraseamento automático de registo informal em registo formal na língua portuguesa. **Linguamática**, Braga e Vigo, v. 10, n.º 2, pp. 53-61, dez. 2018.

CRUZ, Afonso. **Os livros que devoraram meu pai: a estranha e mágica história de Vivaldo Bonfim**. 8.ed. Alfragide: Caminho, 2015. 128pp.

CRUZ, Afonso. **Os livros que devoraram meu pai: a estranha e mágica história de Vivaldo Bonfim**. São Paulo: Leya, 2011. 112pp.

GARCEZ, Isabel. **O papel da edição literária no mercado da cultura**. 2017. 314 f. Tese (Doutoramento em Estudos de Cultura) – Faculdade de Letras da Universidade de Lisboa, Lisboa, 2017.

KILGARRIFF, Adam; RYCHLÝ, Pavel; SMRZ, Pavel; & TUGWELL, David. (2004). **The Sketch Engine**. In: EURALEX INTERNATIONAL CONGRESS, 10., 2004, Lorient, France. Proceedings of the 11th EURALEX International Congress, Lorient: Université de Bretagne-Sud, Faculté des lettres et des sciences humaines, 2004. pp. 105-11.

MOTA, Cristina; BARREIRO, Anabela; RAPOSO, Francisco Afonso; RIBEIRO, Ricardo; CURTO, Sérgio dos Santos Lopes; COHEUR, Luísa. eSPERTo paraphrastic knowledge applied to question-answering and summarization. In: BARONE, Linda; MONTELEONE, Mario; SILBERZTEIN, Max. (Ed.). **NooJ 2016: automatic processing of natural-language electronic texts with NooJ**. Cham: Springer International Publishing AG, 2016. p. 208–220. (Series Communications in Computer and Information Science vol. 667).

NEVES, Maria Helena de Moura. **A gramática funcional**. São Paulo: Martins Fontes, 1997. 160 pp.

REBELO-ARNOLD, Ida; BARREIRO, Anabela; QUARESMA, Paulo; MOTA, Cristina. Alinhamentos parafrásticos PE-PB de construções de predicados verbais com o pronome clítico 'lhe'. **Linguamática**, Braga e Vigo, v. 10, n.º 2, pp. 3-11, dez. 2018.

SANTOS, Diana. Português internacional: alguns argumentos. In: TEIXEIRA, José (ed.). **O Português como Língua num Mundo Global: problemas e potencialidades**. Braga: Centro de Estudos Lusíadas da Universidade do Minho, 2016. pp. 49-66.

WITTMANN, Luzia Helena; PÊGO, Tânia Regina; SANTOS, Diana. Português Brasileiro e Português de Portugal: algumas observações. In: ENCONTRO NACIONAL DA APL, 11., 1995, Lisboa, Portugal. **Atas do XI Encontro Nacional da APLB**. [S.l.:s.n.], 1995. pp. 465-487.