

One Book, Two Language Varieties^{*}

Anabela Barreiro¹, Ida Rebelo-Arnold², Fernando Batista³, Isabel Garcez³,
and Tanara Zingano Kuhn³

¹ INESC-ID, Lisboa anabela.barreiro@inesc-id.pt

² Universidad de Valladolid

³ Instituto Universitário de Lisboa (ISCTE-IUL) and INESC-ID Lisboa

⁴ Universidade de Lisboa

⁵ Universidade de Coimbra

Abstract. This paper presents a comparative study of alignment pairs, either contrasting expressions or stylistic variants of the same expression in the European (EP) and the Brazilian (BP) varieties of Portuguese. The alignments were collected semi-automatically using the CLUE-Aligner tool, which allows to record all pairs of paraphrastic units resulting from the alignment task in a database. The corpus used was a children’s literature book *Os Livros Que Devoraram o Meu Pai* (The Books that Devoured My Father) by the Portuguese author Afonso Cruz and the Brazilian adaptation of this book. The main goal of the work presented here is to gather equivalent phrasal expressions and different syntactic constructions, which convey the same meaning in EP and BP, and contribute to the optimisation of editorial processes compulsory in the adaptation of texts, but which are suitable for any type of editorial process. This study provides a scientific basis for future work in the area of editing, proofreading and converting text to and from any variety of Portuguese from a computational point of view, namely to be used in a paraphrasing system with a variety adaptation functionality, even in the case of a literary text. We contemplate “challenging” cases, from a literary point of view, looking for alternatives that do not tamper with the imagery richness of the original version.

Keywords: European and Brazilian Portuguese · Stylistic Choices · Variety Distinction · Variety Adaptation Tool · Paraphrasing · Text Adaptation · Editing and Proofreading · Editorial Process.

1 Introduction

This paper focuses on comparing alignments, i.e., pairs of paraphrastic units used in the Portuguese (original version) and Brazilian (adapted version) editions of the children’s literature book *Os Livros Que Devoraram o Meu Pai* (*The Books that Devoured My Father*) by the Portuguese author Afonso Cruz. The comparative analysis intends to verify what kind of alterations were made in

^{*} This work was supported by national funds through Fundação para a Ciência e a Tecnologia (FCT) with reference UID/CEC/50021/2019.

the adaptation process. The methodology used in this research work consists of: (i) at an early stage, building sentence-level alignments of the entire book, contemplating, on the one side, EP, and, on the other side, BP; (ii) feed the CLUE-Aligner tool with the parallel corpus; (iii) perform multi-layer lexical unit, multiword, or expression-level alignments (not only with lexical, but also with syntactic and stylistic implications); (iv) analyse and categorise each alignment from a contrastive linguistics point-of-view, i.e., from an approach that seeks to describe the differences and similarities between the pair of language varieties; (v) establish a typology of contrasts between EP and BP; and finally, (vi) discuss the implications of linguistic changes in the constitution of the new text, in semantic terms, also occasionally analyzed from the literary and/or cultural perspectives, i.e., does the conversion respect the intention of the author’s original idea? Does it change the nature of the meaning intended by the author? Etc.

From a linguistic point of view, the specific objective of our work is to identify, analyze, categorise and discuss a selection of simple lexical units, multiwords, phrases, and expressions, which are alternating in the two varieties, i.e., which present differences between the author’s choices in the EP variety (original version) and the corresponding solutions adopted in the conversion to BP. The chosen methodology focuses on the functional grammar-based contrastive linguistic analysis [11], implemented with the aid of computational treatment tools based on the eSPERTo project and using semiautomatic alignments generated by the CLUE-Aligner [5].⁶

From a literary point of view, and based on the same linguistic selection, we will analyze the validity of the editorial choices in the BP edition on two fronts: 1) the rationale of the author’s choices in the EP edition according to their polysemy, intentional “strangeness”, rhythmic richness, and network of meaning relations in the context in which they arise and with the work as a whole; 2) evaluation of the qualitative equivalence in the reception of both editions regarding the same parameters stated in 1), considering that, both the EP and the BP editions can be worked in classroom context by a target audience of the same age group. Based on the results achieved in this work, we also intend to: a) encourage the gathering of linguistic resources for the purposes of mediation of literature in Portuguese, mother tongue and/or foreign language; b) question the importance of literary mediation, whether personalized by the editor, the reviewer or the teacher, with regard to literary reception; c) present a range of editorial and reading mediation alternatives that will preserve the text of the original edition and thus d) increase and enrich the level of linguistic, literary and cultural knowledge among readers of both the EP and the BP varieties of Portuguese.

⁶ The corpora of the study are the EP and the BP editions of the aforementioned book, building on previous research work, where we have already discussed in detail certain linguistic phenomena, such as the position of clitic pronouns in EP and BP [12] or the contrast between languages (in this case, language varieties) that tend to use more contractions than others, such as contractions of prepositions with determiners in the case of EP [3].

The editorial process has significant implications in the quality of a literary text, as well as in the quality of reception. Therefore, a careful proofreading is indispensable in any editorial task. We highlight the literary implications in the examples.

The present study is part of a broader project that proposes lines of comparative analysis of Brazilian editions of literary works by Portuguese authors and Portuguese editions of literary works by Brazilian authors. We argue that the development of the comparative study of these editions plays an important role in a more fruitful transatlantic relationship regarding literary creation in Portuguese and the promotion of cultural and linguistic diversity, especially empowering language users to access the full richness of a language spoken on both sides of the Atlantic Ocean. In the near future, it will be pertinent to foster the development of tools to assist editors, reviewers and teachers, as well as students and readers from the Portuguese-speaking countries in several continents, either from the point of view of mediation and literary reception or from the point of view of awareness of the linguistic contrasts between the language varieties in the Portuguese-speaking communities spread around the world.

2 Related Work

In recent years, editorial revision processes have benefited not only from linguistic reflections and orientations, but also from natural language processing tools, which can serve to develop corpora analyses, generation and summarisation of texts, translation, paraphrasing, among others. In an attempt to complement tools that aid in writing systems available in the market (mostly spelling and superficial grammar correction), [2] created a paraphrase generation system called eSPERTO, an acronym for *Sistema de Parafaseamento para Edição e Revisão de Texto* (System of Paraphrasing in Editing and Revision of Text [9], which aims to assist in various tasks, such as summarizing and answering questions ([10]), language teaching, support for machine translation ([2]; controlled technical writing, editing and proofreading, stylistic adaptation ([12], and text conversion between language varieties ([4]), among others. In the last case study, the use of the tool allowed us to analyze the paraphrastic units of a complete literary work in the EP and BP varieties of Portuguese, following some previously performed experimental works of adaptation between them (cf. [4]; [12])). In this context, we outline a more global, dynamic and respectful course for the divergences and convergences that make Portuguese one of the richest languages on our planet, taking a further step towards its internationalization ([14]).

3 Parallel Corpus

Our corpus consists of the full text of two editions of Afonso Cruz's book, namely, the original in EP and its version in BP. The reasons that led to the choice of this parallel corpus are related to: 1) it is a literary work targeting children and a young public, and for this reason it can be argued that the adaptation

of the text requires a more incisive treatment, since the understanding of the text, in the adapted version, should be equivalent to that of the original text's target audience; 2) it is included in the lists of the Portuguese National Reading Plan, recommended for autonomous reading of the 3rd cycle (12-14 years); 3) it integrates the lists of the National Library School Program in 2013 in Brazil, as recommended reading in the final years of elementary school; 4) the text of the Brazilian edition presents significant differences in relation to that of the Portuguese edition.

4 Methodology

The parallel corpus contains 1,260 pairs of sentences.⁷ Phrases are displayed in grid-like CLUE-Aligner, where the EP phrase appears vertically forming multiple columns and the BP phrase appears horizontally forming multiple rows. The intersection of words in columns and rows forms a matrix where each word corresponds to a cell. The alignment task involves mapping boundaries, according to linguistic analysis criteria. Figure 1 presents the mapping of the alignment of the paraphrastic units of the sentences in example (1).

- (1) *EP* - *Por vezes a voz dela fica um pouco amarrotada.*
BP - *Por vezes, a voz dela fica um pouco enrrouquecida.*
EN - *Sometimes, her voice gets a bit wrinkled/hoarse.*

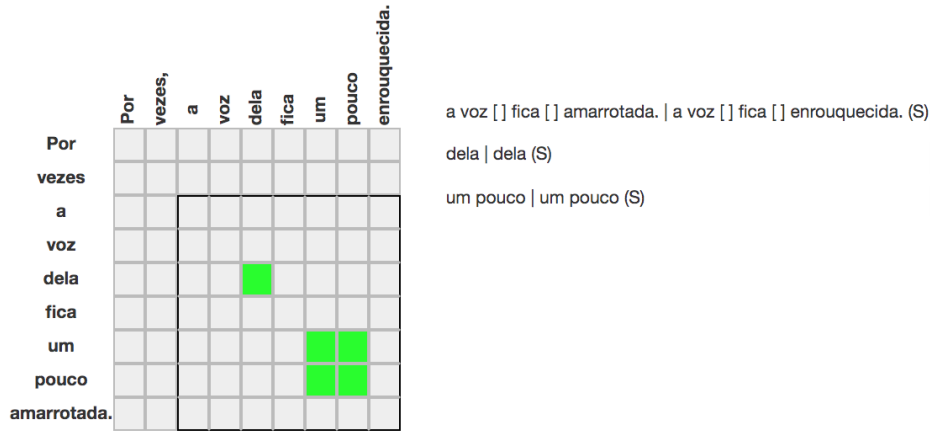


Fig. 1. EP versus BP paraphrastic units.

Figure 1 illustrates the alignment of the EP–BP structures *a voz [] fica [] amarrotada* — *a voz [] fica [] enrrouquecida*, where the only difference is in the

⁷ The most frequent linguistic contrasts of this parallel corpus have been listed and discussed in [13], whose focus was on word sequences analyses.

adjectival forms: *amarrotada* (wrinkled) in EP, and *enrouquecida* (hoarse) in BP. The way to represent the basic structure is to extract the insertions (*dela* (her) and *um pouco* (a bit)), whose alignment mapping is represented in green), which get automatically aligned by design of the CLUE-Aligner tool.

5 Types of Linguistic Phenomena in EP–BP Alignments

The alignment task resulted in a wide variety of contrastive phenomena, some purely lexical (e.g., *apanhei o comboio—peguei o trem* (I caught the train)), other purely syntactic (e.g., *a trabalhar—trabalhando* (walking)) while others, both lexical and syntactic (e.g., *correu mal—deu tudo errado* (it went (all) wrong)). The most common contrastive phenomena found in the EP–BP paraphrastic alignments were lexical, either **Variety Distinction**, or **Lexical Choice**.

Variety Distinction (VAR) represents contrastive alignments where one of the elements in the pair is either only used or common in one of the varieties of Portuguese, either EP or BP (e.g., *apelido—sobrenome*). **Lexical Choice** (LEX-choice) represents the writer’s/editor’s choice as a variety user, even if both expressions would be virtually accepted by readers in either one of the speakers communities. Lexical Choice covers a large amount of paraphrastic units in the corpus, and sometimes it is hard to explain why the editor has changed the expression used in the source text. Some of those unnecessary changes could derive into a misinterpretation of what is variety distinction and what variety distinction implies, i.e., when looking at the EP–BP alignments from the perspective of a student of Portuguese as a foreign language, one might interpret the changes in BP as “necessary” changes due to the low acceptability of the original expression in EP (variety distinction). However, it is a matter of choices by the Portuguese author and the Brazilian editor as users of each variety of Portuguese, not a variety distinction. This will be the subject of further research on variety adaptation in the field of editing and proofreading.

The most well-know variants are of lexical and spelling nature (e.g., *crônicas—crônicas* (stories); *relva—grama* (grass); *casa de banho—banheiro* (bathroom); *rapariga—garota* (girl, young lady), *carteirista—ladrão de carteira* (pickpocket), etc.). These variants have been widely described, including in early EP-BP contrastive works developed by computational linguists [16], [6], and can be easily found in a contrastive dictionary. Also vastly studied within the Framework of the Lexicon-Grammar have been the contrasts between support verb constructions and morpho-syntactic related single verbs (e.g., *entrou dentro dum livro—adentrou um livro* (went inside a book)). Therefore, we will focus, from now on, on contrasts that involve different types of lexical choices and syntactic transformations. It should be noted that the following presentation of the analysed linguistic phenomena does not reflect corpus-frequency order, but rather a set of interesting alignments; some of them happen to be also highly frequent.

Verb Tenses — One of the most consistent pervasive phenomena is the use of different verb tenses between EP and BP, especially the use of the Past Perfect

Compound versus the use of the Past Tense, as demonstrated in example (2). Although we consider that there is no justification for this amendment, because it is always preferable to respect the original text, and we are not in the presence of an error, this substitution of verb tenses consists of a stylistic choice, which does not really represent any change in meaning.

- (2) *EP* - *o meu pai **tinha previsto** uma coisa destas.*
BP - *ele **previu** uma coisa dessas.*
EN - *my father/he (had) predicted such a thing.*

VINF (CLITIC) vs PRO-RELque (CLITIC) VSUBJ — Example (3) illustrates a productive phenomenon that consists of a stylistic choice, used in any one of the Portuguese varieties. The syntactic structure of the paraphrasing pair contains two clauses, a main clause and a dependent one. The verb of the main clause can select an infinitive construction (*para me sentar*) (to sit down) or a relative clause (*que me sentasse*) (that I sit down). However, literature is also made of details and, in this case, the verb *dizer* (say) is more assertive than the verb *pedir* (ask), so our editorial choice would be to keep the verb of the EP version. However, on the other hand, in BP, the expression with the verb *dizer* has a connotation of “order”, and that might be the reason why the Brazilian editor chose the verb *pedir*.

- (3) *EP* - *Ela disse-me **para me sentar***
BP - *Ela pediu **que eu me sentasse***
EN - *She said/asked me to sit down*

PREP-a VINF vs VGER-ndo — One of the most frequent phenomenon contrasting EP and BP is the typically known variety contrast between the use of the verbal construction constituted by the preposition *a* with the infinitive form of the verb in EP and the use of the gerundive form of the verb [PREP(a) VINF vs VGER-ndo] (e.g., *a olhar—olhando* (looking) or as in example (4), *a caminhar—caminhando* (walking)).

- (4) *EP* - *senti-me **a caminhar** pela grande Avenida Nevski*
BP - *senti-me **caminhando** pela grande avenida Nevski*
EN - *I felt walking through the large Nevski Avenue*

CONTR-PREPDET vs PREP DET — In general, the use of contractions is more prolific in EP than in BP. As illustrated in (5), contractions (e.g., *dum*) occur mostly between a preposition (e.g., *de*) and a determiner, an article (*um*). By replacing the verb *preguiçava* (lazed) in EP by *descansava* (relaxed) in BP, the reader loses an incisive and more visual effect, especially occurring with the subject: *o nariz* (the nose).

- (5) *EP* - *O nariz preguiçava em cima **dum** bigode preto*
BP - *O nariz descansava em cima **de um** bigode preto*
EN - *The nose lazed/relaxed on the top of a black moustache*

2nd vs 3rd PERS-SG — Example (6) shows the use of the second versus the third person singular form of the verb (e.g., *Areve-te*—*Atreva-se* (Dare ('yourself'))). The absence of a subject pronoun or noun both in EP and BP is due to the use of the Imperative verb tense.

- (6) *EP* - *Atreve-te*
BP - *Atreva-se*
EN - *Dare ('yourself')*

PRO-DROP V PREP vs PRO V — EP tends to be a stronger pro-drop variety than BP. Example (pro-drop) contrasts the absence of the subject pronoun in EP with the realisation of the pronoun *vocês* (you) in BP.

- (7) *EP* - *É a verdade objetiva como haverão de perceber*
BP - *É a verdade objetiva, como **vocês** vão perceber*
EN - *It is the objective truth, as you will understand*

DET vs NO-DET B4 PRO-POSS — Example (8) illustrates the use of a determiner versus the use of no-determiner, in the context of an antecedent of a possessive pronoun (e.g., *os seus gestos*—*seus gestos* (his gestures)). In general, EP tends to use the determiner, while BP tends to avoid it. The change from *cuidados* (elegant) to *cuidadosos* (careful) is unnecessary (the meanings of these words is equivalent in both varieties of Portuguese) and amends the original meaning, because *cuidados* is not an unquestionable synonym of *cuidadosos*.

- (8) *EP* - *fico sempre fascinado a olhar para **os seus** gestos cuidados*
BP - *sempre fico fascinado observando **seus** gestos cuidadosos*
EN - *I always get fascinated when I look at his elegant/careful gestures*

Word Order — A common conversion strategy between the EP and the BP versions is the word order, which implies words from different parts of speech. Example (9) illustrates the change in the word order of the quantifier *todos* (all). However, examples in the corpus exemplify changes in the word order of adjectives (e.g., *mundo enfadonho*—*enfadonho mundo* (boring world)), adverbs (e.g., *aqui me trouxe*—*me trouxe aqui* (brought me here)), or subjects, either pronouns (e.g., *Que poderia eu*—*Que eu poderia* (What could I) or nouns *Como se chama o senhor?*—*Como o senhor se chama?* (What's your name?)). These editorial choices has different literary implications, which we will not explain in the present work.

- (9) *EP* - *com aqueles livros **todos** sentados nas prateleiras*
BP - *com **todos** aqueles livros sentados nas prateleiras*
EN - *with all those books (all) sitting in the shelves*

Word Reduction and Word Augmentation — The corpus presents many cases of word reduction and word augmentation in a phrase or expression from EP into BP. In example (10), the author of the original text seeks proximity with his young readers through the use of the interrogative cleft structure *é que* (where *is it that* he is?), which is typically used in the spoken language modality. Several authors have pointed out that the interrogative cleft structure is easily found in children’s language (cf. [8] and [1], among others), and the target of this book are young readers, so we would maintain the original form in EP.

- (10) *EP* - Onde **é que** ele está?
BP - Onde ele está?
EN - Where is (it that) he (is)?

In example (11), the Portuguese author seeks to build a fun word game by following the participial adjective *encolhida* (shrunk) by the title of a book *A Servidão Humana* (The Human Servitude). This diversion effect is lost in the BP version with the introduction of the noun *volume* (volume) before the book’s title. We interpret this option of the BP editor as an attempt to make clear to the young reader what the sentence is about.

- (11) *EP* - *encolhida entre o A Servidão Humana de Sommerset Maugham e um livro de Herberto Helder*
BP - *encolhida entre o volume de A Servidão Humana, de Sommerset Maugham, e um livro de Herberto Helder*
EN - *crouched between the (volume of) Human Servitude by Sommerset Maugham and a book by Herberto Helder*

SG vs PL N or NPs — Several instances of the contrast between the use of singular versus the use of the plural form of a noun phrase happen in the corpus, as in example 12 (e.g., *no cabelo* — *nos cabelos* (in the hair), where the author of the original text chooses the singular form often used in the spoken modality to more easily approach his target audience. On the other hand, the reviewer chooses the plural form, used in a higher register. The use of plural and singular in nominal agreement and its relation to higher or lower registers and spoken modality have been discussed in the literature (cf. [7] and [15], among others).

- (12) *EP* - O meu melhor amigo não tinha óleo **no cabelo**
BP - Meu melhor amigo não tinha óleo **nos cabelos**
EN - My best friend didn’t have oil in the hair

6 Quantitative Results

The quantitative results in Table 1 show that there are 4 major types of linguistic phenomena in the EP–BP alignments: (i) the use of different syntactic structures (SYNT) (37.4%); (2) the use of different lexical choices (LEX-choice) (30,6%); (iii) different lexicon (including words, and multiword units) based on variety

distinctions (VAR) (15,1%); and (iv) the combination of the use of different lexical choices plus different syntactic structures (LEX-choice+SYNT) (10.8%). A considerable percentage of alignments were identical (SAME), either due to the design of CLUE-Aligner, which automatically aligns insertions in non-contiguous multiwords, whether these insertions are different or identical in EP and BP (it often happens that they are identical and they add up to type SAME). There are a few counts of the combination of the use of variety contrasts in distinct syntactic structures (VAR+SYNT), and of alignments where there is a clear evidence of semantic deviation (SEM DEV) from the original text in the edited text, for reasons which only the editor knows (or does not know, i.e., natural human errors, lack of knowledge of the exact meaning of the word, multiword, or expression in the original or the lack of understanding of the author’s intended meaning).

Table 1.

Distribution of linguistic phenomena in EP–BP Alignments by most common types					
Type	Count	Percent.	Alignment	Example	Translation
SYNT	325	37.4%	<i>dir-lhe-ei—lhe direi</i>	(I) will tell him/her	
LEX-choice	266	30.6%	<i>esquecidos—absortos</i>	forgotten—absorbed	
VAR	131	15.1%	<i>apanhei o comboio</i>	(I) caught the train	
			<i>—peguei o trem</i>	—(I) caught the train	
LEX-choice+SYNT	94	10.8%	<i>vê-se</i>	it can be seen	
			<i>—está aparente</i>	— it is clear/obvious	
SAME	38	4.4%	<i>depois—depois</i>	then	
VAR+SYNT	8	0.9%	<i>que vê ali na relva</i>	that he sees in the grass	
				— there in the grass	
SEM DEV	7	0.8%	<i>mais nada</i>	nothing else/more	
			<i>—tudo</i>	— everything	
Total	869	100%			

For the sake of space, we will not present here the distribution of linguistic phenomena in the EP–BP alignments by the most common **subtypes**. However, we present, in Table 2, the most frequent EP–BP alignments within the subtype LEX-choice (5 or more occurrences). From these, with more than 10%, there are noun, verb, adjective and preposition differences, and verb tense differences. The table is self-explanatory.

7 Conclusions

We compared the EP and BP editions of the children’s literature book *Os Livros que Devoraram o Meu Pai* (The Books that Devoured my Father), by Portuguese author Afonso Cruz, and analyzed the equivalence between the pairs of paraphrastic units found in this literary text. The alignment was done semi-manually by Portuguese and Brazilian linguists. We classified the pairs of paraphrastic

Table 2.

Most frequent EP–BP Alignments in the LEX-choice type					
LEX-choice	Count	Percent.	Alignment Example	Translation	
N-diff	38	19.9%	<i>cheiro—odor</i>	smell—odour	
V-diff	36	18.8%	<i>tremi—estremeci</i> (I)	shook—shuddered	
ADJ-diff	27	14.1%	<i>massuda—corpulenta</i>	chunky—corpulent	
V-tense	26	13.6%	<i>havia saído—saíra</i>	had left—left	
PREP-diff	24	12.6%	<i>atrás de—após</i>	afterwards—after	
EXPR	17	8.9%	<i>em suores—transpirado</i>	in sweat—perspired	
PRO-DEM-diff	9	4.7%	<i>estas—essas</i>	these—those	
ADV-diff	8	4.2%	<i>depressa—rapidamente</i>	fast—quickly	
V-diff+V-tense	6	3.1%	<i>aconteceu—acontecera</i>	it happened	
Total		100%			

units according to their syntactic-semantic, stylistic or varietal nature. The results are of wide interest, but it is impossible to describe them all in this paper. Therefore, we selected just narrower pervasive linguistic phenomena, namely the conversion of/contrast between verb tenses, nominal agreement, word order and collocation of elements in the phrase, such as determiners and pronouns, in the original EP text and in its Brazilian version, and discuss it from the linguistic point of view, with the perspective of literary text adaptation in mind. We put together a set of paraphrastic pairs that can enrich paraphrasing tools including text conversion between the EP and the BP varieties. The problems diagnosed are very similar to those that occur in translations, where editors often require a conventional type of translation writing, even when authors write in an unusual and unconventional way, thus losing the linguistic richness of the original work. As such, many literary options, which were intentionally chosen to arouse strangeness or used for playful reasons, are eventually lost or made explicit in such a way that they are emptied of their initial “richness”. In the same way, some expressive resources in the EP original edition, and, therefore, naturally closer to the author’s choices, are lost in the BP edition.

References

1. Cavalcante de Araújo, R.: Sobre as interrogativas clivadas (básicas) QU e polares. *Fórum Linguístico* **16**(1), 3530–3544 (2019)
2. Barreiro, A.: Make it Simple with Paraphrases: Automated Paraphrasing for Authoring Aids and Machine Translation. Ph.D. thesis, Universidade do Porto, Porto, Portugal (2009)
3. Barreiro, A., Batista, F.: Contractions: to align or not to align, that is the question. In: Proceedings of the COLING-Workshop on Linguistic Resources for NLP (2018)
4. Barreiro, A., Mota, C.: Paraphrastic Variance between European and Brazilian Portuguese. In: Zampieri, M., Nakov, P., Ljubešić, N., Tiedemann, J., Malmasi, S., Ali, A. (eds.) Proceedings of the Fifth Workshop on NLP for Similar Languages, Varieties and Dialects (VarDial). COLING 2018, Association for Computational Linguistics (2018)

5. Barreiro, A., Raposo, F., Luís, T.: CLUE-Aligner: An Alignment Tool to Annotate Pairs of Paraphrastic and Translation Units. In: Proceedings of the 10th Edition of the Language Resources and Evaluation Conference. pp. 7–13. LREC 2016, European Language Resources Association (2016)
6. Barreiro, A., Wittmann, L., Pereira, M.: Lexical differences between European and Brazilian Portuguese. *The INESC Journal of Research & Development* **5.2**, 75–101 (12 1996)
7. Castilho, A.: *Nova Gramática do Português Brasileiro*. Contexto (2010)
8. Kato, M.A., Ribeiro, I.: Cleft sentences from Old Portuguese to Modern Portuguese, pp. 123–154. John Benjamins (2009)
9. Mota, C., Barreiro, A., Raposo, F., Ribeiro, R., Curto, S., Coheur, L.: eSPERTo's Paraphrastic Knowledge Applied to Question-Answering and Summarization. In: Barone, L., Monteleone, M., Silberstein, M. (eds.) *Automatic Processing of Natural-Language Electronic Texts with NooJ: 10th International Conference, NooJ 2016, České Budějovice, Czech Republic, June 9-11, 2016, Revised Selected Papers*. pp. 208–220. Springer International Publishing, Cham (2016)
10. Mota, C., Carvalho, P., Raposo, F., Barreiro, A.: Generating Paraphrases of Human Intransitive Adjective Constructions with Port4NooJ. In: Okrut, T., Hetsevich, Y., Silberstein, M., Stanislavenka, H. (eds.) *Automatic Processing of Natural-Language Electronic Texts with NooJ*. pp. 107–122. Springer International Publishing, Cham (2016)
11. Neves, M.H.M.: *A gramática funcional. Texto e linguagem*, Martins Fontes (1997), <https://books.google.pt/books?id=auYbAQAAIAAJ>
12. Rebelo-Arnold, I., Barreiro, A.: EP–BP Paraphrastic Alignments of Verbal Constructions Involving the Clitic Pronoun *lhe*. In: *International Conference on Computational Processing of Portuguese. PROPOR 2018*, Springer (2018)
13. Rebelo-Arnold, I., Barreiro, A., Kuhn, T.Z., Garcez, I., Batista, F.: III Simpósio Internacional História, Cultura e Relações de Poder: Revoluções no Mundo Lusófono, chap. *Análise Comparativa das Edições Portuguesa e Brasileira da Obra Os Livros Que Devoraram o meu Pai*, de Afonso Cruz. Editora LiberArs (2019), <http://www.lusosofia.net/>
14. Santos, D.: *O Português como Língua num Mundo Global: problemas e potencialidades*, chap. *Português internacional: alguns argumentos*, pp. 49–66. Centro de Estudos Lusíadas da Universidade do Minho (2016)
15. Scherre, M.M.P.: *Reanálise da concordância nominal em português*. Ph.D. thesis, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brasil (1988)
16. Wittmann, L., Pêgo, T., Santos, D.: *Português Brasileiro e Português de Portugal: algumas observações*. In: *Atas do XI Encontro Nacional da APL*. pp. 465–487. Associação Portuguesa de Linguística, Lisboa, Portugal (1995)