






Paraphrasing Emotions in Portuguese

Cristina Mota^{1,3} , Diana Santos^{2,3} , and Anabela Barreiro^{1,3} 

¹ INESC-ID, Lisbon, Portugal

`cmota@ist.utl.pt`, `anabela.barreiro@inesc-id.pt`

² ILOS-UiO, Oslo, Norway

`d.s.m.santos@ilos.uio`

³ Linguateca, Aveiro, Portugal

Abstract. In this paper we describe the cooperation between the NooJ-based paraphrasing system eSPERTo and the broad-coverage emotion annotation of Linguateca's corpora, Emocionário, to develop strategies for analyzing the paraphrasing of emotions. We start by briefly introducing these two projects. Then, we describe how we selected five parallel texts in European and Brazilian Portuguese, automatically annotated them with Emocionário, and carefully revised the annotation, while paying special attention to multiword units and other phrases representing emotion paraphrastic units, and created a set of putative paraphrasing grammars in NooJ. We also comment on the translations and discuss disagreements on the annotation of emotions, as well as on formalization strategies used in the internal structure of the paraphrases.

Keywords: Emotions · Paraphrastic units · Paraphrase generation

1 Motivation

Emotions are described from several perspectives [9, 21] and, according to Wierzbicka [24, p. 276, 302–4], the variability in the forms of expression in language is very high.

This paper is the result of collaboration between two projects: **Emocionário**, which aims at organizing emotions in Portuguese and annotating them in corpora, and **eSPERTo**, the goal of which is to develop a large-scale paraphrasing system in which the NooJ linguistic engine, grammars, and lexicons are used together with other paraphrase acquisition techniques.

The aims of this collaboration were fivefold. From Emocionário's point of view, it would be particularly useful to have an emotion paraphraser to help identify more cases of emotions in corpora, while from eSPERTo's point of view, adding emotion paraphrases would considerably enhance its paraphrasing power. Thirdly, the application of the emotion classification to a hitherto unused application domain would be a good way to evaluate Emocionário's capabilities and shortcomings, and both projects would gain from learning more about real paraphrases of emotion in text. Finally, we examine the interesting question of how to assess the methodology employed to harvest emotion paraphrases from parallel text.

2 The Two Projects in a Nutshell

Emocionário¹ is a subproject of Linguateca [10, 16]. For more than twenty years, Linguateca has been making annotated corpora publicly available for Portuguese and has for fifteen years dealt with semantic annotation (of some domains) as well [19]. Linguateca's team started looking at emotions in 2014. Before the work described here, 24 emotion groups had been identified, covering more than 4,000 lexemes that can refer to an emotion in Portuguese (see [22] for the list of the groups present in a 20-million-word literary corpus in October 2019).

It is important to highlight the point that this project is concerned with references to – not expressions of – emotion. To clarify, an expression such as *Oh, my God!* expresses emotion but does not refer to it.

One of the goals of Emocionário is to understand what a reference to emotion in Portuguese is, and how to go about it systematically [15]. Also, we aim to gather arguments for the groupings suggested [17].

eSPERTO² is a project that aims at the development of a paraphrasing system using enhanced resources [11] developed within the NooJ [23] linguistic engine]. This system provides suggestions of semantically similar expressions for text (re)writing, and has been used, for example, in machine translation [1] and in the conversion of informal to formal language and vice-versa [4]. Within the eSPERTO project, we integrated and improved three lexicon-grammar tables which were already available for Portuguese. These linguistic resources resulted in (i) paraphrases of human intransitive adjectives (HIAs) [14]; (ii) paraphrases of support verb constructions (SVCs) with *fazer* [12]; and (iii) paraphrases of SVCs with *ser de* [13].

3 The Actual Study

In order to study the paraphrasing of emotions in Portuguese, we did the following:

- (1) Automatically annotated parallel texts with emotions using Emocionário;
- (2) Manually revised the result to guarantee the correctness of the conclusions and to assess the quality of the automatic annotation;
- (3) Identified cases (a) in which emotions (in an emotion group) had been maintained but with other words, and (b) in which different emotions had been chosen;
- (4) Coded them as paraphrastic resources in NooJ (Fig. 1).

¹ <https://www.linguateca.pt/Emocionario/>.

² <https://esperto.hlt.inesc-id.pt/esperto/esperto/demo.pl>.

eSPERTO - System for Paraphrasing in Editing and Revision of Text

Parameters

Demo mode

Interface idiom English

Resources idiom Portuguese

Dictionary PT-Dict_NEW

Sample text SAN

Paraphrasing Check all Uncheck all

GN.old

Active > Passive

Passive > Active

Simple adverb > Compound

Compound adverb > Simple

Nominal/adjectival predicate > Verb

Nominal/adjectival predicate > Nominal/adjectival predicate

Verb > Nominal/adjectival predicate

Relative construct > Adjective

Possessive > Relative construct

Synonyms

Human intransitive adjective

Predicate nouns with Vsup fazer

Predicate nouns with Vsup ser de

PE > PB

PB > PE

SAL

Informal > Formal

Emotions

Debug

Process results

Input file or text (click to show/hide)

Choose file: Browse file

Insert text in the text box

Subitamente, o escritor sorriu com essa ideia.

Process results

Results (click to show/hide)

Subitamente, [o escritor sorriu com essa ideia] .

Save paraphrased text

a ideia fez o escritor sorrir

a ideia fê-lo sorrir

o pensamento fez o escritor sorrir

o pensamento fê-lo sorrir

Suggest your own paraphrase >

Fig. 1. Paraphrasing emotions in eSPERTO

3.1 The Material

It is not easy to obtain “parallel material” in the same language. In fact, most of the cases we could find came from different translations of an original book from another language. We had five texts, each published in two varieties of Portuguese (the one used in Portugal, or PP, and the one used in Brazil, or PB), already sentence-aligned in previous projects, namely, the COMPARA [7] and PANTERA [20] parallel corpora of Linguateca, and used in a contrastive study of language varieties [2, 5]:

- (1) The independent translations of David Lodge’s *Therapy* (DL1), translated by Maria do Carmo Figueira into PP and by Lúcia Cavalcante-Luther into PB;
- (2) The independent translations of Lodge’s *Changing Places* (DL3), translated by Helena Cardoso into PP and by Lúcia Cavalcante-Luther into PB;
- (3) The independent translations of Jostein Gaarder’s *Vita Brevis* (JG), translated by Pedro Maia Soares into PB and Maria Luísa Ringstad into PP;
- (4) The independent translations of Karl Ove Knausgaard’s *Min Kamp 1* (KOK) into PP by João Reis and into PB by Leonardo Pinto Silva;
- (5) The adaptation of the book *Os livros que devoraram o meu pai* (AC) by Afonso Cruz, a Portuguese author, into PB.

Table 1. Quantitative overview: *alu* stands for alignment unit, *alu emo* means alignment unit with an emotion, *rev.* For revised

Text	No. of words	No. of alu	alu emo	emo rev.	Alu emo	rev. Emo
AC _{PP}	18,946	1,261	225	264	246	301
AC _{PB}	18,851		234	272	251	303
DL1 _{PP}	17,132	652	221	290	230	311
DL1 _{PB}	17,076		214	285	230	311
DL3 _{PP}	10,891	417	139	203	147	227
DL3 _{PB}	11,512		123	179	147	226
JG _{PP}	3,802	215	64	91	63	94
JG _{PB}	4,001		70	95	66	93
KOK _{PP}	7,026	309	57	83	64	101
KOK _{PB}	7,062		61	85	65	98
Total _{PP}	57,797	2,854	706	931	748	1,042
Total _{PB}	58,502		702	916	760	1,031

3.2 The Revision Process

The revision process showed clearly that agreement on what constitutes an emotion is not straightforward. In fact, the first human revision produced an annotation that added about 30% more emotion occurrences to the automatic one. However, a second revision, done by those responsible for the automatic revision, reduced the emotion surplus significantly. The main recommendations are reported here:

- in all cases where an MWU with the support verb *sentir* ‘to feel’ was employed, only one emotion (and not two) should be marked, and the same was true for *sentimento de X* or *sensação de X* when X was an emotion, corresponding respectively to *feeling* or *sensation* in English.
- the same should be the case when *sorriso* ‘smile’, *riso* ‘laughter’ or *choro* ‘cry’ were modified by an emotion word.
- words representing feelings used conventionally in other settings, like *sofrer* ‘suffer’, should not be marked as emotions.
- physical descriptions that are not conventionally used to refer to a feeling should not be marked as emotion. Still, some cases depend on the reader’s interpretation, as in *tapou a cara com o jornal* ‘she covered her face with the newspaper’, which was interpreted as describing shyness.
- luck, bad luck, success, and failure should not be considered emotions (although there can be other opinions on that).

There were still a few cases of meaning differences that had to be decided in a third round. But, in general, if there was disagreement, we did not mark it as emotion. The

only exceptions, because they were very frequent in the material, and which were classed in the OTHER emotion bin, were the many verbs that mean *make fun of* in Portuguese. One could consider this action as bringing shame to the patient and satisfaction to the agent, so one could argue that two emotions are actually involved.

The three previously described revisions looked independently at both sides of the corpora. Then we looked at the pairing and were able to find some instances of inconsistency across the collections, as well as cases that had been overlooked. It is the result of this fourth – and final – revision that is presented in this paper.

3.3 First Findings

Table 1 shows that we had at our disposal 2,854 alignment units (each of which corresponds to one sentence in the original text), and that 760 PB and 748 PP alignment units have at least a word referring to an emotion after revision. It also shows that, in the end, human revision added 12% of emotions compared to those found by the automatic system. (To be more precise, this is just the net result. Many cases of “automatic” emotions were removed, and therefore, many more emotions were actually added.)

There are some differences between the texts, with DL3 having the highest percentage of alignment units with at least one emotion (35.3%), and AC the lowest (19.5%), while for the average number of emotions per emotion-encoding alignment unit, the winner is KOK (1.58), with the lowest again being AC (1.23). It was surprising to observe that an original text in Portuguese had fewer emotions than any of the translations, but the sample is too small to generalize.

Another unexpected finding was the considerable number of cases in which the number of emotions referred to were not the same in the two halves, as can be seen in Table 2. Therefore, in 252 cases (out of 2,854, thus 9%) one translator referred to an emotion and the other did not. While some examples can be explained by different discursive strategies, most of them “say the same thing” in a different way, such as in example (1).

- (1) *PP* - *e diz **esperar** que a viagem lhes tenha **agradado*** ‘and he says he hopes that the trip has pleased them’
PB - *e ele **desejava** que tivessem tido uma boa viagem* ‘and he wished that they had a good trip’
*EN*³ - *and he **hopes** they have **enjoyed** the flight*

Table 2. Difference in number of emotions per alignment unit that featured at least one emotion

Difference	0	1	2	2	-2	3	-3
Number of alus	578	115	107	7	5	1	1

³ EN is the original text.

After restricting our counts to the 578 cases in which the number of emotions were the same, we detected 483 alignment units (corresponding to 589 emotions) with the same group of emotions (84%). Of those, only 321 alignment units had exactly the same emotion words. After careful scrutiny, since several of the differences are due to trivial changes in person, tense, the presence of clitics, and even diathesis, we found out that, of all the cases expressed by the same emotion group, 195 involved different lexemes. This would suggest that this method is a good source of (at least) inspiration for paraphrasing.

In addition to the above-mentioned surprisingly frequent addition or removal of emotions in the alignment units, several (78) alignment units with different emotion groups were found, as example (2) illustrates.

- (2) *PP - só teria confirmado os seus profundos **receios***
*PB - apenas confirmado suas **suspeitas** mais profundas*
*EN - merely confirmed his deepest **misgivings***

It should be noted that *misgivings* is a very difficult term to translate into Portuguese, and that the two translators produced satisfactory, albeit different, results, with *receios* ‘fears’ in PP and *suspeitas* ‘suspicions’ in PB.

Example (3) shows how difficult it is to delimit emotion groups: one could argue that *liberdade* ‘freedom’ is ‘release from pressure’, and therefore can be equivalent, in a way, to *alívio* ‘relief’. However, we did not consider the two as belonging to the same group of emotions.

- (3) *PP - Foi igualmente a expressão de um **sentimento de liberdade** porque, enfim, podíamos agir livremente num país longe de Mónica.* ‘it was also the expression of a **sense of freedom** since we, at last, could act freely ...’
*PB - Acho também que pode ter sido uma expressão de **alívio**, porque final-mente podíamos nos mover com liberdade em uma terra distante de Mônica* ‘I think it may have been an expression of **relief**, because we could finally ...’

3.4 Multiword Units

Another striking feature we noted was the need to add (or take into consideration) multiword units (MWUs). In fact, although multiword units are marked and dealt with in Linguateca’s corpora (see, e.g., body annotation [8]), they are not extensively used for emotion. But, since MWUs are extremely relevant for paraphrasing, their presence was manually added in the human annotation, as examples (4) and (5) illustrate. Example (4) features the use of *fazer beicinho* ‘pout’, an MWU that expresses emotions of sadness, disappointment, and/or disagreement.

- (4) *PP - Ela **fez beicinho** e disse: – É que eu tenho uma ideia fabulosa para uma série muito original, uma espécie de versão inglesa do Twin Peaks.*
*PB - Ela **fez um beicinho** e disse: – O negócio é que tive uma idéia sensacional para uma novela diferente, tipo Twin Peaks inglesa.*
*EN - She **pouted** a little and said, «It’s just that I have this fabulous idea for an offbeat soap, a kind of English Twin Peaks.*

Example (5), involving three emotions on each side, is interesting for several reasons. First of all, it shows two cases in which there is an MWU on one side and one word on the other side of the aligned sentence. In Portuguese, the expressions *rogar* [] *pragas* and *dizer* [] *palavrões* are MWUs that express the emotion of anger and can be translated approximately as ‘beg for plagues’ and ‘use curse words’, respectively, and mean wishing for bad things to happen to someone. Then, it presents a case where the English formulation is so Germanic that it has no easy translation in Portuguese: *shock into silence*. Faced with this phrase, the two translators chose widely different solutions, which in fact refer to two different emotion groups: *chocar* ‘shock’ as a bad surprise, and *vociferação* ‘cursing’ as angry language. Finally, the example also illustrates a case in which one paraphrase only works in PB, since the verb *xingar* is only used in Brazil.

- (5) *PP - Arranquei a venda e pus-me aos saltos na sala, a rogartantas pragase a dízertantos palavrõesque acabei por ficar chocadoe me calei.*
PB - Arranquei a venda dos olhos e pulei de um lado para o outro ao redor da sala, xingando e blasfemando coisas tão terríveis que finalmente minha vociferação me fez silenciar
EN - I tore off the blindfold and hopped round the room cursing and blaspheming so terribly I finallyshocked myself into silence.

All the emotion paraphrases were subsequently aligned with CLUE-Aligner [3], depicted in Fig. 2. This tool, developed in the context of eSPERTo, creates entries that can be included in NooJ dictionaries and/or grammars.

From this alignment task, we were able to obtain a list of 463 emotion paraphrases (out of ca. 1,300 paraphrastic units). After removing repetitions and creative or divergent translations, we selected 191, which are publicly available.⁴

Let us explain here what we considered a (real) emotion paraphrase (or "paraphrastic unit," as we call it, since it is not a full sentence). Basically, this is a different way of saying something that does not involve a crucial difference in the emotion expressed, restricted in this case to having more than one word in one of the sides. So, cases like *medo* ↔ *terror* ‘fear’, ‘terror’, which are one-word synonyms, and which were also obtained by our procedure, are not counted here, nor are different morphological features of the same lexeme (such as differences in person or in tense).

3.5 NooJ Resources

The last step was the enrichment of the paraphrastic resources of Port4NooJ, the Portuguese module of NooJ, with the emotion information gathered, and then the application of those resources to feed eSPERTo. Although at the time of writing this article, we have only been able to do a few tests to explore the best way of formalizing this knowledge within NooJ by creating or updating relevant dictionary entries and grammars, some of our ideas can be presented here.

We identified five main groups of pairs of paraphrastic units that, due to their specific properties, need to be formalized in different ways in NooJ:

⁴ <https://www.linguateca.pt/Gramateca/ExemplosParafrasesEmocao.html>.

	isto	não	teria	causado	nenhuma	surpresa	a	Morris	Zapp	-	apenas	confirmado	suas	suspeitas	mais	profundas	.
isso																	
não																	
teria																	
surpreendido																	
Morris																	
Zapp																	
por																	
ai																	
além																	
,																	
só																	
teria																	
confirmado																	
os																	
seus																	
mais																	
profundos																	
receios																	
.																	

teria surpreendido | teria causado [] surpresa a (S)

profundos receios | suspeitas [] profundas (S)

Fig. 2. Using the CLUE-Aligner

None of the Words of the Pair of Paraphrastic Units Varies. The words do not inflect, are not permutable with other words, and to a certain extent they are frozen (with a greater or lesser degree of idiomaticity). In this case, for each paraphrastic unit of the pair, we will create one dictionary entry that will be associated with the corresponding paraphrastic unit of the pair. Then, we need to create a grammar that identifies one paraphrastic unit and rewrites that unit with all other paraphrastic units in pairs that include it. For example, the pair *à beira de um ataque de nervos* ↔ *com os nervos à flor da pele* ‘(X’s) nerves (BE) on edge’ will originate the following two dictionary entries:

- *à beira de um ataque de nervos*, N + EMO:NERVOSO + PARA = *com os nervos à flor da pele*
- *com os nervos à flor da pele*, N + EMO:NERVOSO + PARA = *à beira de um ataque de nervos*

One or More Words of Either Paraphrastic Unit in the Pair Vary. Pairs of paraphrastic units in which at least one word of one of the units can inflect or may be permutable with more than one word will be represented with a grammar (and subgrammars), which may or may not need to represent a wider context to be able to inflect the words of the paraphrastic unit or to permute them.

In the example *entreguei-me aos prazeres* ↔ *direcionei a minha atenção aos interesses* ‘I gave myself to the pleasures / I directed my interests to the pleasures’, the grammar does not need a wider context to make sure the verbs will have the same inflectional traits, or to represent the correct pronoun on both sides – the clitic pronoun *me* in PP and the possessive pronoun *minha* in PB.

On the other hand, for the pair *a transbordar de alegria* ↔ *doido de felicidade* ‘brimming over / crazy with happiness’, in order to properly rephrase the first paraphrastic unit, the grammar needs to represent the subject and accordingly inflect the adjective *doido* ‘crazy’.

The grammar that formalizes the paraphrastic unit *a ideia fez-me sorrir* (*emo:feliz*) ↔ *sorri*(*emo:feliz*) *com esse pensamento* ‘I smiled at the thought’, presented in Fig. 3, does not need a wider context to guarantee the same inflectional traits in the auxiliary verb *fazer* ‘make’ and the verb *sorrir* ‘smile’, but to make sure that the clitic pronoun agrees with the person of *sorrir*, it does. Also, the words *ideia* ‘idea’ and *pensamento* ‘thought’ are permutable.

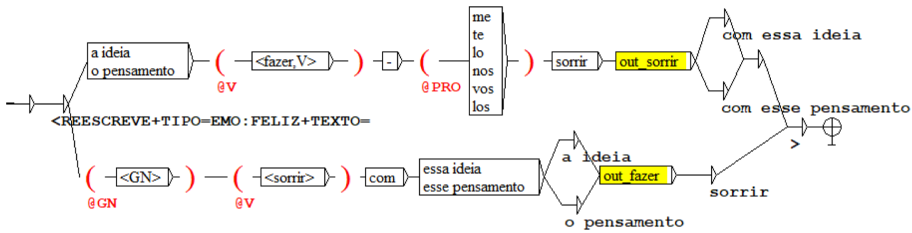


Fig. 3. Grammar to paraphrase the pair of paraphrastic units: *a ideia fez-me sorrir* (*emo:feliz*) ↔ *sorri*(*emo:feliz*) *com esse pensamento*

Several paraphrastic units include arguments which invert positions in the paraphrase. For example, the pairs *N0 achar N1 perturbante* ↔ *N1 deixar N0 desconcertado* ‘N0 be disturbed with N1’ and *N1 agradar minimamente a N0* ↔ *N0 estar razoavelmente satisfeito com N1* ‘N0 be ADV pleased with N1’ are symmetric.

Likewise, grammars can also be developed for constructions involving nouns related to body parts, fluids, or actions, such as *sorriso* ‘smile’, *riso* ‘laugh’, *olhos* ‘eyes’, *choro* ‘cry’, *ombros* ‘shoulders - shrug’, *sobrolho* ‘eyebrows - frown’, thus capitalizing from previous work on the interaction of body and emotion [16].

If one of the paraphrastic units in the pair involves a construction with a human intransitive adjective or nominal predicates supported by the verb *ser de* ‘be of’ or *fazer* ‘do / make’, already included and formalized in the corresponding lexicon-grammars integrated into Port4NooJ dictionaries, we will add the mark + EMO: subcategory to those entries in the dictionaries, which will make the grammar inherit the + EMO feature from the predicate and assign it to its paraphrases.

For example, in the pair *ficou muito envergonhada* ↔ *ficou muito constrangida* ‘became very shy’, the adjectives *envergonhada* and *constrangida* will add the emotion property + EMO: VERGONHA to the corresponding (simplified) dictionary entries:

- envergonhado, A + FLX = ALTO + AV + state + EN = bashful + DRV = AVDRV01:RAPIDAMENTE + EMO:VERGONHA + IH + Table = SEAH3 + Nhum + Vcopser = ser [...]
- envergonhado, A + FLX = ALTO + PChappyType + EN = ashamed + DRV = AVDRV01:RAPIDAMENTE + EMO:VERGONHA + IH + Table = SEAH3 + Nhum + Vcopser = ser[...] + UMNclas + UmModif + AdvQuant + Superlativo + NAdj + em-Vinf [...]
- constrangido, A + FLX = ALTO + IH + Table = EAHP3 + Nhum + Vcopestar = estar + Vcopandar = andar + Vcopficar = ficar + Vcopencontrarse + Vcopsentir-se + Vcopmostrar-se + Vcopprevalarse + AdvQuant + Superlativo + NAdj + DRV = A2N1:ANO + DRV = A2V2:DIRIGIR

In this way, the grammar that handles the alternation of the copulative verbs will generate paraphrases where Vcop will be *ser/estar/ficar/andar/permanecer/sentir-se/mostrar-se/revelar-se/tornar-se*. Finally, it will be necessary to develop new grammars not only to represent the pair of paraphrastic units, but also to extend the pair of paraphrastic units to other constructions involving that nominal or adjectival predicate. We should like to explain that, in our previous example, we want to be able to represent *ficou muito envergonhada* ↔ *ficou muito constrangida* as a pair of paraphrastic units, but also to demonstrate that we can alternate the copulative verbs in this pair of paraphrastic units and that, consequently, *estar muito envergonhada* ↔ *estar muito constrangida*, *sentiu-se muito envergonhada* ↔ *sentiu-se muito constrangida*, etc. are also pairs of paraphrastic units.

The Paraphrastic Units Are Variety-Dependent. When at least one of the paraphrastic units of the pair is mostly used in one of the Portuguese varieties, this needs to be marked with the Portuguese variety (PP or PB), thus enriching eSPERTO's paraphrastic resources across Portuguese from Brazil and from Portugal [18].

The Paraphrastic Units Are Approximate Paraphrases. The paraphrastic units are often not absolutely identical semantically [6], and present subtle or less subtle deviations of meaning that eSPERTO users need to be aware of to be able to decide whether to proceed with rewriting, and to choose the most suitable for the context or the most stylistically appealing to the user. For this reason, some paraphrastic pairs, such as *competir pela honra de* ↔ *competir pela oportunidade de* ('compete for the honour of' ↔ 'compete for the opportunity to') and *ter coragem para* ↔ *ter a cara de pau de* ('dare' ↔ 'have the courage to') should be marked with the feature + APROX.

4 Future Work and Conclusions

We intend to use the pilot study presented in this paper to further our research in the near future. One important task is to produce a more systematic analysis and annotation of MWUs and phrases expressing emotion in Emocionário's large corpora. Another is to design and build high coverage paraphrastic grammars in NooJ to deal with the following phenomena:

- **emotion-conveying idiomatic expressions**, which may correspond to a single word with identical meaning, e.g., *ter pena de* ‘(to) pity’, *sentir falta de* ‘(to) miss’, *encolher os ombros* ‘(to) shrug (one’s) shoulders’, *fazer beicinho* ‘(to) pout’, *(falar) com o coração nas mãos* ‘(to) speak with (one’s) heart in (one’s) hands’, *de cabelos em pé* ‘hair raising’;
- **emotion quantification**, e.g., *gostar muito*, paraphrase of *amar* ‘like a lot = love’; *não gostar nada*, paraphrase of *detestar* ‘do not like at all = hate’, also using **opposite emotions**;
- **combination of different emotions**, e.g., *sorriso triste* ‘sad smile’; *amor violento* ‘violent love’.

From the point of view of each of the initial projects, namely eSPERTo and Emocionário, the cooperation was very productive. For eSPERTo, a new set of paraphrases expressing emotions was collected and can be of use in language learning applications, for example. For Emocionário, it provided an independent evaluation of its procedures and of several new candidates, especially MWUs expressing emotion.

Acknowledgments. This work was partially supported by Portuguese national funds through FCT, Fundação para a Ciência e a Tecnologia, under project UIDB/50021/2020. We also thank UNINETT Sigma2 – the National Infrastructure for High Performance Computing and Data Storage in Norway – for use of their computational resources, as well as FCCN for hosting Linguateca in their servers. We thank Belinda Maia for having read a previous version and suggested considerable improvements in style and content.

References

1. Barreiro, A.: Make it simple with paraphrases: automated paraphrasing for authoring aids and machine translation. Ph.D. thesis, Universidade do Porto, Porto, Portugal (2009)
2. Barreiro, A., Mota, C.: Paraphrastic variance between European and Brazilian Portuguese. In: Zampieri, M., Nakov, P., Ljubei, N., Tiedemann, J., Malmasi, S., Ali, A. (eds.) Proceedings of the Fifth Workshop on NLP for Similar Languages, Varieties and Dialects (VarDial). COLING 2018, pp. 111–121. Association for Computational Linguistics (2018)
3. 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. LREC 2016, pp. 7–13. European Language Resources Association (2016)
4. Barreiro, A., Rebelo-Arnold, I., Baptista, J., Mota, C.: Parafraseamento automático de registo informal em registo formal na Língua portuguesa. *Linguamática* **10**(2), 53–61 (2018)
5. Barreiro, A., Rebelo-Arnold, I., Batista, F., Garcez, I., Kuhn, T.Z.: One book, two language varieties. In: Quaresma, P., Vieira, R., Aluísio, S., Moniz, H., Batista, F., Gonçalves, T. (eds.) PROPOR 2020. LNCS (LNAI), vol. 12037, pp. 379–389. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-41505-1_36
6. Barzilay, R., McKeown, K.: Extracting paraphrases from a parallel corpus. In: Proceedings of the ACL/EACL (Toulouse, 2001), pp 50–57 (2001)
7. Frankenberg-Garcia, A., Santos, D.: Introducing COMPARA: the Portuguese-English parallel corpus. In: Zanettin, F., Bernardini, S., Stewart, D. (eds.) Corpora in Translator Education, pp. 71–87. JSt Jerome, Manchester (2003)

8. Freitas, C., Santos, D., Carriço, B., Mota, C., Jansen, H.: O léxico do corpo e anotação de sentidos em grandes corpora – o projeto Esqueleto. *Revista de Estudos da Linguagem* **23**(3), 641–680 (2015)
9. Maia, B., Santos, D.: Language, Emotion, and the emotions: multi-disciplinary and linguistic background. *Lang. Linguist. Compass* **12**(5) (2018). <https://doi.org/10.1111/lnc3.12280>
10. Mota, C., Santos, D.: Emotions in natural language: a broad-coverage perspective. Technical report, Linguateca (2015). <https://www.linguateca.pt/acesso/EmotionsBC.pdf>
11. Mota, C., Carvalho, P., Barreiro, A.: Port4NooJ v3.0: integrated linguistic resources for Portuguese NLP. In: Proceedings of the Tenth International Conference on Language Resources and Evaluation LREC 2016. LREC 2016, pp. 1264–1269 (2016)
12. Mota, C., Chacoto, L., Barreiro, A.: Integrating the lexicon-grammar of predicate nouns with support verb fazer into Port4NooJ. In: Mourchid, M., Silberztein, M. (eds.) Mbarki S, pp. 29–39. Formalizing Natural Languages with NooJ and Its Natural Language Processing Applications, Springer (2017)
13. Mota, C., Baptista, J., Barreiro, A.: The lexicon-grammar of predicate nouns with ser de. In: Port4NooJ: 12th International Conference, NooJ 2018, Palermo, Italy, 20–22 June 2018, Revised Selected Papers, pp. 124–137 (2019)
14. Mota, C., Carvalho, P., Raposo, F., Barreiro, A.: Generating paraphrases of human intransitive adjective constructions with Port4NooJ. In: Okrut, T., Hetsevich, Y., Silberztein, M., Stanislavenka, H. (eds.) NooJ 2015. CCIS, vol. 607, pp. 107–122. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-42471-2_10
15. Ramos, B., Freitas, C.: “Sentimento de quê?” uma lista de sentimentos para a Análise de Sentimentos. In: STIL - Symposium in Information and Human Language Technology, Salvador, BA, 15–18 October, 2019, pp. 38–47 (2019)
16. Ramos, B., Santos, D., Freitas, C.: Looking at body expressions to enrich emotion clusters. In: Finatto, M.J.B., Luz, S., Pollak, S., Vieira, R. (eds.) 2020 Proceedings of the Digital Humanities and Natural Language Processing Workshop at the 14th International Conference on the Computational Processing of Portuguese Language, pp. 57–62 (2020)
17. Ramos, B.C.: Leitura Distante e categorização de sentimentos representados pelo corpo humano na Literatura (2019). <https://www.linguateca.pt/ELD/aprELDBarbaraRamos.pdf>
18. Rebelo-Arnold, I., Barreiro, A., Quaresma, P., Mota, C.: EP–BP Alinhamentos Parafrásticos PE–PB de Construções de Predicados Verbais com o Pronome Clítico ‘lhe.’ *Linguamática* **10**(2), 3–11 (2018)
19. Santos, D.: Corpora at Linguateca: vision and roads taken. In: Berber Sardinha, T., de Lurdes Sao Bento Ferreira, T. (eds.) Working with Portuguese Corpora, Bloomsbury, pp. 219–236 (2014)
20. Santos, D.: PANTERA: a parallel corpus to study translation between Portuguese and Norwegian. *BELLS* **10**(1), 1–13 (2019)
21. Santos, D., Maia, B.: Language, emotion and emotions: a computational overview. *Linguist. Lang. Compass* **12**(5) (2018). <https://doi.org/10.1111/lnc3.12279>
22. Santos, D., Simões, A.: Towards a computational environment for studying literature in Portuguese (2019). <https://www.linguateca.pt/Diana/download/PresentationBudapestSantosSimoes.pdf>
23. Silberztein, M.: Formalizing Natural Languages: The NooJ Approach. Wiley, Hoboken (2016)
24. Wierzbicka, A.: Emotions Across Languages and Cultures: Diversity and Universals. Cambridge University Press, Cambridge (1999)