

PANTERA: a parallel corpus to study translation between Portuguese and Norwegian

Diana Santos

Abstract. This paper presents an on-going project, PANTERA, dealing with the Portuguese-Norwegian language pair. The PANTERA project aims a) to identify all translations ever published between the two languages Portuguese and Norwegian, and b) make a sample of each available and searchable for the study of translation between the two languages, in the PANTERA parallel corpus. After describing the methodology and processing used to create the corpus, I discuss briefly its contents from a translation studies perspective, and proceed to give examples of its actual use in the context of linguistic and cultural studies, ending with its possibilities as a teaching aid.

The particular subjects discussed are the concept of respect in the two languages, the semantic field of body parts and the identification of possessive datives and null objects in Portuguese using the translations into Norwegian, as a technique to elicit complex syntactic phenomena based on contrastive patterns.

1. Presentation

This paper presents an on-going project, PANTERA, which aims to identify all translations ever published between the two languages Portuguese and Norwegian, and make a sample of each available and searchable for the study of translation between the two languages.

PANTERA stands for *Portuguese And Norwegian Texts for Education, Research and Acquisition of relevant knowledge...* and started in the autumn of 2013 as a cooperation between Linguateca and the University of Oslo. Although the corpus is relatively small, it is probably already one of the most diversified parallel corpora in the world, considering the diversity of its sources in number of authors, time stamps and genres, as we will presently show.

PANTERA is meant to foster the study of translation and of the similarities and differences between the two languages and is publicly available for search at <http://www.linguateca.pt/PANTERA/>.

2. Technical makeup

PANTERA's development was inspired by the English-Norwegian Parallel Corpus (ENPC) (Oksefjell 1999), and by COMPARA (Frankenberg-Garcia and Santos 2003). Just like other similar projects such as CorTrad and PoNTE, it uses DISPARA (Santos 2002) as the underlying system.

Building the corpus involves the following steps: we 1) revise the digitized material in the two languages (digitizing it first if not already electronically available), then 2) apply automatic alignment using the Open CWB workbench, 3) do human revision of the alignment and sentence separation, and 4) perform automatic (syntactic) annotation in the two languages. PALAVRAS (Bick 2000) is used for Portuguese and the Oslo-Bergen-tagger (Johannessen et al. 2012) for Norwegian. We also apply some semantic annotation (so far only) to the Portuguese part, described in more detail in e.g. Santos (2014).

3. The textual contents

As of October 2017, 371 text pairs have been identified for inclusion in PANTERA: 192 translations of Norwegian texts into Portuguese, and 179 translations of texts in Portuguese into Norwegian. A detailed description of these texts, including the sizes of the excerpts used in the corpus, is constantly updated at PANTERA's site as is therefore not necessary to freeze here. Further information about each translation instance (163 different authors and more than 167 translators) is also being stored in the STIG system, in progress (<https://stig.hf.uio.no/pantera/>). STIG, named after Stig Johansson, one of the pioneers of parallel corpora research, aims to be a platform for all interested in translation studies in general, and is being thoroughly tested in the case

of the Portuguese-Norwegian language pair.

<insert fig 1 here>

In figure 1 we give a visual presentation of PANTERA's contents, based on the publication date of the original text. Above the axis you can see the translations published in Portugal, below the axis the ones published in Brazil. In figure 2 we provide another presentation of the same data, this time marking the (first) translation date.

<insert fig 2 here>

Although this may seem a trivial piece of information for the few who have tried to create corpora, not everything is equally easy to ascertain. Let me give some examples: there are some cases where the original language of the texts is not the native language of the author. So, for example the text *Contos e lendas dos vikings* by Lars Haraldson, was initially presumed to be deriving from a (published) Norwegian original text, but *Contes et legends. Les vikings*, 2002 was in fact only published in French by Nathan, Paris. And Thor Heyerdahl's *Sjøveier i Polynesia*, although published in Norwegian by Gyldendal in 1968, had been originally published in German, of which both the Norwegian and the Portuguese versions were translations.

And there are also those special cases where a particular piece of national literature is written in another language, as is probably the case of the *Kjærlighetsbrev fra den portugisiske nonne søster Marianne til Grev de Chamilly*, written... in French. Or, when specific originals are translated into a language in which they are displayed or performed, but the original text is not: This seems to be the case of *Nupen: viagem sem fim* by Per Hovdenakk, published in 2001, by Galleri Wang, Oslo, and the play by Arne Ingolv Sunde Lygre translated as *Homem sem rumo*, published both by SESC Avenida Paulista in São Paulo in 2007, and by Teatro da Comuna, in Lisboa in 2008.

None of these cases were therefore included in PANTERA, although they may find their way into STIG.

Then, the proper identification of a text pair can also be complicated, in cases merging the translation of more than one source book into one target book, or selecting parts of a book into an independent translation. Examples are *Vitória; O sonhador*, by Knut Hamsun, published by Boa Leitura editora of São Paulo in 1961, -- whose originals come from *Victoria*, 1898 and *Sværmere*, 1904. Conversely, Hamsun's *A morte de Glahn*, published by publisher Itatiaia, of Belo Horizonte, was, after some research, identified as ... the epilogue to *Pan*, 1894.

We included these cases in PANTERA, as two distinct translations in the first case and one further translation of Pan in the second. But one might have done the opposite, following Toury's spirit of the primacy of the translated object, see e.g. Toury (1995). One could have used the translated objects as the norm, and have defined one more work by Hamsun in the latter case, and one book/one translation from two sources in the former.

Finally, it is well known that, especially in the past, publishers may omit crucial information about the way translation was done: sometimes it is just our knowledge of the individual translators that can make us posit an intermediary language. Still, even nowadays it can happen that not even the translator's identity is revealed, as was the case with Adriana Lisboa's *Snø fra Brasil* published in 2015 by the publisher Tigerforlaget.

When we are reasonably sure that the original texts are in the right language, we decided to include also indirect translations in PANTERA, specifying whenever possible the intermediate language(s). Much more information on this detective work will be present in STIG, where the sources for the

conjectures can be included.

Another borderline case are bilingual (or multilingual) editions. When it is clearly indicated which is which and who the translator is, we opted to include them as well in PANTERA, as was the case of the book *Bacalao-bacalhau-baccalà* published by the publisher Orkana in 2003 in a Norwegian-Portuguese bilingual edition (although with a trilingual title).

This is not the place to develop the topic of analysis of the translation between the two linguacultures, but some short remarks are in order here.

First, translation into Portuguese clearly preceded translation into Norwegian (as an anecdote, the first translation ever into Norwegian of a Portuguese original seems to be due to its author living in Oslo as a diplomat at the time).

Second, the choices of works to be translated into Portuguese seem to follow the general European and American trend – the classics, the Nobel prizes, the texts who have shown to be good market in other countries, so that the translation pattern is very similar in Brazil and in Portugal.

Third, translation into Norwegian (differently from into Portuguese) seems to have been boosted by a cultural actor (Leif Sletsjøe, professor of Portuguese language and literature at the University of Oslo), who translated works he deemed culturally invaluable, and not due to market pressure, as is the case of five plays of Gil Vicente, an important Renaissance playwright. After Sletsjøe's death, the pattern of translation into Norwegian became very much the same as into Portuguese: Writers with Nobel and other prizes and/or bestsellers, plus a special interest in football.

Finally, the bulk of published translation in either direction concerns fiction, something which is also the case for other foreign languages in Norway and in Portugal. There are few books or texts otherwise, but it is interesting to note that the choices are different: into Portuguese, it is mainly travel and history, with one example in law and the aforementioned cooking book; into Norwegian, football and politics, only one concerning history. Given the importance of Portuguese history for global history, this gap is specially noteworthy.

It is our hope that PANTERA, and especially the possibilities offered for searching and adding information to translation process in STIG, will foster considerable research and study in both directions in the future.

4. Using the corpus for linguistic research

After a short overview of the translation practice between the two languages, I turn to the use of the aligned language material to do linguistic research, illustrating its potentialities with three different case studies. Note that the actual number of corpus pairs is significantly lower (so far) compared to the listing of all works identified, so I indicate the size of the corpus and the query data.

For the studies reported here, the corpus comprised 132 text pairs, 51 in the Portuguese-Norwegian direction and 63 in the Norwegian-Portuguese direction. In total this corresponds to **40,607** alignment units (uas), and **684,849** tokens (words and punctuation marks) in Portuguese and **648,317** in Norwegian. Tables 1 and 2 show the distribution in terms of language variety in original and translated texts respectively.

Table 1. Distribution of the source corpus material by language variety (or målform)

Variety	Size in tokens
angolana (AN)	5,442
bokmål (BM)	288,916

brasileira (BR)	96,261
caboverdiana (CV)	4,724
dansk-norsk (DN)	61,055
guineense (GB)	14,379
moçambicana (MZ)	3,450
nynorsk (NN)	19,885
lusa (PT)	104,863
timoreense (TL)	21,358

Table 2. Distribution of the translated corpus material by language variety (or målform)

Variety	Size in tokens
bokmål (BM)	272,921
brasileira (BR)	75,516
nynorsk (NN)	5,322
lusa (PT)	358,563

Table 3. Distribution of the corpus material by genre of the original texts

Genre	Size in tokens
Novel (romance, roman)	306,644
Drama	113,011
Non-fiction	88,207
Crime	77,509
Children-young	61,306
Small novel (novela, nouvelle)	18,597
Short story	11,774
Autobiography	4,406
Essay	3,104

4.1. Does *respeito* (PT) mean the same as *respekt* (NO)?

My intuition as native speaker of Portuguese and speaking Norwegian as a foreign language, is that *respeito* and *respekt* do not mean the same, and that they rather illustrate well culture differences reflected in language. But how can one answer such a question based on linguistic data?

PANTERA may aid us in that quest. One may find that this semantic field is much more used in one of the languages, and/or with different meanings (actualized by different translations), and/or in different contexts and co-texts...

By using the following query

```
[lema="respeitável|respeito|respeitar|respeitinho|respeitado|respeitoso|desrespeito|desrespeitar|respeitavelmente|desrespeitoso|desrespeitosamente" & sema="emo.*"]
```

which selects all words morphologically associated with *respeito* in Portuguese and requires that they have been tagged as emotion, we get 70 cases. (If one did not care about the semantic annotation, one would get 130 cases, because of the common expressions *a respeito de* (about) and *dizer respeito* (concern). This is why, incidentally, that semantic and syntactic annotation are so relevant.) These boil in fact to even less than 70 cases, for in some cases (n-to-1 alignment) the translated sentence is repeated n times.

By doing a simple distribution by whether the occurrences come from original text or translated, we

get the following result in Table 4:

Table 4. Distribution of Respect-words in Portuguese by translated/original text

trad	38	444919
ori	32	250592

So, there are 1.28 in 10,000 words in original text in Portuguese, and only .95 in 10,000 words in translated text (if we remove the 4 duplicates, we get .76, but to be perfectly accurate we should also have removed the duplicate text, so we will go on with the slightly inflated .95 number.)

But this does not seem a very significant difference, and it is not statistically significant. If we take into consideration the distribution of these references, we see that this semantic notion occurs in 19 different texts in Portuguese, and in 18 different ones in Norwegian. But... does it? What if it is the translation into Portuguese that adds the respect?

If we look into the Norwegian language side, by searching *.*respekt.**, we get only 48 cases, and they are by far on the translated side, as can be seen in Table 5.

Table 5. Distribution of Respect-words in Norwegian by translated/original text

trad	34	444919
ori	14	250592

But not even here the difference is statistically significant. But they correspond to 10 sources in Norwegian, and 20 in Portuguese.

Let us therefore recapitulate: 32 out of 51 texts (text excerpts) written originally in Portuguese have some mention to respect, while only 14 out of 63 texts (text excerpts) written originally in Norwegian use it. Now, this difference is highly significant.

If one looks in addition at the contexts where these “typically” Norwegian-language cases appear vs. those typically Portuguese-language cases appear, the difference, I claim, gets substantiated.¹

For lack of space I present only two examples:

1. O estádio é imponente, verdadeiramente monumental, e, a regurgitar de gente, impõe respeito. -> Wembley Stadion er imponerende, virkelig monumental og, når den er full av mennesker, helt overveldende.
2. Hvem er det som våger det hun gjør? Sende ungen sin alene over fjellene for å møte pappaen sin? Ikke mange. Det står det respekt av. -> Poucos. Isso impõe respeito.

These two examples use the very same expression in Portuguese (*impor respeito*). In Portuguese it is properties and the essence of things that invokes respect; in Norwegian, it is actions that do it. In other words, respect is an essential property in Portuguese, while in Norwegian it is a temporary feeling.

Of course two examples are not enough to demonstrate this. A lot more can and should be done to describe this field, not least the related emotions and/or lexical items that can be identified through close reading of these examples and by observing the other translation choices involved. This will have to be dealt with in a different occasion.

¹The full list of translation examples (created 12 August 2017) can be found at <http://www.linguateca.pt/Diana/download/RespeitoPANTERA.pdf>.

4. 2. Conceptualization of body parts in the two languages: the case of *dedo* (PT) versus *tå* and *finger* (NO)

Germanic languages use separate words for naming the extremities of human limbs, while Romance languages use the same word, which can be distinguished by a specifier if needed. A current explanation for this vagueness in Latin and the derived languages compared to Germanic ones is the geographical, and thus climatic difference. Colder places have more detailed descriptions for the parts of the body they have to clothe, as claimed by Witowski & Brown (1985).

No matter whether this is the cause for this distinction in the particular language pair we are concerned here, it is undeniable that different lexical resources imply different language practices.

This is what a cursory examination of the words related to fingers and toes shows in both languages: [lema="tå"] gives 8 hits, only two of them from a Portuguese original, and one from an idiom, *da cabeça aos pés*, (fra topp til tå) which literally means from the head to the feet. So, of the six cases which referred to actual toes, only two had the specification *dos pés*. There is one case of *tånegler* which is similarly rendered only by *unhas* (nails).

The verb *tå* in Norwegian is not translated by anything related to fingers in Portuguese. Either it gets rendered by *em bicos dos pés* or *de fininho.*, expressions with the same meaning (describing a form of walking) but not the same lexical logic.

As for [lema="finger"] there are 99 results, shown in Table 6.

Table 6. The lemma *finger* in PANTERA, in original or translated Norwegian

ori	66	250592
trad	33	444919

And here it is crystal clear that again fingers are much more referred to in Norwegian originals than in translations. There are also 8 additional cases of compound words involving finger as first element of the compound: *fingerupp* and *fingerspiss* (ponta dos dedos, “finger tip”), *fingerknokke* (nó dos dedos, “finger knot”) and *fingervante* (a sort of gloves), as well as 27 cases of finger as the last element: mostly specification of which finger, and which often do not include the word *dedo* in Portuguese (*indicador, mindinho, médio, polegar, dedão, anelar*). Again, it is interesting to note is that in as many as nine cases the Norwegian text specifies which finger while in the Portuguese text appears only a reference to *dedo*. And there is one example of *dedo de fogo* which is translated by *ildfinger*.

If we look from the other direction, [lema="dedo"], among the 181 hits we can point some interesting cases. First, the fact that Norwegian has also totally different words for parts of fingers, which in Portuguese are rendered by a noun compound as the ones mentioned above (examples 1-2). Second, the well known lexical richness of Germanic languages associated to manner of movement, and description of sight and sound, which forces the translator to use an explanation and not a corresponding lexical item:

1. som klemte om armlenene så **knokene** hennes hvitnet -> que apertava os braços do assento com tanta força que os **nós dos dedos** se empalideceram
2. Den var rød over **knokene** -> Tinha as **articulações dos dedos** vermelhas
3. Sarah massajou os olhos com os **dedos** -> Sarah **gned** seg i øynene
4. mørk, ser hen for seg og **trommer** på bordet -> triste, olha em frente **tamborilando os dedos** sobre a mesa

5. for nå **pekte** han opp på meg igjen. -> **Apontou o dedo** para mim
6. **pekte** på dem som fikk komme inn. -> e **indicavam a dedo** os que tinham direito a entrar
7. Han **drodlet** på blokka ved siden av -> **Tamborilou com os dedos** no bloco que tinha perto.
8. I bilen hjemover **trommet** han på rattet -> No carro, no caminho de volta, ia **tamborilando com os dedos** no volante
9. han stirret på Peter, vantro, **pekte**, hele hånden skalv. -> ele fitava Peter, incrédulo, e **apontou-lhe o dedo**, a sua mão tremendo por completo
10. o piano coberto com um chale espanhol, lembrança de grandezas passadas, **gemia nos dedos** de Magda as notas de «Prima Carezza», -> Magda **klunket** i vei med Det Første Kjærtegn

In cases 11 and 12, the translator in fact adds specification that one could argue unnecessary:

11. Det kunne være godt å ha noe å **holde** i -> Seria agradável ter algo a que **agarrar os dedos** (something to hold gets translated into something to get your finger into)
12. Hun **dreide** vannglasset **i hånden**. -> **Girou** o copo de água **entre os dedos**. (turn in the hand gets translated into turn between the fingers)

Conversely, and also in line with the general spirit of the language, mentions of fingers in Portuguese are often abstract, metaphorical or conventional:

1. trocar dois **dedos** de prosa -> veksle noen ord med ham (dois dedos = a little)
2. não aceita qualquer, para ele as escolhidas a **dedo** -> han valgte dem ut med omhu (escolher a dedo: = choose carefully)
3. filhos sem pais a quem ensinam a apontar com o **dedo** o nosso retrato -> som de lærer til å peke på bildet vårt (apontar com o dedo = recognize)
4. que ficam a chuchar no **dedo** que se lixam. -> siden kan dere bli blå for at dere skal få noe mer av denne dama (chuchar no dedo = suck as a compensation for something you did not get)

However, there are also a few cases where the image is more specific in Portuguese than in Norwegian, going therefore against the general pattern. So, a very common gesture in both cultures, is described as doing something to the fingers in Portuguese² and to the hands in Norwegian in 5., just like in English “fold hands”. In 6, the position of a bird in Tereza's fingers turns into Tereza's hand.

5. Em seguida, vi-a endireitar a cabeça, cruzar os **dedos** e -> Om litt så jeg henne løfte hodet, folde hendene
6. Logo o pássaro, nos **dedos** de Tereza, -> Snart reiste fuglen seg stolt i Terezas **hånd**,

4.3. Studies of missing grammatical constituents

Corpora are notoriously hard to use to find missing things, and there is actually disagreement how to deal with these issues. However, when one knows that missing constituents in one language are obligatory in another, or when one phenomenon which is difficult to search for in one language is explicitly marked in another, one can use parallel corpora to ease the search (and check as well if the hypothesised contrastive difference holds).

This is what has been done by Jansen (2016) to find null objects in Portuguese. She searched

² The more common expression in Portuguese is *cruzar as mãos* (literally “cross hands”), which also makes reference to the hands, but this alignment unit demonstrates that also the expression *cruzar os dedos* is possible.

PANTERA for translation units with pronominal objects in Norwegian and without in Portuguese, and identifying which cases corresponded to null objects. Although I will not discuss here the complexity of the issue, it should be obvious that in addition to wrong parsing in either of the two languages, or wrong alignment, there would also be cases of interruptions, of different lexical items that require different constituents, and – the most common of all – use of a non-pronominal object.

Some (new) examples are shown in what follows, obtained from searching PANTERA by [lema="det" & func="obj"] and looking at the Portuguese counterparts (1765 cases). Cases 1 to 4 (from translation from Norwegian to Portuguese) and 5 to 8 (from Portuguese to Norwegian) are clear cases of null object in Portuguese.

1. Men jeg hadde ikke kommet hvis jeg ikke var nødt, **det** håper jeg du skjønner. -> Mas não teria vindo, se não fosse necessário. -- Espero que compreendas.
2. Jeg håper for Eddies skyld at dere finner han som gjorde det, **det** håper jeg virkelig. -> Pelo Eddie, espero que descubram quem cometeu esse crime, espero sinceramente.
3. -- Ja, skjønner De ikke **det**? -> Pois. Não está a perceber.
4. -- Men jeg liker **det** jo! sa jeg og dro margarinen over brødsboksen med kniven. -- -- Mas eu gosto! -- disse eu, pondo com a faca margarina no pão.
5. -- Só lá é que poderei saber. - **Det** kan jeg ikke vite før jeg kommer dit.
6. verá o pedreiro traçando o seu sinal pessoal, suavemente, batendo no escopo a jeito de não torcer o risco, nada mais fácil. -> Han vil se for seg steinhuggeren som meisler inn sitt personlige merke, hamrer varsomt for ikke å risse **det** opp skjjevt, en ganske enkel sak.
7. Estamos entre o Minho e o Douro, eu sei! -> Vi er mellom elvene Minho og Douro, jeg vet **det**.
8. Outras vezes juntávamo-nos na praia para nadar melhor do que os outros e deixar o sol queimar quem mais merecesse. -> Andre ganger møttes vi på stranden for å svømme bedre enn de andre og la sola varme den som fortjente **det** mest.

On the other hand, cases 9 to 12 are examples of different syntactic structures in which although there is a *det*-object in Norwegian there is no null object in Portuguese.

9. Jeg fortalte **det** ved kveldsbordet, og det ble stille. -> À mesa do jantar o silêncio caiu quando reporteí o incidente.
10. og at hun hadde **det** best hjemme i Skien med mor og far. -> e que ela estaria melhor em Skien junto com a mãe e o pai.
11. -- Jamen tror jeg ikke at han mener **det**, mumlet sirkusdirektøren for seg selv. -> -- Macacos me mordam se ele não está a falar a sério -- disse o director do circo para si mesmo.
12. Mamma har kanskje ikke ønsket noe annet for meg enn at jeg skulle få **det** bra... -> A minha mãe talvez só tenha querido que tudo me corresse bem?

This is offered here as an illustration, of course. More thorough work would have to consider the distribution of all possible cases obtained with this simple query, of which probably just 10-20% would be real null objects.

A previous study, published in Santos (2015), concerned the identification of possessive datives in Portuguese, a syntactic phenomenon who cuts across language families and exists for example in Portuguese and German, but not in French, English or Norwegian. Although the contrastive study published in the aforementioned paper was mainly based on the contrast with English, some data were also presented on Norwegian³, and – after having considerably extended PANTERA (instead of the 80 cases reported, we get 1002), we can reapply the method and find additional data for possessive datives.

3 The data is available from <http://www.linguateca.pt/Diana/dados/datposs/>

Again as an illustration, I present some new translation pairs, from Portuguese to Norwegian:

1. -- Em vez de ter morrido numa cruz, por ti, antes tivesse pegado na lança que **me** abriu o peito, para com ella te rasgar os olhos da cara. -> -- Snarere enn å ha dødd for deg på korset, skulle jeg ha grepet spydet som åpnet opp brystet **mitt**, for med det å stikke øynene ut av hodet på deg.
2. não dispõe de bases económicas que **lhe** garantam a tomada do poder. -> rår ikke over det økonomiske grunnlag som kan garantere **deres** maktovertakelse.
3. Não o via, mas ouvia-**lhe** a voz ali ao pé de mim, -> Jeg så ham ikke, men hørte stemmen **hans** like ved meg.
4. E a melindrosa desconhecida largou-lhe o braço com delicadeza, e retirara-se, apertando-lhe a mão. -> Og den ukjente slapp armen **hans** nesten uten at han merket det og bød ham hånden til avskjed.
5. O público invadiu o campo e levou-nos em triunfo, depois de **nos** ter feito em pedaços as camisolas. -> Publikum stormet inn på banen og bar oss bort i triumf etter å ha stykket opp trøyene **våre** i småbiter.

And from Norwegian to Portuguese:

1. Bernt Helle la hånden på underarmen **hennes**, og ga henne et flyktig kyss på kinnet. --> Bernt Helle pousou-**lhe** a mão no antebraço e deu-lhe um beijo rápido na face.
2. Han grep etter hånden **hennes** da hun gikk forbi ham for å sette seg i den andre sofaen. -> Agarrou-**lhe** na mão quando ela passou para ir sentar-se no outro sofá.
3. Litt for mye akevitt som nå lå over panna **hans**, som en kjetting. -> Um pouco de aguardente a mais que agora **lhe** pesava na frente como uma corrente.
4. På en måte stjeler det tankene **mine** å vite at Norge er medlem av Australia Group. -> Saber que a Noruega faz parte do Australia Group rouba-**me** de certa forma alguns pensamentos.
5. Hun fant seg noe å se på mens Dordi stemplet kortet **hennes** og spurte om hun lånte for seg selv, eller for mora. -> Encontrou algo para onde olhar enquanto Dordi **lhe** carimbava o cartão e lhe perguntava se estava a levar os livros para si ou para a mãe.
6. Han ble oppglødd i ansiktet, småløp ut av dusjen mot benken med bagen, bøyde seg -- og jeg så hvordan lårmusklene stramma seg under rumpa mens det dryppet fra kroppen **hans** -> Ele se animou, correu para o banco onde estava a bolsa, se inclinou -- e vi os músculos da coxa se contraírem embaixo da bunda enquanto **lhe** pingava água do corpo

5. Using PANTERA in teaching

Finally, I mention the use of the corpus for the semi-automatic creation of exercises and other teaching materials with the Ensinador paralelo tool described in Santos & Simões (2015), in the context of my daily work as university teacher of Portuguese for Norwegians.

I use selected concordances to illustrate, with authentic and varied examples, contrastive issues which are not well described, or even at all, in Portuguese or Norwegian grammars – precisely because they are more interesting contrastively.

Let me provide here some examples:

- modals in Norwegian, and their frequent correspondence to the subjunctive mood in Portuguese;
- verbs of perception and imagination (see Santos, ;
- sound and appearance (see Snell-Hornby, 1983);
- movement and position (Talmy, 1983)
- description of some emotions

I also use some of these concordances as exercises for further training in translation, but given that I

want to emphasize that there is not one single translation possibility in general but several, the criteria to choose the exercises are more complex and will not be discussed here. Anyway, I hope that this paper will allow others to use PANTERA also for teaching purposes, since everything is available.

References

- Bick, Eckhard. 2000. *The Parsing System "Palavras": Automatic Grammatical Analysis of Portuguese in a Constraint Grammar Framework*. Aarhus: Aarhus University Press.
- Frankenberg-Garcia, Ana, and Diana Santos. 2003. "Introducing COMPARA, the Portuguese-English parallel translation corpus". In *Corpora in Translation Education*, edited by Federico Zanettin, Silvia Bernardini and Dominic Stewart (eds.), pp. 71-87. Manchester: St. Jerome Publishing.
- Jansen, Heidi. 2016. "Objeto nulo no português - Observações sobre a sua problemática". Master thesis, University of Oslo.
- Johannessen, Janne Bondi, Kristin Hagen, André Lynum, and Anders Nøklestad. 2012 "OBT+stat. A combined rule-based and statistical tagger." In *Exploring Newspaper Language. Corpus compilation and research based on the Norwegian Newspaper Corpus*, edited by Gisle Andersen, pp. 51-65. John Benjamins Publishing Company.
- Oksefjell, Signe. 1999. "A description of the English-Norwegian parallel Corpus: Compilation and Further Developments". *International Journal of Corpus Linguistics* 4, no. 2: 197-216.
- Santos, Diana. 1999. "Perception verbs in English and Portuguese". In *Corpora and Cross-linguistic Research: Theory, Method, and Case Studies*, edited by Stig Johansson and Signe Oksefjell (eds.), pp. 319-342. Amsterdam: Rodopi.
- Santos, Diana. 2002. "DISPARA, a system for distributing parallel corpora on the Web". In *Advances in Natural Language Processing (Third International Conference, PorTAL 2002, Faro, Portugal, June 2002, Proceedings)*, edited by Elisabete Ranchhod and Nuno J. Mamede, pp. 209-18. Springer.
- Santos, Diana. 2014. "Corpora at Linguateca: Vision and roads taken". In *Working with Portuguese Corpora*, edited by Tony Berber Sardinha, and Telma de Lurdes São Bento Ferreira, pp. 219-236. Bloomsbury.
- Santos, Diana. 2015. "Os possessivos estão-me a complicar o ensino ;-) Um estudo do dativo possessivo baseado em corpos". *Linguística : Revista de Estudos Linguísticos da Universidade do Porto*, 10, 2015, pp. 107-130.
- Santos, Diana, and Alberto Simões. 2015. "Ensinador paralelo: Alicerces para uma pedagogia nova". In *Linguística, Informática e Tradução: Mundos que se Cruzam. Homenagem a Belinda Maia*, edited by Alberto Simões, Anabela Barreiro, Diana Santos, Rui Sousa-Silva and Stella E.O. Tagnin, pp. 235-52. Oslo: OSLa, Vol 7, No 1 (2015).
- Snell-Hornby, Mary. 1983. *Verb-descriptivity in German and English: A contrastive study in semantic fields*, Carl Winter Universitätsverlag, Heidelberg.
- Talmy, Leonard. "How language structures space". In H. Pick & L. Acredolo (eds.), *Spatial orientation: theory, research, and application*. New York, Plenum Press, 1983, pp. 225-282.
- Toury, Gideon. *Descriptive translation studies and beyond*, Amsterdam: Benjamins, 1995.
- Witkowski, Stanley R. & Cecil H. Brown. 1985. "Climate, clothing and body-part nomenclature". *Ethnology* 24, 3, 1985, pp. 197-214.