

Looking at body expressions to enrich emotion clusters

Barbara Ramos Diana Santos Cláudia Freitas

`barbaracmpramos@gmail.com`

`d.s.m.santos@ilos.uio.no`

`maclaudia.freitas@gmail.com`

3 March 2020

Two projects

In Linguateca we have been developing semantic annotation of corpora in Portuguese for some time. In particular

Esqueleto How is the body represented and used in Portuguese

Emocionário How are emotions referred to and used in Portuguese

In this paper, we set up to investigate possible connections between the two areas, and the two projects.

Studying the body in language

- a lexicon of body parts and expressions
- classified as to which (major) part of the body they pertain
- classified as to many types of uses, such as `corpo:sentimento`, `corpo:vegetal`, `corpo:faculdade`, `corpo:opinião`, etc.: ten categories plus one “others”

See Freitas et al. (2015) for further details and Freitas (2015) for extensive documentation for each category. See also <http://www.linguateca.pt/acesso/Esqueleto/>

Para eles, a sociedade não tem ouvidos[sema="corpo:faculdade"]
nem olhos[sema="corpo:faculdade"]; tampouco tem coração[sema="

Quem está de cabelo em pé[sema="corpo:sentimento"] é o zagueiro
João Carlos

Se a criança apresenta pé chato[sema="corpo:doença"], o uso da
palmilha não fez bem

- First phase: lexicon and first clusters (Santos & Mota, 2015)
- Second phase: slight improvements and more clusters (Santos et al., 2018)
- Third and current phase: critical assessment and reorganization as one of the aims of Barbara's PhD, involving the creation of *Emocionário*, a precise documentation of the emotion annotation

Emotion annotation: data in literary texts (from Santos & Simões (2019))

amor	107,203	desejo	72,242
feliz	61,541	infeliz	61,488
medo	38,473	gen	30,237
vergonha	28,135	orgulho	25,439
feliz & satisfeito	22,889	coragem	21,117
surpresa	20,637	humildade	20,282
odio	19,502	esperanca	19,259
furia	14,943	satisfeito	14,651
desespero	14,601	saudade	13,037

General issues about the emotion annotation project

- Document what the Portuguese language considers as an emotion
- Get reliable criteria for considering that a particular word is being used with an emotional meaning
- Get clear criteria for assigning a particular word occurrence to a particular cluster (emotion group)
- Decide how many and which groups should the annotation cater for

All these are complex questions, and accept many answers.

Can we gain anything from crossing the two projects?

There are several reasons to embark in this line of action

- Emotion is inextricably connected with the body
 - Emotions have often a bodily counterpart, like change in temperature or heart rate
 - We perceive emotion in others by their body movements, expressions, postures and changes in physical qualities
 - Often we express emotions with our body, like smiling, crying or shouting
- All (?) cultures associate some parts of the body with emotions and other faculties
- We had already recognized two kinds of body expressions with feelings or opinions in *Esqueleto*
- In the emotion annotation we were wondering about how to treat body expressions (*dor de cotovelo*) as well as body conditions that expressed emotions (*chorar*)

In addition to see whether it would be beneficial to somehow merge the two annotations, as we have done before with the annotation of the *dizer* verbs and the the emotion realm, some of the specific questions we wanted to elucidate with this study were

- 1 Are the emotions expressed by body parts or expressions easy to classify in our emotion grid, or are there specific clusters or groups that are highly connecte with the body and belonging to possibly other groups?
- 2 Are (some, all) emotions more associated with parts of the body, in terms of co-occurrence in the text?
- 3 More specifically, would a geographical information system (GIS)-inspired visualization help the understanding of the complex relationships between body and emotion in Portuguese?

- 1 When we tried to assign the body expressions that had been classified as emotions, we were able to assign 70% to the existing clusters.

estar de pé atrás, ficar com a pulga atrás da orelha

But for the body expressions which had been classified as opinions (and which is often called sentiment in NLP), we managed only 25%.

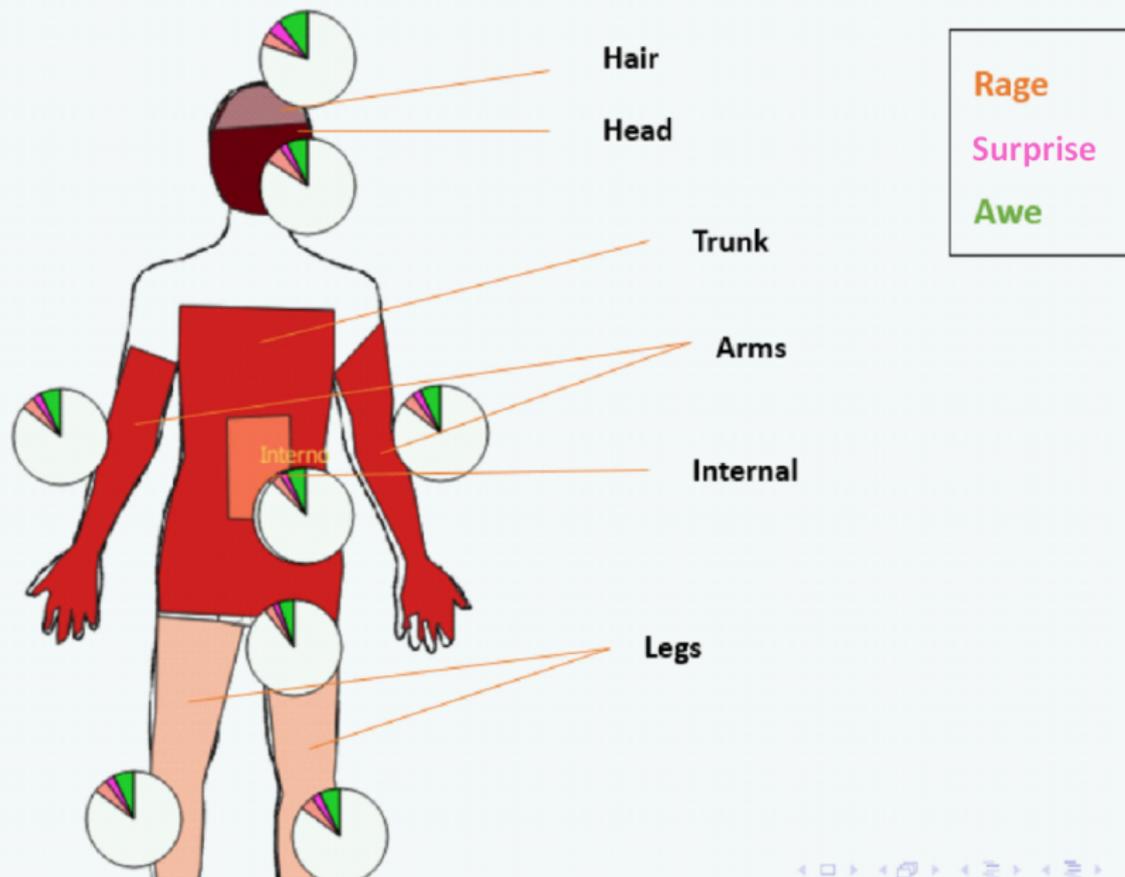
ver com bons olhos, cabeça dura

Could this be a good argument to actually distinguish between the two issues, evaluation and feeling? Or would this be better handled adding a group of evaluation as emotion?

(Preliminary) answers (cont'd)

- 1 Preliminary counts seem to indicate that, if we add up all cases by body part (Head, Trunk, Arms, Legs, Internal and Hair) that have been interpreted literally, the head is the part of the body with the highest co-occurrence of emotion description, followed by the hair, and then the trunk and the arms. This is in contrast (or complementary) with the fact that often body parts are interpreted as feelings.
- 2 We believe to have been the first to provide a bodymap by using GIS technology for visualizing the body, graduated by number of occurrences. We hope this may be useful as another way to present results, although we lose in the next figure the possibility to see which specific cases of emotion words or expressions are more connected to the feelings.

A body map for three emotions: rage, surprise and awe



Concluding remarks

- Initial results seem to be promising: it makes sense to look at both domains (body and emotions) together
- A wide array of different alleys can be followed: for example, create distorted maps per emotion, or, for a specific group of emotions, see if they distribute differently around the body
- The visualization we proposed is just the initial, plainest case. Other kinds of bodies (and in particular the feminine, the masculine and the child bodies) should be investigated and mapped separately
- We hope to have shown that distant reading the body and the emotions, thanks to annotated corpora, is worth while to further understand language