

“Own” work with emotions

Diana Santos

ILOS

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

June 2015, day 5



Why am I here?

- I am responsible for a large infrastructure for Portuguese corpus processing, AC/DC and Gramateca, both under the largest venue of Linguateca, where, in addition to other semantic fields, we decided to annotate emotions
- Also the contrastive (cultural and grammatical) issues are extremely important in my teaching and research activity, and I work in several parallel corpora

Note: Most if not all the work here is done in collaboration with my colleagues in Linguateca.

Annotate with emotions?

What does this mean? Note that we are talking here of 1.3 billion words of several genres.

- decide on which families (group of emotions)
- decide on which members in each family
- decide what to do with ambiguous words, and with vague words

We used roughly Belinda's classes

- We used Belinda's classes, but also added some others, due to disagreement or due to specific interest in that area.
- Then we used several thesauri for getting members of each family, as well as (as an annotator marking those which were obviously ambiguous.
- Then we automatically annotated all instances of those lexemes with the label of either **emomax** or **emomin**. See Mota & Santos (2015) for details.
- And (for some specific cases, in some corpora) created some rules for disambiguating emotion.

List of emotion families

19 cases

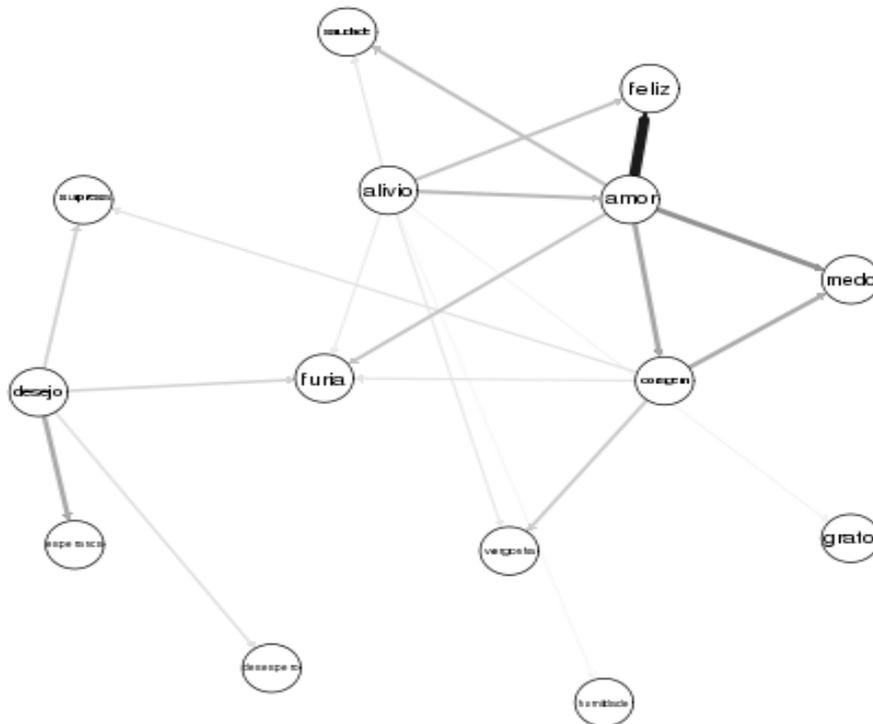
alívio	relief	admirar	admiration
amor	love	coragem	courage
desejo	desire	desespero	despair
esperança	hope	feliz	happy
fúria	anger	grato	grateful
infeliz	unhappy	humildade	humbleness
insatisfeito	insatisfaction	medo	fear
orgulho	pride	satisfeito	content
saudade	missing, nostalgia	surpresa	surprise
vergonha	shame		

Then we were able to study some cases

- Fear and courage: Maia & Santos (2013) at ICAME
- Surprise vs. admiration: Santos & Mota (2015)
- Emotion per genre and passives of emotion verbs in Santos (2015)
- Correlation between polarity annotated in (book) reviews and emotion words, Santos et. al. (2015)

New ideas

Emotion network: which emotion families co-occur in text? (In vercial)



Emotion infrastructure

Simply measuring the co-occurrence of the families may mask issues of frequency and others. The advantage of having other annotations in the corpora is that one can also

- crossreference with tense
- crossreference with passive
- crossreference with person
- crossreference with subject and target
- crossreference with genre
- crossreference with body and colour

Emotion infrastructure: examples

- Which cases associated with body are labeled “emotion” or “opinion”?
corpo:sentimento meu coração... de cortar o coração, ódio no coração, trazia nos olhos, três pedras na mão, cara fechada, sangue ferve, arejar a cabeça
corpo:opiniaio cara de..., desmiolada, bater palmas
corpo:atitude rojar-se aos pés, encolher os ombros
- Which cases have both body and emotion in a sentence? Examples next slide

Emotion infrastructure: examples og body and emotion

sorriso no rosto
felicidade estampada nos rostos
cabelos em fúria
coragem de dar as mãos
agitada, olhos arregalados
rancor irradiava-lhe dos olhos
de pé, terrível, pálido da raiva
sentir a adrenalina no sangue
não querem mais derramamentos de sangue
a melancolia da separação dera-lhe
cabelos brancos
Agressivamente um dos vendedores
resmungou entre dentes
mal seguro nas pernas, que vergavam
sob o peso da ira

smile in the face
happiness written in the face
hair in anger
courage to give hands
frenetic, with wide-opened eyes
contempt emanated from his eyes
standing, awful, pale of hate
feel adrenalin in the blood
do not want no more blood sheddings
sadness of separation turned his hair
white
aggressively mumbling with gritted
teeth
hardly on his legs, falling with the
hate weight

Concluding remarks

- infrastructure: what does this mean?
- lots of things to investigate: endless exploratory studies
- does it pay? Or precisely the emotions/feelings are lacking in (written) corpora?
- how consensual are the emotions chosen? (in the sense of allowing for different applications)
- how to evaluate independently the value of these large annotation projects? How to define a user-visible task (to be evaluated) instead of user-transparent ones, in terms of Gaizauskas (1998)?